Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 1–10.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 5–16.

ATOMS, MOLECULES, OPTICS

Temporal Self-Action and Compression of Intense Ultrashort Laser Pulses in Hollow Photonic-Crystal Waveguides

A. D. Bessonov, E. E. Serebryannikov, and A. M. Zheltikov

Physics Department, Lomonosov Moscow State University, Vorob'evy gory, Moscow, 119992 Russia
International Laser Center, Lomonosov Moscow State University, Vorob'evy gory, Moscow, 119992 Russia
e-mail: zheltikov@top.phys.msu.ru

Received June 28, 2004

Abstract—Hollow-core waveguides with a periodic (photonic-crystal) cladding are shown to allow efficient temporal compression of high-intensity ultrashort laser pulses and formation of megawatt soliton-like features in the regime of robust isolated guided modes. We numerically analyze the temporal envelope evolution and spectral transformation of the light field in air-guided modes of gas-filled hollow coaxial periodic Bragg waveguides. Based on this analysis, we define optimal compression regimes, permitting high compression ratios (of about six) and high compression efficiencies (up to 73%) to be achieved for microjoule laser pulses with an initial pulse length of 80–400 fs. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Temporal self-action of ultrashort laser pulses is one of the most interesting phenomena in ultrafast nonlinear optics. Self-phase modulation (SPM) [1, 2], related to the light-induced, intensity-dependent change in the refractive index of a nonlinear material, allows extremely short pulses to be produced and the phase of laser radiation to be controlled [3]. This effect also plays an important role in the formation of optical solitons [4–6]. The common strategy of short-pulse compression involves SPM-induced spectral broadening of a laser pulse within a large propagation length (typically, in an optical fiber) with subsequent chirp compensation using prism compressors, diffraction gratings, or chirped mirrors [3].

The laser power transmitted through an optical fiber is limited by self-focusing [1], which leads to an optical damage of the fiber above a certain critical laser power.

Hollow waveguides [7, 8] represent a powerful and convenient tool for the transmission and nonlinear-optical transformation of high-power laser pulses. The threshold of optical breakdown for gases filling the core of such waveguides is much higher than typical breakdown thresholds for dielectrics, with the radiation fluence on waveguide walls usually being several orders of magnitude lower than the radiation fluence at the center of the waveguide core. Due to this advantageous combination of properties, hollow waveguides made it possible to perform several interesting and important experiments dealing with the physics of high-intensity ultrashort laser pulses [9]. Hollow-core fibers are intensely used, in particular, in modern laser systems to increase the length of nonlinear-optical interactions of laser pulses and to enhance nonlinear-optical processes. Fibers of this type allow high-intensity ultrashort laser pulses to be spectrally transformed through nonlinear-optical processes without a laser breakdown in the fiber core. Self-phase modulation in a gas filling the core of a hollow fiber makes it possible to produce pulses shorter than 5 fs [10, 11]. Stimulated Raman scattering (SRS) of laser pulses in hollow fibers filled with Raman-active gases results in an efficient generation of multiple Raman sidebands. In the regime of locked phases, these Raman sidebands can be employed to synthesize pulses shorter than 4 fs [12]. Hollow waveguides can radically enhance high-order harmonic generation [13–16] and improve the sensitivity of gas-phase analysis based on four-wave mixing (FWM) spectroscopy [17–19].

The modes of standard hollow fibers with a solid dielectric cladding are leaky [7], with the magnitude of optical losses increasing for these modes as λ^2/a^3 with a decrease in the radius a of the hollow core (λ is the radiation wavelength). There is no way, therefore, to use standard hollow fibers with very small inner diameters for laser experiments, which usually operate with hollow fibers with core diameters ranging from 100 up to 500 μm. Such fibers are essentially multimode. The losses of guided modes in standard, solid-cladding hollow fibers with smaller core diameters are typically unacceptably high for the transmission and nonlinear-optical transformations of laser pulses. This circumstance prevents the above-described strategies of pulse compression, based on standard hollow-core fibers,

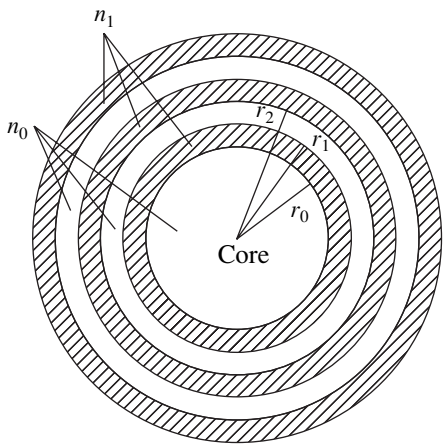

Fig. 1. A sketch of the cross section of a coaxial periodic waveguide. Unshaded and shaded areas correspond to materials with low (n_0) and high (n_1) refractive indices, respectively.

from being extended to femtosecond pulses with energies ranging from a few up to tens of microjoules.

Recently developed hollow-core photonic-crystal fibers (PCFs) [20, 21] provide a unique opportunity to transmit high-intensity laser fields in guided modes with transverse sizes of 10–50 μm. These fibers can confine waveguide modes of electromagnetic radiation within the low-index area of the hollow core due to the high reflectivity of a two-dimensionally periodic (photonic-crystal) cladding within photonic bandgaps (PBGs) [20–26]. In these frequency ranges, electromagnetic field cannot exist in the form of waves propagating inside the periodic structure of the photonic-crystal cladding. The reflection coefficient of a periodic structure within PBGs is much higher than the reflection coefficient of the material of the cladding, substantially reducing optical losses of air-guided modes in hollow fibers. PBGs of the PCF cladding are thus mapped onto passbands in fiber transmission.

Due to the high intensities of laser pulses attainable in the hollow core of PCFs without an optical breakdown of the fiber and because of the large interaction length provided by the waveguide geometry, hollow PCFs can radically enhance nonlinear-optical processes, including stimulated Raman scattering [27], four-wave mixing [28, 29], and self-phase modulation [30, 31]. As demonstrated by experiments [32, 33], hollow PCFs can transmit ultrashort laser pulses in the regime of temporal solitons. Such fibers can be employed for the laser guiding of microspecies and atoms [34], creation of optical switches and limiters for high-intensity laser pulses [35], and transportation of high-energy laser pulses for technological [36, 37] and biomedical [38] applications.

In this work, we will show that hollow PCFs allow an efficient temporal compression of high-intensity ultrashort laser pulses, as well as formation of strong-field solitonlike features in the regime of robust isolated guided modes. We will perform a numerical analysis of temporal envelope evolution and the spectral transformation of the light field in air-guided modes of gas-filled hollow coaxial periodic Bragg waveguides. Based on this analysis, we will identify optimal compression regimes, permitting high compression ratios (of about six) and high compression efficiencies (up to 73%) to be achieved for microjoule laser pulses with an initial pulse length of 80–400 fs.

2. MODEL OF A PERIODIC COAXIAL WAVEGUIDE

For a qualitative analysis of temporal self-action of femtosecond pulses in hollow PCFs, we employ a model of a periodic coaxial Bragg waveguide. The cladding in such a waveguide consists of a stack of coaxial cylinders with a periodically alternating refractive index. Physically, the mechanism behind guided-mode formation in coaxial Bragg waveguides is in many respects similar to the mechanism of waveguiding in hollow-core PCFs, as electromagnetic radiation is confined to the low-index hollow core in both cases due to PBGs of the fiber cladding. The modes of coaxial Bragg waveguides have been studied in earlier works [39–43]. In recent years, this effort was, at least partially, motivated by the fabrication and successful demonstration of dielectric coaxial Bragg waveguides [44]. Obviously, the model of a coaxial Bragg waveguide cannot provide an accurate quantitative description of guided modes in hollow PCFs. However, this model allows the basic features of dispersion properties and transmission spectra of hollow PCFs to be understood in a simple and illustrative way, providing also a general insight into the properties of field intensity profiles in waveguide modes localized in a hollow core of a PCF [45].

The cross-sectional structure of a hollow-core coaxial Bragg waveguide is sketched in Fig. 1. The waveguide has a hollow core with a radius a surrounded by a stack of coaxial cylinders with alternating refractive indices and thicknesses b and c. The refractive index is assumed to remain constant within the core and each of the cladding layers. Equations for the eigenmodes of such a waveguide are written as

$$[\Delta_\perp + k^2 n_i^2]\mathbf{E}_i = \beta^2 \mathbf{E}_i, \quad (1)$$

$$[\Delta_\perp + k^2 n_i^2]\mathbf{H}_i = \beta^2 \mathbf{H}_i, \quad (2)$$

where n_i is the refractive index of the ith region; k is the wavenumber; β is the propagation constant; and $\mathbf{E}_i$ and $\mathbf{H}_i$ are the electric and magnetic fields in the ith region, respectively.

The electric and magnetic fields are coupled by the boundary conditions:

$$\mathbf{E}_i^\tau(r_i, \theta, \omega) = \mathbf{E}_{i+1}^\tau(r_i, \theta, \omega), \quad (3)$$

$$\mathbf{H}_i^\tau(r_i, \theta, \omega) = \mathbf{H}_{i+1}^\tau(r_i, \theta, \omega), \quad (4)$$

where ω is the frequency; θ is the angle in the cylindrical frame of reference; and r_i is the radius of the circular interface between the ith and $(i + 1)$th regions,

$$r_i = a + i(b + c) \text{ for even } i,$$

$$r_i = a + b + (i - 1)(b + c) \text{ for odd } i.$$

The longitudinal components of the electric and magnetic fields in the ith layer of the waveguide are written as [40–42]

$$E_{iz}(r, \theta) = (A_i J_m(u_i r) + B_i N_m(u_i r))\cos(m\theta), \quad (5)$$

$$H_{iz}(r, \theta) = (C_i J_m(u_i r) + D_i N_m(u_i r))\sin(m\theta), \quad (6)$$

where J_m and N_m are the Bessel functions of the first and second kind; A_i, B_i, C_i, and D_i are the coefficients related to each other by the relevant boundary conditions; and u_i is the transverse part of the propagation constant for the ith waveguide mode.

The boundary conditions of Eqs. (3) and (4) can be represented in the form of matrix equations for unknown coefficients:

$$T_i(r_i)U_i = T_{i+1}(r_i)U_{i+1}, \quad (7)$$

where

$$U_i = (A_i B_i C_i D_i), \quad (8)$$

and

$$T_i(r) = \begin{pmatrix} J_m(u_i r) & N_m(u_i r) & 0 & 0 \\ 0 & 0 & J_m(u_i r) & N_m(u_i r) \\ -\frac{\beta m}{u_i^2 r} J_m(u_i r) & -\frac{\beta m}{u_i^2 r} N_m(u_i r) & \frac{\omega\mu_0}{u_i} J'_m(u_i r) & \frac{\omega\mu_0}{u_i} N'_m(u_i r) \\ \frac{n_i^2 \varepsilon_0 \omega}{u_i} J'_m(u_i r) & \frac{n_i^2 \varepsilon_0 \omega}{u_i} N'_m(u_i r) & -\frac{\beta m}{u_i^2 r} J_m(u_i r) & -\frac{\beta m}{u_i^2 r} N_m(u_i r) \end{pmatrix}. \quad (9)$$

The boundary conditions (3) and (4) should be supplemented by the requirement that the field should remain finite for $r = 0$, i.e., $B_0 = 0$ and $D_0 = 0$. To complete the set of equations for unknown coefficients, we also introduce a boundary condition for the field on the external boundary of the outer layer. The form of the field along this contour is physically insignificant because the field amplitude is small on this interface. The magnitude of losses for the considered type of waveguide can be found through the calculation of the coefficient of reflection from the periodic cladding:

$$R = 1 - \frac{(A_N)^2 + (C_N)^2}{(A_0)^2 + (C_0)^2}. \quad (10)$$

The attenuation coefficient can then be derived by using the ray approach to the description of waveguide modes [45, 46]:

$$\tan\varphi = \frac{u_0}{\beta}, \quad (11)$$

$$N_R = \frac{\tan\varphi}{4a}, \quad (12)$$

$$\gamma = -N_R \ln R. \quad (13)$$

Here, φ is the angle between the ray representing the guided mode and the z axis, N_R is the number of reflections of the ray from waveguide walls per unit length, and γ is the sought-for attenuation coefficient. Formulas (11)–(13) define the losses of guided modes in a hollow-core waveguide with a periodic cladding.

3. EVOLUTION OF ULTRASHORT PULSES

In this section, we present a model of the temporal self-action of ultrashort laser pulses in a hollow waveguide. Analysis of the evolution of high-power ultrashort field waveforms in a hollow waveguide is a complex nonlinear problem. In the case of very short pulse lengths, the standard slowly varying envelope approximation (SVEA) becomes inapplicable. To introduce our approach, we start with the generic equation for the field $\mathbf{E}(\mathbf{r}, \omega)$ propagating in a nonlinear medium [1]:

$$\begin{aligned} &\mathrm{rot}[\mathrm{rot}\mathbf{E}(\mathbf{r}, \omega)] \\ &= k^2 n^2 \mathbf{E}(\mathbf{r}, \omega) + \mu_0 \omega^2 \mathbf{P}_{nl}(\mathbf{r}, \omega), \end{aligned} \quad (14)$$

where μ_0 is the magnetic permeability and $\mathbf{P}_{nl}(\mathbf{r}, \omega)$ is the nonlinear part of the polarization of the medium.

We split the differential operator appearing in Eq. (14) into the longitudinal and transverse parts and separate variables assuming the excitation of only one waveguide mode:

$$\mathbf{E}(\mathbf{r}, \omega) = E(z, \omega)\mathbf{E}(x, y, \omega), \quad (15)$$

$$\frac{\partial^2 \mathbf{E}(\mathbf{r}, \omega)}{\partial z^2} + \beta^2 \mathbf{E}(\mathbf{r}, \omega) + \mu_0 \omega^2 \mathbf{P}_{nl}(\mathbf{r}, \omega) = 0. \quad (16)$$

Using the solution to the linear problem in the nonlinear regime in the case under consideration, we, in fact, neglect spatial self-action. Assuming that the nonlinear polarization is parallel to the applied field, we derive

$$\frac{\partial^2 \mathbf{E}(\mathbf{r}, \omega)}{\partial z^2} + \beta_{nl}^2(z, \omega)\mathbf{E}(\mathbf{r}, \omega) = 0, \quad (17)$$

$$\beta_{nl}^2(z, \omega) = \beta^2(\omega) + \mu_0 \omega^2 B_{nl}(z, \omega), \quad (18)$$

where β_{nl} is the nonlinear propagation constant and

$$\mathbf{P}_{nl}(\mathbf{r}, \omega) = \mathbf{E}(\mathbf{r}, \omega) B_{nl}(z, \omega).$$

The differential operator acting on the Fourier transform of the field in Eq. (17) can be represented as

$$\frac{\partial^2}{\partial z^2} + \beta_{nl}^2 = \left(\frac{\partial}{\partial z} - i\beta_{nl}\right)\left(\frac{\partial}{\partial z} + i\beta_{nl}\right) - i\frac{\partial \beta_{nl}}{\partial z}. \quad (19)$$

The identity (19) shows that the solution to Eq. (17) in the linear case is given by a sum of counterpropagating waves:

$$E(z, \omega) = E^+(z, \omega) + E^-(z, \omega).$$

We are interested in the forward wave,

$$E^+(z, \omega) \propto E_0(\omega)\exp(i\beta_{nl}(\omega, z)z).$$

The backward wave is induced by the nonlinear, intensity-dependent modulation of the refractive index along the z axis, giving rise to the reflection of the forward wave. This effect is represented by the last term on the right-hand side of Eq. (19). In the regime when

$$\frac{\partial \beta_{nl}}{\partial z} \ll \beta_{nl}^2,$$

this term is negligible and effects related to the reflected wave can be ignored. In this case, Eq. (17) is reduced to

$$\frac{\partial \mathbf{E}(\mathbf{r}, \omega)}{\partial z} = i\beta_{nl}(z, \omega)\mathbf{E}(\mathbf{r}, \omega). \quad (20)$$

We now expand β_{nl} as a Taylor series including only the first two terms in this expansion:

$$\beta_{nl} \approx \beta + \frac{\mu_0 \omega^2}{2\beta} B_{nl}. \quad (21)$$

Our equation for the field then reads

$$\frac{\partial \mathbf{E}(\mathbf{r}, \omega)}{\partial z} = i\beta(z, \omega)\mathbf{E}(\mathbf{r}, \omega) + i\frac{\mu_0 \omega^2}{2\beta(z, \omega)}\mathbf{P}_{nl}(\mathbf{r}, \omega). \quad (22)$$

Introducing the retarded time

$$\eta = t - z/V_g,$$

where V_g is the group velocity for the central wavelength of the pulse, we rewrite Eq. (22) as

$$\frac{\partial \mathbf{E}(\mathbf{r}, \omega)}{\partial z} = i\left(\beta(z, \omega) - \frac{\omega}{V_g}\right)\mathbf{E}(\mathbf{r}, \omega) + i\frac{\mu_0 \omega^2}{2\beta(z, \omega)}\mathbf{P}_{nl}(\mathbf{r}, \omega). \quad (23)$$

For ultrashort pulses with broad spectra, the SVEA approach becomes inadequate. We describe the nonlinear polarization in the approximation of an instantaneous, nondispersive third-order nonlinear response at the frequency of the incident field, neglecting third-harmonic generation:

$$\mathbf{P}_{nl}(\mathbf{r}, \omega) = \mathrm{FT}\{\mathbf{P}_{nl}(\mathbf{r}, \eta)\}, \quad (24)$$

$$\mathbf{P}_{nl}(\mathbf{r}, \eta) = \varepsilon_0 \chi^{(3)} |\mathbf{E}(\mathbf{r}, \eta)|^2 \mathbf{E}(\mathbf{r}, \eta), \quad (25)$$

$$\mathbf{E}(\mathbf{r}, \eta) = \mathrm{FT}^{-1}\{\mathbf{E}(\mathbf{r}, \eta)\}, \quad (26)$$

where $\chi^{(3)}$ is the cubic nonlinear susceptibility of the medium related to the Kerr nonlinearity and FT{ } stands for the Fourier transform.

Multiplying both parts of Eq. (23) by $\mathbf{E}(x, y, \omega)$ and performing integration over the cross section of the waveguide, we arrive at the following one-dimensional equation:

$$\frac{\partial E(z, \omega)}{\partial z} = i\left(\beta(z, \omega) - \frac{\omega}{V_g}\right)E(z, \omega) + i\frac{\mu_0 \omega^2}{2\beta(z, \omega)}P_{nl}(z, \omega), \quad (27)$$

where

$$P_{nl}(z, \omega) = \alpha\varepsilon_0 \mathrm{FT}\{|E(z, \eta)|^2 E(z, \eta)\}, \quad (28)$$

$$E(z, \eta) = \mathrm{FT}^{-1}\{E(z, \omega)\}, \quad (29)$$

$$\alpha = \frac{\iint \chi^{(3)}(x, y)|\mathbf{E}(x, y, \omega_0)|^4 dS}{\iint |\mathbf{E}(x, y, \omega_0)|^2 dS}. \quad (30)$$

Equation (27) adapts the Maxwell equations to the self-phase modulation of broadband field waveforms,

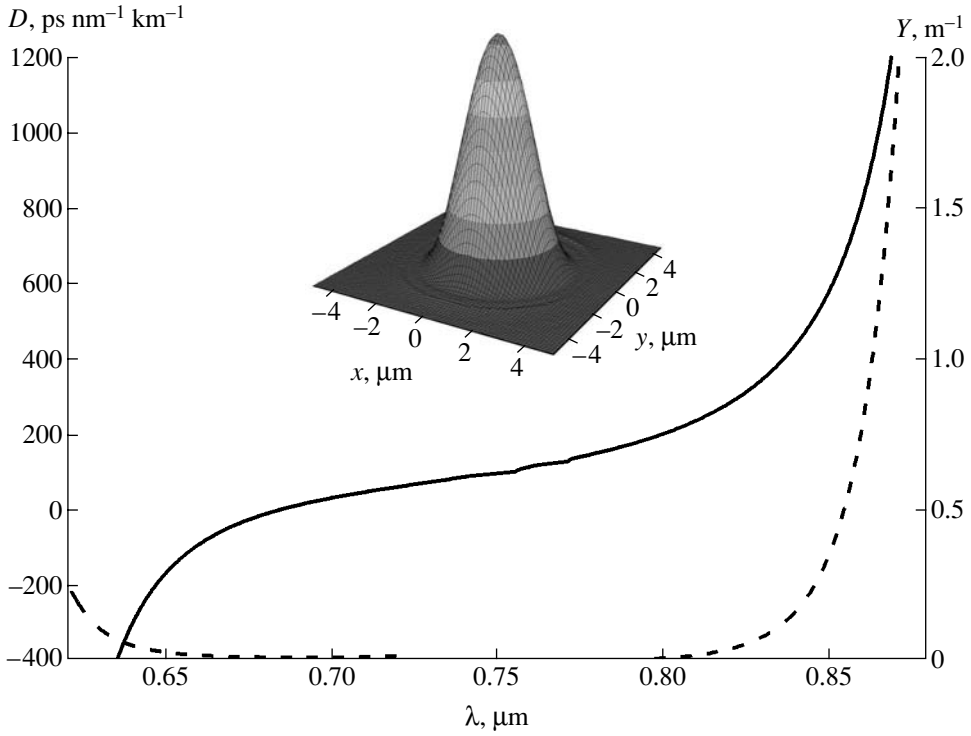

Fig. 2. Group-velocity dispersion D (the solid line) and the coefficient of losses Y (the dashed line) for the fundamental mode of a hollow periodic coaxial Bragg waveguide with $a = 3.0$ μm, $b = 0.5$ μm, and $c = 1.0$ μm. The inset shows the transverse distribution of the field intensity in the fundamental mode of the waveguide.

making it possible to examine the temporal self-action of ultrashort laser pulses in hollow waveguides beyond the SVEA approach.

We now complete our model by adding a damping term into Eq. (27). This operation gives rise to an imaginary part of the propagation constant β. Along with waveguide losses, damping may include also material absorption. However, this material part of losses will be neglected in our analysis. The final form of our evolution equation is thus written as

$$\frac{\partial E(z,\omega)}{\partial z} = i\left(\beta(z,\omega) - \frac{\omega}{V_g}\right)E(z,\omega) + i\frac{\mu_0\omega^2}{2\beta(z,\omega)}P_{nl}(z,\omega) - \gamma(\omega)E(z,\omega). \quad (31)$$

Equations similar to Eq. (31), but derived in a slightly different fashion, have been earlier employed to analyze supercontinuum generation and four-wave mixing in microstructure fibers [47, 48]. We numerically solved Eq. (31) by using a standard finite-difference procedure based on the approximation of the electric field with a discrete function and replacement of the field derivative in the z-coordinate by the relevant finite difference. The initial condition was defined by the spectrum of the pulse at the input of the waveguide at $z = 0$ (see Figs. 7 and 8 in Section 5 below). Consumption of computation time for this procedure was dominated by the fast Fourier transform for nonlinear polarization at each step of integration.

4. PULSE SELF-COMPRESSION

Self-phase modulation in a gas-filled hollow waveguide results in a spectral broadening of a laser

Compression of microjoule femtosecond pulses in the fundamental mode of a hollow-core coaxial Bragg waveguide

Central wavelength, μm	0.865	0.850	0.840
Initial pulse length, fs	400	200	80
Initial energy, μJ	1.8	1.0	1.4
Group-velocity dispersion at the central wavelength, ps nm^{-1} km^{-1}	1010	582	440
Propagation distance, cm	20	12	3.0
Dispersion length, cm	12.7	5.7	1.24
Nonlinear length, cm	9.7	8.5	2.4
Minimal pulse length, fs	68	32	14
Energy within the central peak, mJ	608	601	1022
Energy fraction in side peaks	47%	35%	27%
Compression efficiency in pulse energy	34%	60%	73%

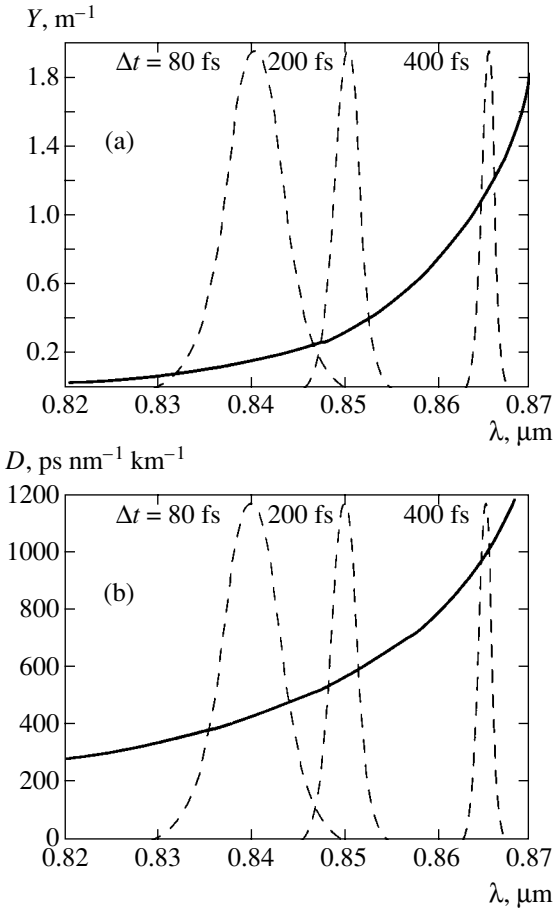

Fig. 3. Spectra of the input pulses (dashed lines) coupled into the fundamental mode of a hollow periodic coaxial Bragg waveguide with a = 3.0 μm, b = 0.5 μm, and c = 1.0 μm. The solid lines display the magnitude of losses (a) and the group-velocity dispersion (b) for the fundamental mode of this waveguide.

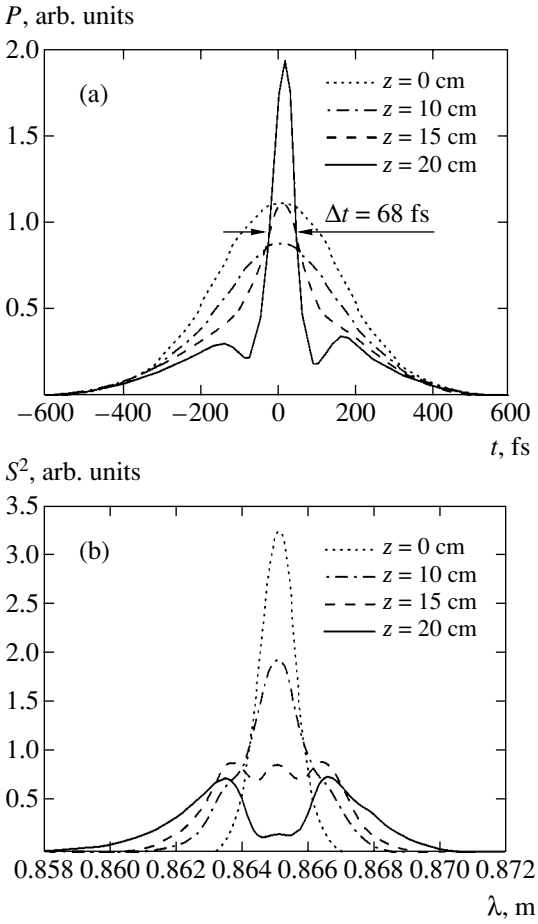

Fig. 4. Evolution of the temporal envelope (a) and the spectrum (b) of a femtosecond pulse in the fundamental mode of a hollow periodic coaxial Bragg waveguide with a = 3.0 μm, b = 0.5 μm, and c = 1.0 μm. The input pulse has a Gaussian temporal envelope with an initial pulse length of 400 fs, the central wavelength of 0.865 μm, and the energy of 1.8 μJ.

pulse propagating through the waveguide and puts a chirp on this pulse. Under certain conditions, the resulting chirp can be compensated by the waveguide dispersion. The elementary theory of self-phase modulation for narrowband pulses [1, 2] yields the following expression for the SPM-induced nonlinear phase shift:

$$\varphi(t, z) = -k_0 n_2 I(t) z, \tag{32}$$

where k_0 is the wavenumber for the central frequency and n_2 is the nonlinear refractive index of the medium. According to Eq. (32), the parameter of the SPM-induced chirp is given by

$$\alpha_{SPM}(t, z) = \frac{\partial^2 \varphi(t, z)}{\partial t^2} = -k_0 n_2 z \frac{\partial^2 I}{\partial t^2}. \tag{33}$$

On the other hand, solution of the linear evolution equation for a Gaussian pulse in the second order of dispersion theory yields the following expression for the parameter of the dispersion-induced chirp [49]:

$$\alpha_d(t, z) = z k_2^{-1} (z^2 + L_d^2)^{-1}, \tag{34}$$

where

$$k^2 = \frac{d^2 k}{d\omega^2},$$

L_d is the dispersion length.

Based on these simple arguments, one can expect that an anomalous group-velocity dispersion (GVD) would be needed to compensate for a chirp induced by SPM in a medium with a positive nonlinear refractive index.

For the fundamental air-guided mode of a hollow coaxial Bragg waveguide, the group-velocity dis-

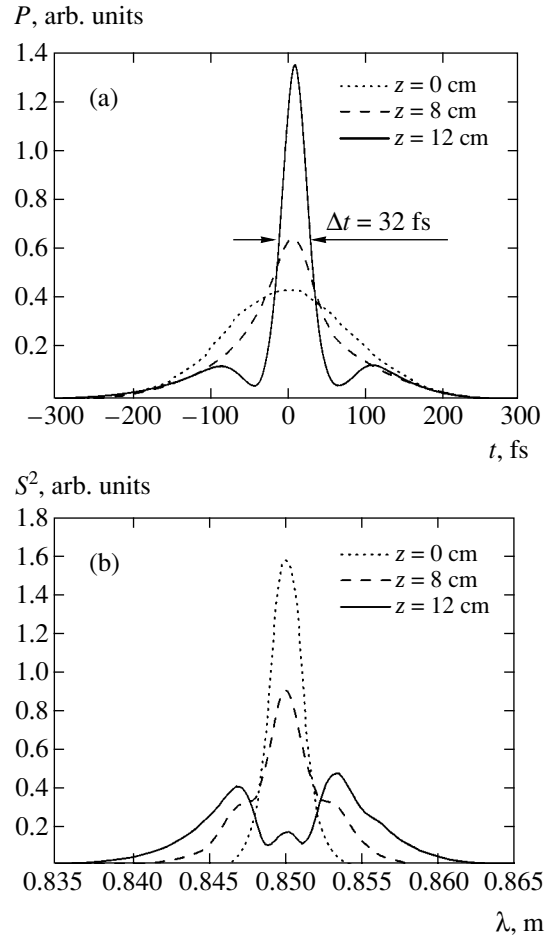

Fig. 5. Evolution of the temporal envelope (a) and the spectrum (b) of a femtosecond pulse in the fundamental mode of a hollow periodic coaxial Bragg waveguide with a = 3.0 μm, b = 0.5 μm, and c = 1.0 μm. The input pulse has a Gaussian temporal envelope with an initial pulse length of 200 fs, the central wavelength of 0.850 μm, and the energy of 1.0 μJ.

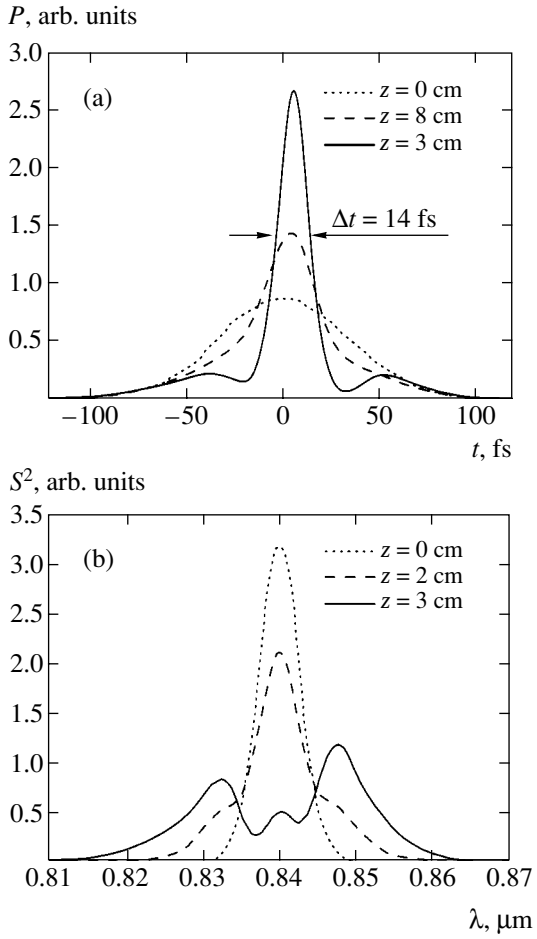

Fig. 6. Evolution of the temporal envelope (a) and the spectrum (b) of a femtosecond pulse in the fundamental mode of a hollow periodic coaxial Bragg waveguide with a = 3.0 μm, b = 0.5 μm, and c = 1.0 μm. The input pulse has a Gaussian temporal envelope with an initial pulse length of 80 fs, the central wavelength of 0.840 μm, and the energy of 1.4 μJ.

persion can take both positive and negative values. Figure 2 shows the group-velocity dispersion for the fundamental mode of a hollow coaxial Bragg fused silica waveguide filled with argon at a pressure of 1 atm. Geometric parameters of the waveguide are a = 3 μm, b = 0.5 μm, and c = 1.0 μm. The inset to Fig. 2 displays the transverse field intensity profile in the fundamental mode of the waveguide. Self-compression of laser pulses in a medium with a positive nonlinear refractive index requires anomalous group-velocity dispersion, which is achieved on the right-hand side of the GVD plot in Fig. 2.

Numerical simulations were performed for laser pulses with an initial pulse length ranging from 50 up to 500 fs. The initial pulse shape was assumed to be Gaussian. Based on the results of numerical simulations, we defined the optimal pulse energy and central frequency providing the maximum efficiency of pulse compression. Tuning the central frequency shifts the spectrum of the laser pulse relative to the GVD curve (Fig. 3), thus modifying the effective dispersion profile sensed by the most powerful components in the spectrum of the laser pulse.

Optimization on the space of laser-pulse and waveguide parameters for the maximum efficiency of pulse compression is a quite complicated problem, involving a careful matching of the initial parameters of the laser pulse with the characteristics of the waveguide. The table presents three sets of initial parameters of microjoule femtosecond laser pulses allowing efficient pulse compression in a hollow coaxial Bragg waveguide. Figure 3 displays the initial spectra of these pulses against the GVD profile of the fundamental mode in the waveguide under study. The temporal and spectral evolution of the pulses with initial parameters specified in the table and in Fig. 3 is illustrated by Figs. 4–6. With realistic GVD profiles of waveguide modes, self-compression of the laser pulses

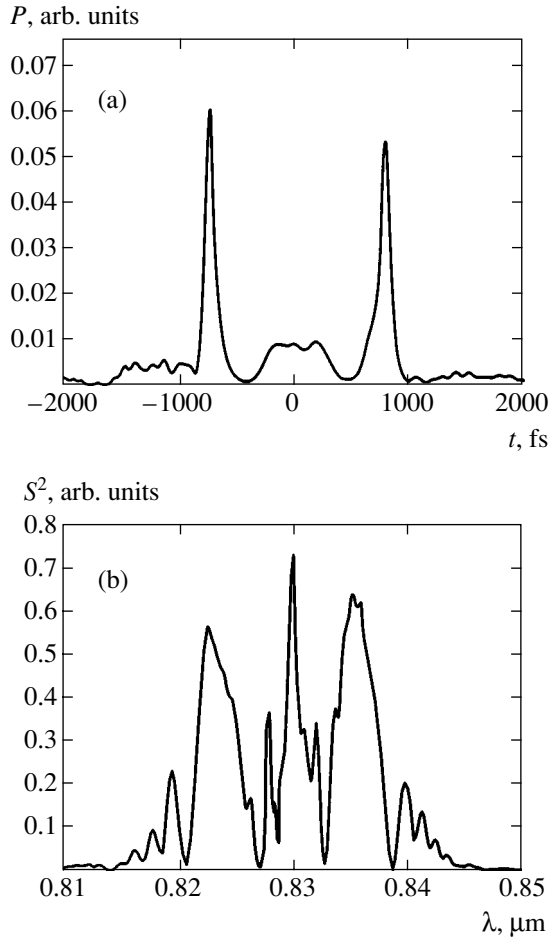

Fig. 7. Formation of megawatt quasi-solitonic features. The temporal envelope (a) and the spectrum (b) of a pulse in the quasi-soliton regime at a distance of 5 cm from the input end of the waveguide. The input pulse has a Gaussian temporal envelope with an initial pulse length of 100 fs, the central wavelength of 0.830 μm, and the energy of 0.8 μJ.

is accompanied by the appearance of off-center radiation peaks and other pulse-shape distortions. However, in all the three cases, the ratio of pulse compression is quite high (about six). The central peak of the pulse corresponding to the highest pulse compression ratio carries up to 73% of the input pulse energy (Fig. 6). These results demonstrate a unique potential of hollow PBG waveguides and PCFs for the compression of high-power ultrashort laser pulses.

5. SOLITON-LIKE FEATURES

In the previous sections, we considered the evolution of laser pulses within propagation distances comparable with dispersion and nonlinear lengths. On such a spatial scale, self-phase modulation leads to a substantial spectral broadening of laser pulses. Our numerical simulations show that, within larger propagation distances, light pulses transmitted through hollow coaxial Bragg waveguides tend to form quasi-stable features, similar to optical solitons. Examples of such soliton-like features are presented in Figs. 7 and 8.

Figure 8 illustrates a quasi-soliton propagation regime where a pulse with a duration of 115 fs and an energy of 320 nJ is transmitted over a distance L = 30 cm, which is many times larger than the dispersion length (1.6 cm) without dramatic changes in its spectrum or temporal envelope. The power of the light field carried by this solitary wave can be as high as 3 MW under the above-specified conditions. Such megawatt solitons in hollow PCFs have been recently demonstrated by Ouzounov *et al.* [32]. In view of substantial losses, it is hardly possible to produce true solitons in waveguide structures considered in this paper. A decrease in the pulse energy due to waveguide losses reduces SPM-induced spectral broadening, eventually leading to a breakup of solitonic features. However, the

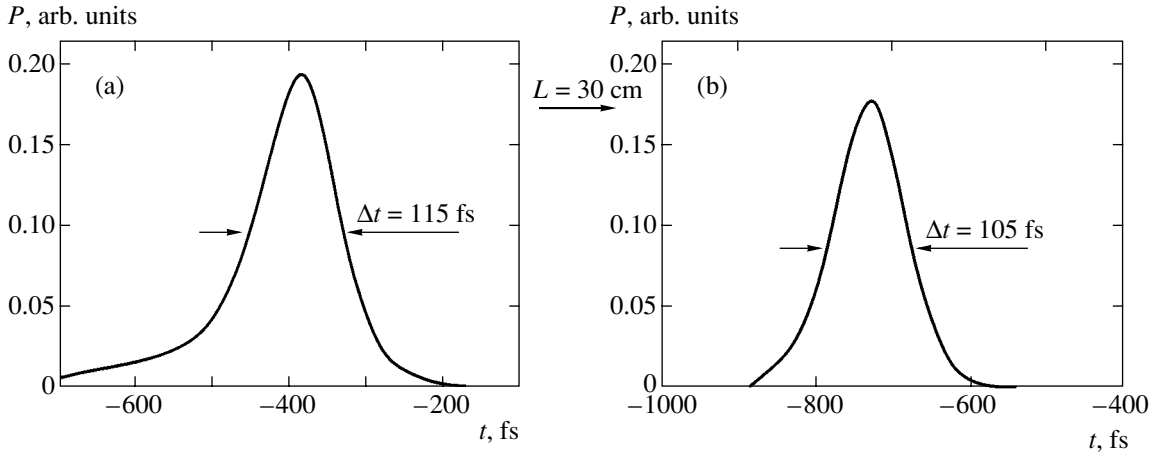

Fig. 8. Evolution of the temporal envelope of a laser pulse in the quasi-soliton regime in a hollow periodic coaxial Bragg waveguide with a = 3.0 μm, b = 0.5 μm, and c = 1.0 μm: (a) the input pulse and (b) the pulse transmitted over a distance L = 30 cm along the waveguide. The central radiation wavelength is 0.850 μm.

stability of the temporal pulse envelope within limited propagation lengths, demonstrated by our simulations, makes this regime of waveguiding promising for the transportation of high-power ultrashort laser pulses.

6. CONCLUSIONS

Numerical analysis performed in this work demonstrates that hollow-core waveguides with a periodic (photonic-crystal) cladding permit a highly efficient temporal compression of high-intensity ultrashort laser pulses and generation of megawatt solitonic features in robust isolated guided modes. Self-compression of ultrashort pulses in such waveguides is based on spectral broadening due to self-phase modulation and compensation of the resulting SPM-induced chirp by a specially designed dispersion profile of guided modes. This approach is shown to allow a sixfold compression of microjoule pulses with an initial pulse length of 80–400 fs. The central peak of the compressed pulse, corresponding to the maximum compression ratio, can carry up to 73% of the input pulse energy. The minimum duration of the compressed pulse at the output of a hollow PCF is mainly limited by the width of the spectral region where the GVD is anomalous. Within propagation lengths exceeding the dispersion and nonlinear lengths, intense ultrashort pulses can break up into several quasi-stable soliton-like features. The radiation power transmitted in such quasi-solitary waves can be as high as several megawatts. The rate of spectral broadening in such a quasi-soltonic regime of propagation is much lower than the rate of spectral broadening characteristic of the initial stage of pulse propagation in the waveguide.

Thus, hollow-core waveguides with a one- or two-dimensional photonic-crystal cladding offer unique possibilities for the compression of high-power femtosecond laser pulses. Due to low losses, the possibility of tailoring the group-velocity dispersion profile, the existence of stable isolated air-guided modes, and the possibility of concentrating the energy of the light field in guided modes with a small transverse size, these new waveguides can provide high efficiencies of pulse compression for both high- and medium-energy laser pulses.

ACKNOWLEDGMENTS

This study was supported in part by the President of Russian Federation Grant MD-42.2003.02, the Russian Foundation for Basic Research (project nos. 03-02-16929, 03-02-20002-BNTS, 04-02-39002-GFEN2004, and 04-02-81036-Bel2004), and INTAS (project nos. 03-51-5037 and 03-51-5288). The research described in this publication was made possible in part by Award no. RP2-2558 of the U.S. Civilian Research and Development Foundation for the Independent States of the Former Soviet Union (CRDF).

REFERENCES

1. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984; Nauka, Moscow, 1989).
2. P. N. Butcher and D. Cotter, *The Principles of Nonlinear Optics* (Cambridge Univ. Press, Cambridge, 1990).
3. T. Brabec and F. Krausz, Rev. Mod. Phys. **72**, 545 (2000).
4. G. P. Agrawal, *Nonlinear Fiber Optics* (Academic, Boston, 1989; Mir, Moscow, 1996).
5. A. Hasegawa and M. Matsumoto, *Optical Solitons in Fibers* (Springer, Berlin, 2003).
6. G. I. Stegeman and M. Segev, Science **286**, 1518 (1999).
7. E. A. J. Marcatili and R. A. Schmeltzer, Bell Syst. Tech. J. **43**, 1783 (1964).
8. M. J. Adams, *An Introduction to Optical Waveguides* (Wiley, New York, 1981; Mir, Moscow, 1984).
9. A. M. Zheltikov, Usp. Fiz. Nauk **172**, 743 (2002) [Phys. Usp. **45**, 687 (2002)].
10. M. Nisoli, S. De Silvestri, and O. Svelto, Appl. Phys. Lett. **68**, 2793 (1996).
11. M. Nisoli, S. De Silvestri, O. Svelto, *et al.*, Opt. Lett. **22**, 522 (1997).
12. N. Zhavoronkov and G. Korn, Phys. Rev. Lett. **88**, 203901 (2002).
13. A. Rundquist, C. G. Durfee III, Z. Chang, *et al.*, Science **280**, 1412 (1998).
14. E. Constant, D. Garzella, P. Breger, *et al.*, Phys. Rev. Lett. **82**, 1668 (1999).
15. C. G. Durfee III, A. R. Rundquist, S. Backus, *et al.*, Phys. Rev. Lett. **83**, 2187 (1999).
16. A. Paul, R. A. Bartels, R. Tobey, *et al.*, Nature **421**, 51 (2003).
17. R. B. Miles, G. Laufer, and G. C. Bjorklund, Appl. Phys. Lett. **30**, 417 (1977).
18. A. B. Fedotov, F. Giammanco, A. N. Naumov, *et al.*, Appl. Phys. B **72**, 575 (2001).
19. A. N. Naumov, F. Giammanco, D. A. Sidorov-Biryukov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 301 (2001) [JETP Lett. **73**, 263 (2001)].
20. R. F. Cregan, B. J. Mangan, J. C. Knight, *et al.*, Science **285**, 1537 (1999).
21. P. St. J. Russell, Science **299**, 358 (2003).
22. A. M. Zheltikov, *Optics of Microstructure Fibers* (Nauka, Moscow, 2004) [in Russian].
23. S. O. Konorov, A. B. Fedotov, O. A. Kolevatova, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 401 (2002) [JETP Lett. **76**, 341 (2002)].
24. G. Bouwmans, F. Luan, J. C. Knight, *et al.*, Opt. Express **11**, 1613 (2003).
25. J. C. Knight, Nature **424**, 847 (2003).
26. C. M. Smith, N. Venkataraman, M. T. Gallagher, *et al.*, Nature **424**, 657 (2003).
27. F. Benabid, J. C. Knight, G. Antonopoulos, and P. St. J. Russell, Science **298**, 399 (2002).
28. S. O. Konorov, A. B. Fedotov, and A. M. Zheltikov, Opt. Lett. **28**, 1448 (2003).

29. S. O. Konorov, E. E. Serebryannikov, A. A. Ivanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **79**, 499 (2004) [JETP Lett. **79**, 395 (2004)].
30. S. O. Konorov, D. A. Sidorov-Biryukov, I. Bugar, *et al.*, Appl. Phys. B **78**, 547 (2004).
31. S. O. Konorov, A. M. Zheltikov, Ping Zhou, *et al.*, Opt. Lett. **29**, 1521 (2004).
32. D. G. Ouzounov, F. R. Ahmad, D. Müller, *et al.*, Science **301**, 1702 (2003).
33. F. Luan, J. C. Knight, P. St. J. Russell, *et al.*, Opt. Express **12**, 835 (2004).
34. F. Benabid, J. C. Knight, and P. St. J. Russell, Opt. Express **10**, 1195 (2002).
35. S. O. Konorov, D. A. Sidorov-Biryukov, I. Bugar, *et al.*, Phys. Rev. A **70**, 023807 (2004).
36. S. O. Konorov, A. B. Fedotov, O. A. Kolevatova, *et al.*, J. Phys. D: Appl. Phys. **36**, 1375 (2003).
37. J. D. Shephard, J. D. C. Jones, D. P. Hand, *et al.*, Opt. Express **12**, 717 (2004).
38. S. O. Konorov, A. B. Fedotov, V. P. Mitrokhin, *et al.*, Appl. Opt. **43**, 2251 (2004).
39. P. Yeh, A. Yariv, and E. Marom, J. Opt. Soc. Am. **68**, 1196 (1978).
40. Yong Xu, R. K. Lee, and A. Yariv, Opt. Lett. **25**, 1756 (2000).
41. G. Ouyang, Yong Xu, and A. Yariv, Opt. Express **9**, 733 (2001).
42. T. Kawanishi and M. Izutsu, Opt. Express **7**, 10 (2000).
43. S. G. Johnson, M. Ibanescu, M. Skorobogatiy, *et al.*, Opt. Express **9**, 748 (2001).
44. M. Ibanescu, Y. Fink, S. Fan, *et al.*, Science **289**, 415 (2000).
45. S. O. Konorov, O. A. Kolevatova, A. B. Fedotov, *et al.*, Zh. Éksp. Teor. Fiz. **123**, 975 (2003) [JETP **96**, 857 (2003)].
46. A. Yariv and P. Yeh, *Optical Waves in Crystals: Propagation and Control of Laser Radiation* (Wiley, New York, 1984; Mir, Moscow, 1987).
47. A. V. Husakou and J. Herrmann, Phys. Rev. Lett. **87**, 203901 (2001).
48. A. V. Husakou and J. Herrmann, J. Opt. Soc. Am. B **19**, 2171 (2002).
49. S. A. Akhmanov, V. A. Vysloukh, and A. S. Chirkin, *The Optics of Femtosecond Laser Pulses* (Nauka, Moscow, 1988) [in Russian].

Translated by A. Zheltikov

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 107–119.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 125–138.

ELECTRONIC PROPERTIES OF SOLIDS

Special Features of the Electric Components of Acoustic Waves in the Vicinity of Nonpiezoactive Directions in Crystals

V. I. Al'shits[a], V. N. Lyubimov[a], and A. Radowicz[b]

[a]*Shubnikov Institute of Crystallography, Russian Academy of Sciences, Moscow, 119333 Russia*
[b]*Kielce Technical University, 25-314 Kielce, Poland*
e-mail: alshits@ns.crys.ras.ru
Received September 15, 2004

Abstract—The conditions of existence of the zero components of electric field **E** and electric induction **D** accompanying a volume acoustic wave propagating in a piezoelectric medium have been studied. General equations describing the positions of the zero-field lines $\mathbf{E}(\mathbf{m}) = 0$ and the zero-induction points $\mathbf{m}_0$, such that $\mathbf{D}(\mathbf{m}_0) = 0$ on the unit sphere ($\mathbf{m}^2 = 1$) of the wave propagation directions, are obtained. General theorems determining the conditions ensuring the existence of such lines and points, even in triclinic crystals, are formulated. The relationship between such directions and various elements of the crystal symmetry is analyzed. The vector fields $\mathbf{D}(\mathbf{m})$, which are always orthogonal to the wave normals **m**, in the vicinity of the zero-induction points $\mathbf{m}_0$ exhibit certain orientational singularities characterized by the Poincaré indices $n = 0, \pm 1, \pm 2$. The general analytical expressions are obtained for the n values in crystals with arbitrary anisotropy and specified for a number of crystals belonging to various symmetry classes. The conditions of stability of the orientational singularities with respect to small perturbations of the material moduli and a change in the crystal symmetry are considered. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

As is known, the acoustic wave of displacements in piezoelectric media is usually accompanied by a quasistatic wave of the electric potential. This implies that, using acoustic waves, electric signals can be transmitted at the velocity of sound over a crystal. This possibility opened the way to numerous applications of acoustic waves in electronic devices and even led to the formation of a special field of science called acoustoelectronics. The applied aspect provides an important stimulus for extensive investigations devoted to various features of acoustic fields in piezocrystals [1]. These investigations are also stimulated by basic interest in the study of new effects in media featuring interactions of electromechanical fields [2, 3]. The acoustics of piezoelectric crystals is still an extensively developing field of solid state physics (see, e.g., review [4]), the more so that even purely basic investigations in this field frequently contain ideas for fruitful, albeit not immediately evident, applications.

It should also be noted that by no means all basic problems in piezoacoustcs have been solved, especially for media with arbitrary anisotropy. The anisotropy often influences the properties of piezoelectric crystals in a nontrivial way, and may sometimes lead to qualitatively new phenomena. In particular, it is very important from the practical standpoint to know the wave propagation directions **m** for which the electric field components possess maximum amplitudes [5] and, on the contrary, to reveal the nonpiezoactive directions [1, 3] in which the electric signals are not transmitted. Taking into account that, irrespective of the anisotropy, the electric field in an acoustic wave is always longitudinal ($\mathbf{E} \parallel \mathbf{m}$) and the electric induction is always transverse ($\mathbf{D} \perp \mathbf{m}$), we have to distinguish [3] between the directions of longitudinal and transverse nonpiezoactivity in which $\mathbf{E} = 0$ and $\mathbf{D} = 0$, respectively, in any crystal. This paper presents the results of investigations aimed at a detailed analysis of the nonpiezoactivity of both types. Previously, only rather fragmentary data have been reported on these issues in the available literature.

Another important aspect of this problem is related to directions **m**, in the vicinity of which the vector fields of the amplitudes of displacements (**u**) and the accompanying electric components (**E**, **D**) exhibit singularities. According to the results obtained in [6, 7], this very situation takes place near the acoustic axes, where the orientational singularities in the degenerate branches of natural waves are observed for the **u** and **D** fields, and the amplitude singularities, for the **E** field. This paper deals with orientational singularities of a new type, which occur in the vicinity of the transverse nonpiezoactivity directions in the vector fields $\mathbf{D}(\mathbf{m})$, that is, around the points $\mathbf{m}_0$ on the unit sphere such that $\mathbf{D}(\mathbf{m}_0) = 0$.

Below we will formulate, proceeding from the general expressions describing the electric components $\mathbf{E}_\alpha$ and $\mathbf{D}_\alpha$ as functions of the direction of wave propagation in a crystal for all three branches of the acoustic spectrum ($\alpha = 1, 2, 3$), the equations determining spe-

cial directions $\mathbf{m}$ for which either $\mathbf{E}_\alpha(\mathbf{m}) = 0$ or $\mathbf{D}_\alpha(\mathbf{m}) = 0$. These directions have different dimensionalities: the typical solutions appear as zero-field lines ($\mathbf{E}_\alpha = 0$) and zero-induction points ($\mathbf{D}_\alpha = 0$) on the unit sphere $\mathbf{m}^2 = 1$. The equations obtained will be analyzed both in the general case and in application to various particular crystal symmetry classes. The two types on nonpiezoactivity are closely related to the crystal symmetry, but they can also exist in triclinic crystals possessing no elements of symmetry. The corresponding theorems of existence are proved.

The possible types of singularities in the vector field $\mathbf{D}_\alpha(\mathbf{m})$ in the vicinity of the transverse nonpiezoactivity directions will be considered. In particular, it will be shown that, depending on the material moduli, the singularity in an isolated point $\mathbf{m}_0$ may be characterized by the Poincaré indices (topological charges) $n = 0, \pm 1, \pm 2$. The general analytical expressions will be obtained for the n values in triclinic crystals with arbitrary anisotropy and specified for a large series of crystals belonging to particular crystal symmetry classes. Only the solutions corresponding to singularities with $n = \pm 1$ are topologically stable, while singularities of the other types either split or disappear upon an arbitrary triclinic perturbation of the material tensors. However, the sum of indices for any splitting must be equal to the initial index n.

2. FORMULATION OF THE PROBLEM AND SOME MAIN EQUATIONS

As was mentioned above, some purely mechanical characteristics, including the elastic displacement vector $\mathbf{u}(\mathbf{r}, t)$, the distortion tensor $\hat{\beta}(\mathbf{r}, t)$, and the stress tensor $\hat{\sigma}(\mathbf{r}, t)$, are related to such electrical quantities as the potential $\phi(\mathbf{r}, t)$ and the electric field strength $\mathbf{E}(\mathbf{r}, t)$, and induction $\mathbf{D}(\mathbf{r}, t)$. The fields of $\hat{\beta}(\mathbf{r}, t)$ and $\mathbf{E}(\mathbf{r}, t)$ can be expressed in terms of their own potentials as

$$\hat{\beta}(\mathbf{r}, t) = \nabla\mathbf{u}(\mathbf{r}, t), \quad \mathbf{E}(\mathbf{r}, t) = -\nabla\phi(\mathbf{r}, t). \tag{1}$$

The interrelation of these characteristics is explicitly expressed by the coupling equations [8]

$$\sigma_{ij} = c_{ijkl}\beta_{kl} + e_{kij}E_k, \quad D_i = e_{ikl}\beta_{kl} - \varepsilon_{ik}E_k, \tag{2}$$

where $c_{ijkl} = \hat{c}$ is the tensor of elastic moduli, $e_{kij} = \hat{e}$ is the tensor of piezoelectric moduli, and $\varepsilon_{ik} = \hat{\varepsilon}$ is the permittivity tensor. In such a piezoelectric medium, the volume acoustic wave with the wavevector $\mathbf{k} = k\mathbf{m}$ and the phase velocity v must be a superposition of mechanical and electric dynamic fields:

$$\{\mathbf{u}, \phi\} = \{\mathbf{u}_0, \phi_0\}\exp[ik(\mathbf{m}\cdot\mathbf{r} - vt)]. \tag{3}$$

These fields obey the usual equations of motion [8]

$$\mathrm{div}\hat{\sigma} = \rho\ddot{\mathbf{u}}, \quad \mathrm{div}\mathbf{D} = 0, \tag{4}$$

where ρ is the density of medium. Here, we use the well-known quasi-static approximation valid to within the terms proportional to the ratio $(v/c)^2 \sim 10^{-10}$ (c is the velocity of light). Combining the above relations, we readily obtain a homogeneous equation for the polarization vector $\mathbf{u}_0$ [3, 9]:

$$\hat{F}(v, \mathbf{m})\mathbf{u}_0 \equiv [\hat{F}^{(0)} + \mathbf{e}\otimes\mathbf{e}/\varepsilon]\mathbf{u}_0 = 0, \tag{5}$$

where

$$\hat{F}^{(0)} = \mathbf{m}\hat{c}\mathbf{m} - \rho v^2\hat{I}, \quad \mathbf{e} = \mathbf{m}\hat{e}\mathbf{m}, \quad \varepsilon = \mathbf{m}\cdot\hat{\varepsilon}\mathbf{m}, \tag{6}$$

symbol $\otimes$ denotes the dyadic product, and $\hat{I}$ is the unit matrix. A necessary condition for the existence of nontrivial solutions of the homogeneous equation (5) is

$$\det\hat{F}(v, \mathbf{m}) = 0. \tag{7}$$

This is a cubic equation for the square phase velocity (v^2), which determines the three branches of the velocity of the volume acoustic waves $v_\alpha(\mathbf{m})$ ($\alpha = 1, 2, 3$).

Orientations of the corresponding mutually orthogonal polarization vectors $\mathbf{u}_{0\alpha}(\mathbf{m})$ of the isonormal natural waves can be expressed in terms of the $\bar{\hat{F}}_\alpha$ matrix, which is adjoint to the matrix $\hat{F}_\alpha(\mathbf{m}) \equiv \hat{F}_\alpha(v_\alpha(\mathbf{m}), \mathbf{m})$ and is determined from the condition $\bar{\hat{F}}_\alpha\hat{F}_\alpha = \hat{I}\det\hat{F}_\alpha$. As can be readily checked, Eq. (5) for any vector $\mathbf{c}$ such that $\bar{\hat{F}}_\alpha\mathbf{c} \neq 0$ is satisfied for

$$\mathbf{u}_{0\alpha} \parallel \bar{\hat{F}}_\alpha(\mathbf{m})\mathbf{c}. \tag{8}$$

It should be emphasized that the direct relation (8) between the polarization vector $\mathbf{u}_{0\alpha}(\mathbf{m})$ and the wave normal $\mathbf{m}$ will be widely used in the subsequent analysis.

Once the field of elastic displacements for a given wave branch $\mathbf{u}_\alpha(\mathbf{r}, t)$ is known, we can also determine the corresponding electric components (see, e.g., [8]). For the subsequent analysis, these components are con-

veniently represented (by analogy with Eqs. (3)–(8)) in a coordinate-free form as

$$\phi_\alpha = \mathbf{e}\cdot\mathbf{u}_\alpha/\varepsilon, \quad \mathbf{E}_\alpha = -ik\phi_\alpha\mathbf{m}, \quad \mathbf{D}_\alpha = \hat{N}\mathbf{u}_\alpha, \quad (9)$$

$$\hat{N} = \hat{e}\mathbf{m} - (\hat{\varepsilon}\mathbf{m})\otimes\mathbf{m}\hat{e}\mathbf{m}/\mathbf{m}\hat{\varepsilon}\mathbf{m}. \quad (10)$$

Relations (9) together with condition (8) essentially determine the functions $\mathbf{E}_\alpha(\mathbf{m})$ and $\mathbf{D}_\alpha(\mathbf{m})$ necessary for the subsequent analysis.

As can be readily seen, $\mathbf{m}\hat{N} \equiv 0$. This relation and the third relation in (9) clearly illustrate the well-known property (see Section 1) according to which the electric field $\mathbf{E}_\alpha(\mathbf{m}_0)$ is purely longitudinal, whereas the induction $\mathbf{D}_\alpha(\mathbf{m}_0)$ is purely transverse:

$$\mathbf{E}_\alpha \parallel \mathbf{m}, \quad \mathbf{D}_\alpha \perp \mathbf{m}. \quad (11)$$

On the other hand, the same identity $\mathbf{m}\hat{N} = 0$ implies one useful property of the $\hat{N}$ matrix:

$$\det\hat{N} = 0, \quad (12)$$

which indicates that this matrix is planar and, hence, can be represented as a sum of two dyads.

3. EXAMPLES OF TRANSVERSELY ISOTROPIC PIEZOELECTRICS

There are three groups of piezoelectrics which exhibit a transverse isotropy of their acoustic properties. These media belong to the following classes of symmetry [10]:

$$\infty 2, 622, \quad (13)$$

$$\infty m, 6mm \quad (14)$$

$$\infty, 6. \quad (15)$$

Owing to the transverse isotropy, the formulas presented below contain only the polar angle θ between the $\mathbf{m}$ vector and the z axis coinciding with the main axis of symmetry. Without loss of generality, we may proceed with the analysis upon selecting any cross section containing the main axis. Here, it is convenient to choose

$$\mathbf{m} = (m_1, 0, m_3) \equiv (\sin\theta, 0, \cos\theta). \quad (16)$$

In these coordinates, the $\hat{F}^{(0)}$ matrix in Eq. (6) for all the six classes of symmetry (13)–(15) have the same quasi-diagonal form [11]:

$$\hat{F}^{(0)} = \begin{Bmatrix} c_{11}m_1^2 + c_{44}m_3^2 - \rho v^2 & 0 & dm_1m_3 \\ 0 & c_{66}m_1^2 + c_{44}m_3^2 - \rho v^2 & 0 \\ dm_1m_3 & 0 & c_{44}m_1^2 + c_{33}m_3^2 - \rho v^2 \end{Bmatrix}, \quad (17)$$

where $d = c_{13} + c_{44}$. The $\hat{\varepsilon}\mathbf{m}$ vector and, hence, the ε scalar in Eq. (6) are also the same for all symmetry classes (13)–(15) [10]:

$$\hat{\varepsilon}\mathbf{m} = (\varepsilon_1 m_1, 0, \varepsilon_3 m_3), \quad \varepsilon = \varepsilon_1 m_1^2 + \varepsilon_3 m_3^2. \quad (18)$$

However, the form of the electric vector $\mathbf{e}$ according to Eq. (6) for the transversely isotropic crystals of three types is different. For the piezoelectric media belonging to classes (13) and (14), the electric vectors are expressed as

$$\mathbf{e} = e_{14}(0, m_1m_3, 0), \quad (19)$$

$$\mathbf{e} = \{(e_{15} + e_{31})m_1m_3, 0, e_{15}m_1^2 + e_{33}m_3^2\}, \quad (20)$$

respectively. For a medium of the symmetry class (15), the electric vector is given by a sum of expressions (19) and (20). Thus, the structure of the $\hat{F}$ matrix (5) for classes (13) and (14) is the same as in (17), but this conclusion is not valid for the $\hat{F}$ matrix of the symmetry classes (15), which contains no vanishing elements. In the same coordinates, the $\hat{N}$ matrix for the piezoelectric media belonging to classes (13) and (14) has the following form:

$$\hat{N} = e_{14}\begin{Bmatrix} 0 & (\varepsilon_3/\varepsilon)m_3^2 & 0 \\ -m_3 & 0 & -m_1 \\ 0 & -(\varepsilon_3/\varepsilon)m_1m_3^2 & 0 \end{Bmatrix}, \quad (21)$$

$$\hat{N} = \begin{Bmatrix} e_{15}m_3 & 0 & e_{15}m_1 \\ 0 & e_{15}m_3 & 0 \\ e_{31}m_1 & 0 & e_{33}m_3 \end{Bmatrix}$$

$$-\frac{1}{\varepsilon}\begin{Bmatrix} (e_{15}+e_{31})\varepsilon_1 m_1^2 m_3 & 0 & (e_{15}m_1^2 + e_{33}m_3^2)\varepsilon_1 m_1 \\ 0 & 0 & 0 \\ (e_{15}+e_{31})\varepsilon_3 m_1 m_3^2 & 0 & (e_{15}m_1^2 + e_{33}m_3^2)\varepsilon_3 m_3 \end{Bmatrix}, \quad (22)$$

respectively. For a medium of the symmetry classes (15), the $\hat{N}$ matrix is (by analogy with vector **e**) given by a sum of expressions (21) and (22). In classes (13) and (14) of higher symmetry, one of the natural wave branches for any direction **m** is purely transverse:

$$\mathbf{u}_t \parallel (0, 1, 0),$$
$$\rho v_t^2 = c_{66}m_1^2 + c_{44}m_3^2 + (e_{14}^2/\varepsilon)m_1^2 m_3^2. \quad (23)$$

Such purely transverse waves of the *t* mode are frequently called SH waves. The other two branches are polarized in the $\{m_1, m_3\}$ plane:

$$\mathbf{u}_{l,t'} \parallel (2dm_1m_3, 0, -\Delta_{14}^- m_1^2 + \Delta_{34}^- m_3^2 \pm R),$$
$$\rho v_{l,t'}^2 = (\Delta_{14}^+ m_1^2 + \Delta_{34}^+ m_3^2 \pm R)/2, \quad (24)$$

where

$$R = \sqrt{(\Delta_{14}^- m_1^2 - \Delta_{34}^- m_3^2)^2 + (2dm_1m_3)^2},$$
$$\Delta_{ij}^{\pm} = c_{ii} \pm c_{jj}. \quad (25)$$

The electric components of the above wave fields can be also determined for an arbitrary direction **m**. For a medium of the symmetry class (13):

$$\phi_t = (e_{14}/\varepsilon)m_1m_3u_t,$$
$$\mathbf{D}_t = e_{14}(\varepsilon_3/\varepsilon)m_3^2(m_3, 0, -m_1)u_t,$$
$$\phi_{l,t'}(\mathbf{m}) \equiv 0,$$
$$\mathbf{D}_{l,t'} = -e_{14}\{0, 1, 0\}[m_1(\mathbf{u}_{l,t'})_3 + m_3(\mathbf{u}_{l,t'})_1]. \quad (26)$$

In a less simple case of the symmetry class (14), we present only the result for the *t* branch:

$$\phi_t(\mathbf{m}) \equiv 0, \quad \mathbf{D}_t = (0, e_{15}m_3, 0)u_t. \quad (27)$$

The structure of acoustic waves in the media belonging to the symmetry classes (15) is more complicated. In this case, even a purely symmetric branch ($\mathbf{u}_t \parallel y$) exists only in the *xy* basis plane.

4. ZERO ELECTRIC FIELD LINES ON THE UNIT SPHERE

According to the second relation in (9), the electric field amplitude distribution on the unit sphere of the wave propagation directions is described by the equation

$$\mathbf{E}_\alpha(\mathbf{m}) = \text{const} \cdot \phi_\alpha(\mathbf{m})\mathbf{m}, \quad (28)$$

which shows that the zero values of $\mathbf{E}_\alpha(\mathbf{m})$ coincide with those of the potential $\phi_\alpha(\mathbf{m}) = \mathbf{e}(\mathbf{m}) \cdot \mathbf{u}_\alpha(\mathbf{m})/\varepsilon$. According to condition (8), these directions are determined by the equation

$$\mathbf{e}(\mathbf{m}) \cdot \bar{\hat{F}}_\alpha(\mathbf{m})\mathbf{c} = 0. \quad (29)$$

The acoustic waves (3) propagating in these directions contain no electrostatic components $\mathbf{E}_\alpha$, as in a nonpiezoelectric medium. Even a nonzero induction field $D_i = e_{ijk}u_{k,j}$ in these directions does not influence the parameters of the displacement wave.

The scalar equation (29) poses only one limitation on the direction of the wave normal $\mathbf{m} \equiv \mathbf{m}(\theta, \varphi)$ as a function of two spherical angular coordinates). In other words, Eq. (29) determines a line (or several lines) of nonpiezoelectric directions (in which $\mathbf{E}_\alpha = 0$) on the sphere $\mathbf{m}^2 = 1$. It should be noted that the condition of longitudinal nonpiezoactivity,

$$\mathbf{e}(\mathbf{m}) \perp \mathbf{u}_\alpha(\mathbf{m}), \quad (30)$$

in some special cases can be satisfied even on the whole $\mathbf{m}^2 = 1$ sphere. This takes place, in particular, in the transversely isotropic crystals belonging to the symmetry classes (13) (for the *l* and *t*' modes (26)) and (14) (for the *t* mode (27)). For all other crystals, including transversely isotropic crystals belonging to the symmetry classes (15), the geometric locus of the longitudinal nonpiezoactivity as the form of lines on the unit sphere $\mathbf{m}^2 = 1$. Such lines also exist in the piezoactive branches of the aforementioned high-symmetry media belonging to symmetry classes (13) and (14). For example, the zero-field lines $\mathbf{E}_\alpha = 0$ in the *l* and *t*' branches of the media of classes (14) and (15) appear at the intersection of the $\mathbf{m}^2 = 1$ sphere with the cones of directions defined by the polar angles θ_l and $\theta_{t'}$ as

$$\tan^2\theta_l \approx -e_{33}/(2e_{15} + e_{31}),$$
$$\tan^2\theta_{t'} \approx (e_{15} + e_{31} - e_{33})/e_{15}. \quad (31)$$

For simplicity, these expressions are written in an approximate form corresponding to the case of a weak electromechanical interaction and a small elastic anisotropy. Nevertheless, one can readily check that the exact condition for the existence of the aforementioned

nonpiezoactivity cones is the positive determinacy of the right-hand parts of the approximate formulas (31).

It is possible to prove that the longitudinal nonpiezoactivity lines in fact exist virtually in all (even triclinic) crystals. Let us consider a crystal with arbitrary anisotropy, which contains at least one acoustic axis of the general (conical) type. Here, it should be noted that no one real crystal without acoustic axes and no one triclinic crystal without conical axes are known so far. As was demonstrated in [12, 13], the polarization fields of elastic displacements $\mathbf{u}_{0\alpha}(\mathbf{m})$ for the volume natural acoustic waves in such a crystal can be arranged on the $\mathbf{m}^2 = 1$ sphere so that one is even,

$$\mathbf{u}_{02}(-\mathbf{m}) = \mathbf{u}_{02}(\mathbf{m}), \quad (32)$$

and two are odd,

$$\mathbf{u}_{01}(-\mathbf{m}) = -\mathbf{u}_{01}(\mathbf{m}), \quad \mathbf{u}_{03}(-\mathbf{m}) = -\mathbf{u}_{03}(\mathbf{m}). \quad (33)$$

The nondegenerate branch $\mathbf{u}_{03}(\mathbf{m})$ is always odd and continuous on the entire sphere of wave directions. As for the degenerate branches, $\mathbf{u}_{01}(\mathbf{m})$ and $\mathbf{u}_{02}(\mathbf{m})$, their evenness depends on the representation and can be changed simultaneously. These branches are continuous at all points of the sphere except for some nonclosed lines on which the $\mathbf{u}_{01}(\mathbf{m})$ and $\mathbf{u}_{02}(\mathbf{m})$ functions change sign. Such lines can be arbitrarily deformed on the unit sphere without changing the positions of terminal points (coinciding with the points of degeneracy). In fact, the representation is chosen by setting certain fixed positions of the sign reversal lines (coinciding for both degenerate branches).

One can readily check that the aforementioned properties of the fields of elastic displacements, which were established in [12, 13] for purely elastic media, are also valid for piezoelectrics. Taking into account that, according to relations (33), the $\mathbf{u}_{03}(\mathbf{m})$ function is odd and the $\mathbf{e}(\mathbf{m})$ function is (by definition (6)) even, we may conclude that the potential

$$\phi_{03}(\mathbf{m}) = \mathbf{e}(\mathbf{m}) \cdot \mathbf{u}_{03}(\mathbf{m})/\varepsilon$$

is an odd function

$$\phi_{03}(-\mathbf{m}) = -\phi_{03}(\mathbf{m}). \quad (34)$$

This result implies that, for any path connecting the opposite points $\mathbf{m}$ and $-\mathbf{m}$ on the unit sphere, there exists at least one point $\mathbf{m}_0$ such that $\phi_{03}(\mathbf{m}_0) = 0$. In scanning the paths on the unit sphere, points $\mathbf{m}_0$ will apparently form a closed line representing a geometric locus of the directions of longitudinal nonpiezoactivity for the nondegenerate branch.

For the degenerate branches $\phi_{01}(\mathbf{m})$ and $\phi_{02}(\mathbf{m})$, the considerations should be somewhat modified, while being still generally analogous to those used in solving a similar problem [12] concerning the existence of the lines of solutions in the degenerate branches of special volume waves in semi-infinite elastic media.[1] Not reproducing these considerations here, we only formulate the result: the longitudinal nonpiezoactivity lines exist in both degenerate branches and pass from one branch to another at the degenerate points. Thus, the following theorem of existence is valid:

All three wave branches in an arbitrary crystal containing conical acoustic axes must possess lines of longitudinal nonpiezoactivity directions on the unit sphere.

It should be also noted that, when a wave propagates along an acoustic axis $\mathbf{m}_d$ of any type, the continuum of possible orientations of the wave polarization $\mathbf{u}$ in the plane of degeneracy always contains a vector orthogonal to the $\mathbf{e}(\mathbf{m}_d)$ direction. This ensures nonpiezoactivity of the corresponding wave. Therefore, the acoustic axes must belong to the lines of longitudinal nonpiezoactivity.

The elements of crystal symmetry can become an independent factor accounting for the phenomenon of nonpiezoactivity under consideration. According to [1], any symmetry axis determines the direction of nonpiezoactivity for purely transverse modes, while a symmetry plane is the geometric locus of nonpiezoactive directions for the corresponding SH waves. Let us consider, for example, a monoclinic piezoelectric crystal belonging to one of the two possible symmetry classes: *m* or 2. In the first case, the electric vector $\mathbf{e}$ of a wave propagating in a plane of symmetry *m* must be orthogonal to the polarization vector of the *t* branch and, hence, occur in this plane. In the second case, the $\mathbf{e}$ vector for a wave normal occurring in the plane perpendicular to the dyad (2-fold) axis of symmetry must be parallel to this axis and, hence, orthogonal to polarization vectors (belonging to said plane) of the *l* and *t*' waves. Naturally, the latter property is valid for any other symmetry axis of even order. In monograph [1], this rule was formulated for planes orthogonal to the tetrad (4-fold) and hexad (6-fold) axes.

In particular, the coordinate planes of the crystal system orthogonal to the tetrad and dyad axes in cubic piezoelectrics (symmetry classes $\bar{4}3m$ and 23) must be nonpiezoactive for the corresponding *l* and *t*' branches. At the same time, the diagonal planes of {110} symmetry are nonpiezoactive for the corresponding *t* waves. One can readily check that, in the vicinity of the coordinate axes, the potential amplitudes for these branches can be represented as (Fig. 1)

$$\phi_{l,t'} \propto \theta^2 \sin 2\varphi, \quad (35)$$

$$\phi_t \propto \theta^2 \cos 2\varphi. \quad (36)$$

[1] The authors are grateful to A.L. Shuvalov for attracting their attention to this circumstance.

5. ZERO-INDUCTION POINTS ON THE UNIT SPHERE

5.1. General Case of Arbitrary Anisotropy

Now let us consider the conditions determining the propagation directions $\mathbf{m}_0$ in which the electric induction vector $\mathbf{D}_\alpha = \hat{N}\mathbf{u}_\alpha$ defined in (9) vanishes. Taking into account identity (12) and the definition of the adjoint tensor

$$\hat{N}\bar{\hat{N}} = \hat{I}\det\hat{N}, \tag{37}$$

one can readily check that $\mathbf{D}_\alpha = 0$ for the directions $\mathbf{m}_0$ such that $\mathbf{u}_\alpha \parallel \bar{\hat{N}}\mathbf{d}$, where $\mathbf{d}$ is any vector obeying the condition $\bar{\hat{N}}\mathbf{d} \neq 0$. For these directions $\mathbf{m}_0$, according to condition (8), we also have

$$\bar{\hat{F}}_\alpha\mathbf{c} \parallel \bar{\hat{N}}\mathbf{d}. \tag{38}$$

In the general case, this condition gives two equations with two unknowns θ and φ, which determine the positions of isolated points $\mathbf{m}_0(\theta, \varphi)$ such that $\mathbf{D}_\alpha = 0$ on the unit sphere $\mathbf{m}^2 = 1$.

There is the well-known Brouwer theorem in the topology, according to which *any continuous transform on a sphere, not mapping any point by its antipode, has at least two stationary points.*

Now let us consider a distribution of vectors $\mathbf{D}_\alpha(\mathbf{m})$ continuous everywhere on the $\mathbf{m}^2 = 1$ sphere. The continuity of $\mathbf{D}_\alpha(\mathbf{m})$ is ensured when the corresponding branch α is nondegenerate. According to Brouwer's theorem, this distribution of $\mathbf{D}_\alpha$ vectors tangent to the sphere must have two stationary points for which $\mathbf{D}_\alpha = 0$. On the other hand, relations (9) and (10) imply that this distribution also possesses an additional property: $\mathbf{D}_\alpha(-\mathbf{m}) \parallel \mathbf{D}_\alpha(\mathbf{m})$. For this reason, the pair of points stipulated by Brouwer's theorem includes the inversion-equivalent stationary points $\mathbf{m}_0$ and $-\mathbf{m}_0$. Thus, the following theorem of existence of the transverse nonpiezoactivity directions is valid:

In any crystal of arbitrary anisotropy, each nondegenerate branch must contain at least one inversion-nonequivalent zero-induction point $\mathbf{m}_0$ *such that* $\mathbf{D}_\alpha(\mathbf{m}) = 0$ *on the unit sphere of directions.*

Therefore, the zero-induction points in a wave field $\mathbf{D}_\alpha(\mathbf{m})$ must exist even in triclinic crystals. Of course, the positions of such points in the general case (i.e., the solutions to Eq. (38) in the general form) cannot be found analytically. However, in some more symmetric crystals, the nonpiezoactive directions $\mathbf{m}_0$ can be determined without cumbersome computations.

5.2. Zero-Induction Points Related to Elements of the Crystal Symmetry

5.2.1. Longitudinal waves propagating along symmetry axes. Let us consider a wave propagating the direction m_0 in a piezoelectric crystal, which coincides with a symmetry axis of any order except for dyad axes (e.g., this can be the 3, 4, $\bar{4}$, 6, or $\bar{6}$-fold axis). As is known [11], any symmetry axis (including a dyad axis) is a longitudinal normal. Evidently, the electric induction $\mathbf{D}_l(\mathbf{m}_0)$ accompanying the given longitudinal wave ($\mathbf{u}_l \parallel \mathbf{m}_0$) must be zero, since otherwise the $\mathbf{D}_l(\mathbf{m}_0)$ vector would possess two equivalent orientations (to contradict the single-valuedness of the third relation in (9)):

$$\mathbf{D}_l = \hat{N}(\mathbf{m}_0)\mathbf{m}_0. \tag{39}$$

It should be noted that these considerations are not valid in the case of transverse branches, which are always degenerate for the selected direction, that is, possessing equal phase velocities ($v_t = v_{t'}$) and, hence, arbitrary orientations of $\mathbf{u}_{t,t'}$ and $\mathbf{D}_{t,t'}$ in the plane:

$$\mathbf{u}_{t,t'}, \mathbf{D}_{t,t'} \perp \mathbf{m}_0. \tag{40}$$

The wave propagating along a dyad axis should be treated separately (albeit with the same result). In the general case, this direction is not an acoustic axis. On the other hand, the $\mathbf{u}_\alpha$ and $\mathbf{D}_\alpha$ vectors are determined to within the sign and, hence, their rotations around the dyad axis cannot be considered as different solutions. One can readily check that, for a wave propagating along the dyad axis, the transverse branches are characterized by nonzero induction vectors. However, now we will show that the longitudinal branch in the same direction always obeys the relation $\mathbf{D}_l(\mathbf{m}_0) = 0$. Com-

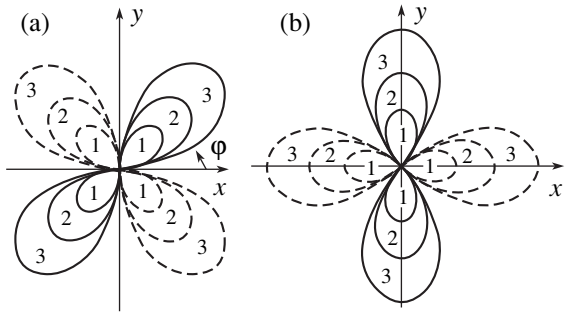

Fig. 1. Polar diagrams of the electric potentials (a) $\phi_{l,t'}(\varphi)$ and (b) $\phi_t(\varphi)$ at θ = const in the vicinity of the (0, 0, 1) direction in a cubic piezoelectric crystal. Numbers 1, 2, and 3 refer to the angles $\theta_1 < \theta_2 < \theta_3$; solid and dashed lines correspond to the potentials of different signs.

bining condition (39) with relation (10) for $\mathbf{m} = \mathbf{m}_0$, we obtain

$$\mathbf{D}_l = \hat{N}(\mathbf{m}_0)\mathbf{m}_0 = (\hat{e}\mathbf{m}_0)\mathbf{m}_0 - \frac{(\hat{\varepsilon}\mathbf{m}_0)[(\mathbf{m}_0\hat{e}\mathbf{m}_0)\mathbf{m}_0]}{\mathbf{m}_0 \cdot \hat{\varepsilon}\mathbf{m}_0}. \quad (41)$$

Let us check that the right-hand part of this expression vanishes even for a monoclinic crystal of the symmetry class 2. Indeed, selecting the z axis in (41) along the dyad axis ($2 \parallel \mathbf{m}_0$), we obtain

$$\mathbf{D}_l = \left\{e_{13} - \frac{\varepsilon_{13}e_{33}}{\varepsilon_{33}}, e_{23} - \frac{\varepsilon_{23}e_{33}}{\varepsilon_{33}}, 0\right\}. \quad (42)$$

However, according to [1, 10], the off-diagonal components of $\hat{e}$ and $\hat{\varepsilon}$ tensors entering into relation (42) for a monoclinic crystal of the symmetry class 2 in this coordinate system are vanishing: $e_{13} = e_{23} = \varepsilon_{13} = \varepsilon_{23} = 0$ and, hence,

$$\mathbf{D}_l(\mathbf{m}_0 \parallel 2) = 0. \quad (43)$$

Evidently, Eq. (43) is valid for all crystals of various classes possessing dyad axes. Thus, the following statement is valid:

A longitudinal wave propagating along any axis of symmetry in a piezoelectric crystal is accompanied by an electric component with zero induction.

For example, let us consider a piezoelectric crystal of the orthorhombic symmetry class 222. According to the above theorem, all three dyad axes in this crystal are the zero-induction directions $\mathbf{m}_0$ for the longitudinal modes. However, it can be shown that another four inversion-inequivalent asymmetric directions $\mathbf{m}_0$ with zero induction ($\mathbf{D}_l = 0$) may exist in a quasi-longitudinal branch of this crystal:

$$(\theta_0, \pm\varphi_0), (\theta_0, \pm\varphi_0 + \pi), \quad (44)$$

$$\theta_0 = \arccot\sqrt{\frac{e_{36}\varepsilon_1\varepsilon_2}{(e_{14}\varepsilon_2 + e_{25}\varepsilon_1)\varepsilon_3}}, \quad (45)$$

$$\varphi_0 = \arctan\sqrt{\frac{e_{25}\varepsilon_1}{e_{14}\varepsilon_2}}. \quad (46)$$

For simplicity, solutions (45) and (46) are written in the approximation of small piezoelectric moduli and weak elastic anisotropy. In this approximation, a necessary condition for the existence of the above series of zero-induction points is that all the piezoelectric moduli entering into relations (45) and (46) must have the same sign (Fig. 2). It should be noted that cubic piezoelectric crystals (symmetry classes $\bar{4}3m$ and 23) always correspond to Fig. 2b, since additional zero-induction directions according to relations (44)–(46) always appear along the triad axes.

5.2.2. Transverse (SH) waves propagating in symmetry planes. Example 1: symmetry class *m*. Let the z axis be perpendicular to the plane of symmetry of a monoclinic crystal ($z \perp m$) and consider the t branch of a wave propagating in this plane:

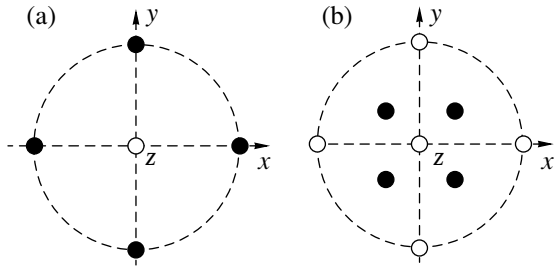

Fig. 2. Diagrams of the directions of propagation of the transversely nonpiezoactive quasi-longitudinal acoustic waves in crystals of the symmetry class 222. The stereographic projections are given for the cases when (a) the sign of the piezoelectric modulus e_{36} is opposite to that of e_{14} and/or e_{25} and (b) all piezoelectric moduli have the same sign.

$$\mathbf{m} = (m_1, m_2, 0), \quad \mathbf{u}_t \parallel (0, 0, 1). \quad (47)$$

One can readily check that in such waves

$$\mathbf{D}_t = (0, 0, e_{35}m_1 + e_{34}m_2)u_t \parallel \mathbf{u}_t. \quad (48)$$

Therefore, the symmetry plane always has a single direction $\mathbf{m}_0$ corresponding to zero induction $\mathbf{D}_t$, which is determined by the azimuthal angle φ_0:

$$\mathbf{m}_0 = (m_{01}, m_{02}, 0), \quad \tan\varphi_0 = \frac{m_{02}}{m_{01}} = -\frac{e_{35}}{e_{34}}. \quad (49)$$

The maximum amplitude of $\mathbf{D}_t$ in this plane corresponds to the direction $\mathbf{m}_{\max} = (m_{02}, -m_{01}, 0)$, which is perpendicular to $\mathbf{m}_0$.

Example 2: symmetry class 3*m*. In trigonal crystals, the situation with transverse nonpiezoactive directions for the t waves in each of the three symmetry planes containing the triad axis is completely analogous to the above case of a monoclinic crystal. For example, in the yz symmetry plane, relations (47)–(49) have to be replaced by

$$\mathbf{m} = (0, m_2, m_3), \quad \mathbf{u}_t \parallel (1, 0, 0), \quad (50)$$

$$\mathbf{D}_t = (-e_{22}m_2 + e_{15}m_3, 0, 0)u_t \parallel \mathbf{u}_t, \quad (51)$$

$$\mathbf{m}_0 = (0, m_{02}, m_{03}), \quad \tan\theta_0 = \frac{m_{02}}{m_{03}} = \frac{e_{15}}{e_{22}}, \quad (52)$$

where θ_0 is the polar angle between $\mathbf{m}_0$ and the triad axis.

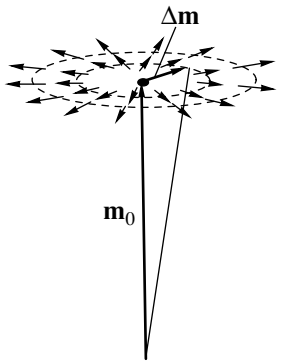

Fig. 3. A singular vector distribution of the $\mathbf{D}_\alpha(\mathbf{m})$ vectors in the vicinity of a zero-induction point $\mathbf{m}_0$.

Example 3: symmetry class *mm*2. For a *t* wave propagating in the *yz* symmetry plane of an orthorhombic crystal, we have

$$\begin{gathered}\mathbf{m} = (0, m_2, m_3), \quad \mathbf{u}_t \parallel (1, 0, 0), \\ \mathbf{D}_t = (e_{15}m_3, 0, 0)u_t \parallel \mathbf{u}_t.\end{gathered} \tag{53}$$

Evidently, in this case

$$\mathbf{m}_0 = (0, 1, 0), \quad \mathbf{m}_{\max} = (0, 0, 1). \tag{54}$$

Relations (53) and (54) are also valid for tetragonal crystals of the symmetry class 4*mm*.

Example 4: symmetry classes $\bar{4}2m$, $\bar{4}3m$, and 23. For the *x* axis parallel to the dyad axis ($x \parallel 2$), transverse waves propagating in the diagonal plane $(1, \bar{1}, 0)$ obey the relations

$$\mathbf{m} = (m_1, -m_1, m_3), \quad \mathbf{u}_t \parallel (1, 1, 0). \tag{55}$$

These waves exhibit electric components with the amplitude of induction

$$\mathbf{D}_t = e_{14}m_3(1, 1, 0)u_t \parallel \mathbf{u}_t \tag{56}$$

and, hence, have the following special directions:

$$\mathbf{m}_0 = (1, -1, 0)/\sqrt{2}, \quad \mathbf{m}_{\max} = (0, 0, 1). \tag{57}$$

Example 5: symmetry classes 6*mm* and ∞*m* (14). Any plane containing the principal symmetry axis in such a crystal is the plane of symmetry *m*. According to relations (27), the electric induction vector of *t* waves propagating in such planes is orthogonal to *m* and proportional to m_3. Therefore, $\mathbf{D}_t$ vanishes along the entire equator $m_3 = 0$:

$$\mathbf{D}_t(m_1, m_2, 0) = 0. \tag{58}$$

It should be noted that, for crystals of the same symmetry classes, the other transverse branch (*t*') polarized along the principal symmetry axis also has a zero-induction line on the same equator:

$$\mathbf{D}_{t'}(m_1, m_2, 0) = 0. \tag{59}$$

One can readily check that relation (58), but not (59), is also valid for crystals of the symmetry classes (13) and (15).

6. SINGULARITIES OF THE INDUCTION VECTOR FIELDS IN THE VICINITY OF $\mathbf{D} = 0$ POINTS

6.1. General Case of Arbitrary Anisotropy

The vector field $\mathbf{D}_\alpha(\mathbf{m})$, which is orthogonal to the wave normal $\mathbf{m}$, may exhibit orientational singularity in the vicinity of the zero-induction points $\mathbf{m}_0$ (Fig 3). Let us consider the $\mathbf{D}_\alpha(\mathbf{m})$ function for $\mathbf{m} = \mathbf{m}_0 + \Delta\mathbf{m}$, where $\Delta\mathbf{m} \perp \mathbf{m}_0$ and $|\Delta\mathbf{m}| \ll 1$. Using condition (8) and the third relation in (9), we obtain

$$\mathbf{D}_\alpha(\mathbf{m}) \parallel \hat{N}(\mathbf{m})\bar{\hat{F}}_\alpha(\mathbf{m})\mathbf{c}. \tag{60}$$

Taking into account that $\mathbf{D}_\alpha(\mathbf{m}_0) = 0$, we have, to a first approximation,

$$\begin{gathered}\mathbf{D}_\alpha(\mathbf{m}) \parallel \Delta\mathbf{m}\hat{Q}_\alpha \\ \equiv \Delta m_i \left\{ \frac{\partial}{\partial m_i}[\hat{N}(\mathbf{m})\bar{\hat{F}}_\alpha(\mathbf{m})\mathbf{c}] \right\}_{\mathbf{m} = \mathbf{m}_0}.\end{gathered} \tag{61}$$

For the transverse $\mathbf{D}_\alpha(\mathbf{m})$ field (see (11)), the asymmetric tensor entering into expression (61),

$$\hat{Q}_\alpha = \nabla_{\boldsymbol{m}} \otimes \hat{N}(\mathbf{m})\bar{\hat{F}}_\alpha(\mathbf{m})\mathbf{c}|_{\mathbf{m} = \mathbf{m}_0}, \tag{62}$$

must be planar, that is, its spectral expansion can be represented as a sum of two dyads:

$$\hat{Q}_\alpha = \lambda_{\alpha 1}\tilde{\mathbf{e}}_{\alpha 1} \otimes \mathbf{e}_{\alpha 1} + \lambda_{\alpha 2}\tilde{\mathbf{e}}_{\alpha 2} \otimes \mathbf{e}_{\alpha 2}, \tag{63}$$

where $\lambda_{\alpha j}$, $\tilde{\mathbf{e}}_{\alpha j}$, and $\mathbf{e}_{\alpha j}$ are the eigenvalues and eigenvectors (left and right) of the $\hat{Q}_\alpha$ tensor ($\mathbf{e}_{\alpha 1}$ and $\mathbf{e}_{\alpha 2}$ must be orthogonal to $\mathbf{m}_0$). Note that the $\tilde{\mathbf{e}}_{\alpha j}$ eigenvectors (in contrast to $\mathbf{e}_{\alpha j}$) in the general case do not belong to a plane orthogonal to $\mathbf{m}_0$, but their components $\tilde{\mathbf{e}}^{\parallel}_{\alpha j}$ oriented along $\mathbf{m}_0$ are insignificant for our analysis.

Let us decompose each $\tilde{\mathbf{e}}_{\alpha j}$ eigenvector into two components

$$\begin{gathered}\tilde{\mathbf{e}}_{\alpha j} = \tilde{\mathbf{e}}^{\parallel}_{\alpha j} + \tilde{\mathbf{e}}^{\perp}_{\alpha j}, \quad \tilde{\mathbf{e}}^{\parallel}_{\alpha j} \parallel \mathbf{m}_0, \\ \tilde{\mathbf{e}}^{\perp}_{\alpha j} = (\hat{I} - \mathbf{m}_0 \otimes \mathbf{m}_0)\tilde{\mathbf{e}}_{\alpha j} \perp \mathbf{m}_0,\end{gathered} \tag{64}$$

and form a more convenient matrix

$$\hat{Q}_\alpha^\perp = (\hat{I} - \mathbf{m}_0 \otimes \mathbf{m}_0)\hat{Q}_\alpha = \lambda_{\alpha 1}\tilde{\mathbf{e}}_{\alpha 1}^\perp \otimes \mathbf{e}_{\alpha 1} + \lambda_{\alpha 2}\tilde{\mathbf{e}}_{\alpha 2}^\perp \otimes \mathbf{e}_{\alpha 2}, \quad (65)$$

which will be used below instead of $\hat{Q}_\alpha$:

$$\mathbf{D}_\alpha \| \Delta\mathbf{m}\hat{Q}_\alpha^\perp. \quad (66)$$

Let the angle of orientation Φ of the $\mathbf{D}_\alpha(\mathbf{m})$ vector be measured from the $\mathbf{e}_{\alpha 1}$ direction, and the analogous angle φ for $\Delta\mathbf{m}$ in the same plane, from the $\tilde{\mathbf{e}}_{\alpha 1}^\perp$ direction (Fig. 4). In these terms, we can write

$$\tan\Phi = \frac{\mathbf{D}_\alpha \cdot \mathbf{e}_{\alpha 2}}{\mathbf{D}_\alpha \cdot \mathbf{e}_{\alpha 1}} = \frac{\lambda_{\alpha 2}}{\lambda_{\alpha 1}}\frac{\Delta\mathbf{m} \cdot \tilde{\mathbf{e}}_{\alpha 2}^\perp}{\Delta\mathbf{m} \cdot \tilde{\mathbf{e}}_{\alpha 1}^\perp} = \frac{\lambda_{\alpha 2}}{\lambda_{\alpha 1}}\tan\varphi. \quad (67)$$

Thus, the complete turn of $\Delta\mathbf{m}$ around $\mathbf{m}_0$ in the plane orthogonal to $\mathbf{m}_0$ implies the compete turn of $\mathbf{D}_\alpha(\mathbf{m})$ in the same or in the opposite direction (depending on the sign of $\det\hat{Q}_\alpha^\perp = \lambda_{\alpha 1}\lambda_{\alpha 2}$), which corresponds to the Poincaré index of the given singular point (Fig. 5)

$$n = \operatorname{sgn}\det\hat{Q}_\alpha^\perp. \quad (68)$$

The above considerations fail to be valid in some particular cases, when one of the eigenvalues ($\lambda_{\alpha 1}$ or $\lambda_{\alpha 2}$) of matrices (63) and (65) vanishes. In such cases, $\det\hat{Q}_\alpha^\perp = 0$, but formula (68) is not applicable. Indeed, let $\lambda_{\alpha 2} = 0$ at $\mathbf{m}_0$. Then,

$$\mathbf{D}_\alpha \| \Delta\mathbf{m}\hat{Q}_\alpha^\perp = \lambda_{\alpha 1}(\Delta\mathbf{m} \cdot \tilde{\mathbf{e}}_{\alpha 1}^\perp)\mathbf{e}_{\alpha 1} \quad (69)$$

and a zero-induction line can pass via m_0 in the direction of $\Delta\mathbf{m} \perp \tilde{\mathbf{e}}_{\alpha 1}^\perp$, but only provided that $\lambda_{\alpha 2} = 0$ is valid. In this situation, the very concept of the Poincaré index is inapplicable. However, if the vanishing of $\lambda_{\alpha 2}$ has a strictly local character and takes place only along $\mathbf{m}_0$, then we are dealing with a very special singularity analogous to a local-wedge degeneracy known in the theory of acoustic axes [6]. It can be shown that a topological charge of the corresponding singularity in the $\mathbf{D}_\alpha(\mathbf{m})$ field in this case can take one of three values: $n = 0, \pm 1$. However, both situations (point and line) with a zero induction amplitude are very exclusive and never encountered in real (even symmetric) crystals. Below we will consider zero-induction lines of this kind in model crystals. It will be demonstrated that the examples of the $\mathbf{D}_\alpha = 0$ lines corresponding to Eqs. (58) and (59) belong to a different type.

On the other hand, the singular points with $n = \pm 1$ depicted in Fig. 5 are rather widely encountered in real crystals. For example, all directions $\mathbf{m}_0$ in Fig. 2b corresponding to orthorhombic (222) or cubic ($\bar{4}3m$ and 23) crystals are characterized by topological charges $n = \pm 1$ (in Fig. 2, filled and empty circles correspond to +1 and –1, respectively). Figure 6 is a schematic diagram of the $\mathbf{D}_l(\mathbf{m})$ distribution in the central region of the circle in Fig. 2b.

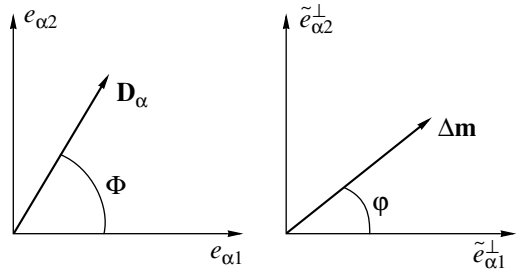

Fig. 4. The angles of orientation of the $\mathbf{D}_\alpha$ and $\Delta\mathbf{m}$ vectors in the plane orthogonal to $\mathbf{m}_0$.

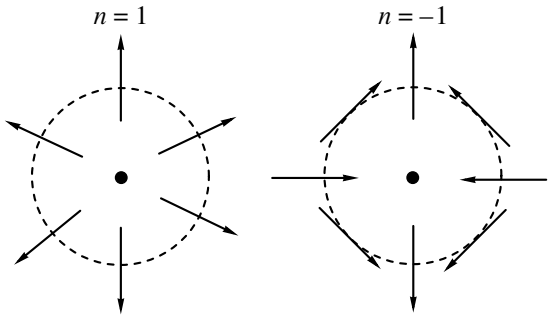

Fig. 5. The two main types of singularities in the vicinity of zero-induction points of the $\mathbf{D}_\alpha(\mathbf{m})$ vectors.

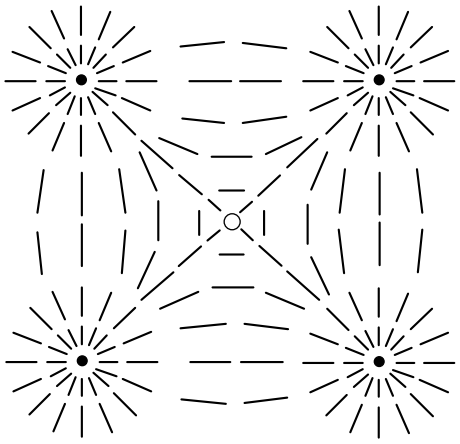

Fig. 6. A schematic image of the $\mathbf{D}_l(\mathbf{m})$ vector field distribution over a group of five singular points $\mathbf{m}_0$ in the central region of a circle in Fig. 2b in the representation of nondirectional segments.

Now let us consider some examples of crystals belonging to particular crystal symmetry systems.

6.1.1. Longitudinal waves propagating along symmetry axes. Example 1. For the $\mathbf{m}_0 \parallel 2$ direction in a monoclinic crystal with dyad axis, we have

$$n = \operatorname{sgn}\{(a_1 + b_1)(a_2 + b_2) - c_1 c_2\}, \tag{70}$$

where

$$a_1 = \frac{e_{14}c_{36}}{\bar{\Delta}_{34}}, \quad b_1 = \frac{e_{15}d_1}{\bar{\Delta}_{35}} - \frac{e_{33}\varepsilon_1}{\varepsilon_3}, \quad c_1 = \frac{e_{14}d_2}{\bar{\Delta}_{34}} + \frac{e_{15}c_{36}}{\bar{\Delta}_{35}}, \tag{71}$$

$$a_2 = \frac{e_{25}c_{36}}{\bar{\Delta}_{35}}, \quad b_2 = \frac{e_{24}d_2}{\bar{\Delta}_{34}} - \frac{e_{33}\varepsilon_2}{\varepsilon_3}, \quad c_2 = \frac{e_{24}c_{36}}{\bar{\Delta}_{34}} + \frac{e_{25}d_1}{\bar{\Delta}_{35}}, \tag{72}$$

$d_i = c_{i3} + c_{33}$. The $\hat{\varepsilon}$ tensor is assumed to be diagonal, which can be ensured by the appropriate choice of the *x*- and *y* axes of the crystallographic coordinate system with the *z* axis parallel to the dyad axis.

Example 2. For the $\mathbf{m}_0 \parallel 2$ direction in an orthorhombic crystal belonging to the symmetry class *mm*2, we have

$$n = \operatorname{sgn}(b_1 b_2). \tag{73}$$

Example 3. For the $\mathbf{m}_0 \parallel 2 \parallel z$ direction in an orthorhombic crystal belonging to the symmetry class 222, we have

$$n^{(z)} = \operatorname{sgn}(a_1 d_2 / a_2 d_1). \tag{74}$$

Analogous formulas for the $n^{(x)}$ and $n^{(y)}$ are obtained from (74) by cyclic rearrangement of the indices. Note that, in the isotropic limit, we obtain

$$d_1/\bar{\Delta}_{35} = d_2/\bar{\Delta}_{34} \longrightarrow 2, \tag{75}$$

and only a very large elastic anisotropy can change the signs of the ratios in formula (74). For this reason, these signs for most orthorhombic crystals are determined only by the piezoelectric moduli $n^{(z)} = \operatorname{sgn}(e_{14}/e_{25})$.

Example 4. As can be readily checked, for the principal symmetry axes in crystals of the symmetry classes 422, 622, ∞22, 4*mm*, 6*mm*, ∞*mm*, 4, 6, ∞, 32, 3*m*, and 3 we have

$$n^{(z)} = 1, \tag{76}$$

and in crystals of the symmetry classes $\bar{4}2m$, $\bar{4}$, $\bar{4}3m$, and 23,

$$n^{(z)} = -1. \tag{77}$$

6.1.2. Transverse (SH) waves propagating in symmetry planes. We will not write lengthy expressions determining the choice between $n = \pm 1$ indices for the waves along $\mathbf{m}_0$ directions in monoclinic and trigonal crystals (see relations (49) and (52), respectively) and instead pass to orthorhombic crystals.

Example 1. For the transverse acoustic waves (53) and (54) propagating in the vicinity of the zero-induction direction $\mathbf{m}_0 = (0, 1, 0)$ in an orthorhombic crystal belonging to the symmetry class *mm*2, the singular induction field is characterized by the Poincaré index

$$n = \operatorname{sgn}\left[\frac{e_{32}(c_{12} + c_{66})/\bar{\Delta}_{26} - e_{31}}{e_{15}}\right]. \tag{78}$$

Example 2. For the same direction in tetragonal crystals belonging to the symmetry class 4*mm*, we have

$$n = \operatorname{sgn}\left(\frac{e_{31}}{e_{15}}\frac{c_{12} - c_{11} + 2c_{66}}{\bar{\Delta}_{16}}\right). \tag{79}$$

Example 3. For the direction $\mathbf{m}_0 = (1, -1, 0)/\sqrt{2}$ in tetragonal crystals belonging to the symmetry class $\bar{4}2m$, we obtain

$$n = -\operatorname{sgn}(e_{14}/e_{36}). \tag{80}$$

Example 4. For the same direction in cubic crystals of the symmetry classes $\bar{4}3m$, 23 in the diagonal symmetry plane, as well as for the symmetry-equivalent direction $\mathbf{m}_0 = (-1, 1, 0)/\sqrt{2}$, we have (for any combinations of the moduli)

$$n = -1. \tag{81}$$

6.2. Special Types of Singularities

The above analysis is exhaustive only provided that the $\hat{Q}_\alpha$ tensor (62) is nonzero. As was shown above, usual systems have $\hat{Q}_\alpha \neq 0$. However, in some very special cases, this tensor may vanish in some special directions because of high symmetry or as a result of vanishing of certain combinations of the material tensor components. In such cases, the general expressions are very lengthy and we only present here some final results. For $\hat{Q}_\alpha(\mathbf{m}_0) = 0$, the distribution of the induction vector field in the vicinity of $\mathbf{m}_0$ has four additional variants depicted in Fig. 7. The first three of these correspond to isolated singular points with the Poincaré indices $n = 0, \pm 2$ (Figs. 7a–7c), while the fourth variant correspond to the existence of a $\mathbf{D} = 0$ line passing via the $\mathbf{m}_0$ point (Fig. 7d). This very situation is observed on the equator $m_3 = 3$ for the transverse tangentially polarized *t* waves (58) in all transverse-isotropic media (13)–(15) and for the trans-

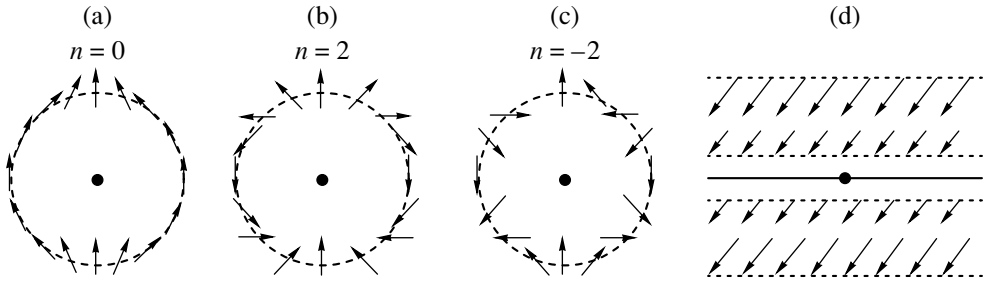

Fig. 7. Four possible types of the $\mathbf{D}_\alpha(\mathbf{m})$ vector field distribution around a zero-induction point $\mathbf{m}_0$, where $\mathbf{D}_\alpha(\mathbf{m}) = \hat{Q}_\alpha(\mathbf{m}_0) = 0$.

verse *t*' waves (59) polarized along the principal symmetry plane in the media of symmetry classes 6*mm* and ∞m. The only alternative example of $\hat{Q}_\alpha$ matrix vanishing is offered by a crystal with hexad axis $\bar{6}$. In this case, all three wave branches are featuring the identical singularities with $n = -2$ (Fig. 7c).

6.2.1. A model crystal of the symmetry class *mm*2. Let us assume that one piezoelectric modulus in the crystal under consideration is much smaller than the other moduli. In particular, we consider a conventional crystallographic coordinate system with the *x* and *y* axes perpendicular to the symmetry planes and the *z* axis parallel to the dyad axis, in which

$$|e_{31}| \ll |e_{15}|, |e_{24}|, |e_{32}|, |e_{33}|. \tag{82}$$

One can readily check that, in a zero-order approximation with $e_{31} = 0$, a quasi-longitudinal nondegenerate wave branch along the $\mathbf{m}_0 = (1, 0, 0)$ direction features the above special situation, whereby simultaneously $\mathbf{D}_l(\mathbf{m}_0) = 0$ and $\hat{Q}_{\alpha 1}(\mathbf{m}_0) = 0$. In this case, the $\mathbf{D}_l(\mathbf{m})$ vector field distribution in the *yz* plane in the vicinity of $\mathbf{m}_0$ is described by the expression

$$\begin{aligned}\mathbf{D}_l \| \{0, g_2 \sin 2\phi, g_1 + \varepsilon_1 e_{32}\gamma_2 \\ -(g_1 - \varepsilon_1 e_{32}\gamma_2)\cos 2\phi\},\end{aligned} \tag{83}$$

where ϕ is a polar angle of the $\mathbf{m}$ direction measured from the *y* axis in the *yz* plane and

$$\begin{aligned} g_1 &= \gamma_1(\varepsilon_1 e_{33} - \varepsilon_3 e_{15}), \\ g_2 &= (\gamma_1 + \gamma_2)\varepsilon_1 e_{24} - \varepsilon_2 e_{15}, \\ \gamma_1 &= \bar{d}_5/\bar{\Delta}_{15}, \quad \gamma_2 = d_6/\Delta_{16}, \\ \bar{d}_5 &= c_{13} + \bar{c}_{55}, \quad \bar{\Delta}_{15} = c_{11} - \bar{c}_{55}, \\ \bar{c}_{55} &= c_{55} + e_{15}^2/\varepsilon_1, \quad d_6 = c_{12} + c_{66}. \end{aligned} \tag{84}$$

Expression (83) shows that, depending on the material constants, the $\mathbf{D}_l(\mathbf{m})$ field always corresponds to one of the possible variants depicted in Fig. 7. The Poincaré indices for the point singularities corresponding to Figs. 7a–7c are as follows:

$$n_l = \begin{cases} 0, & g_1 e_{32}\gamma_2 > 0, \\ 2\,\mathrm{sgn}[(g_1 - \varepsilon_1 e_{32}\gamma_2)g_2], & g_1 e_{32}\gamma_2 < 0. \end{cases} \tag{85}$$

A condition for the existence of zero-induction lines in the $\mathbf{D}_l(\mathbf{m})$ field for Fig. 7d is

$$g_1 e_{32}\gamma_2 = 0 \text{ or } g_2 = 0, \quad g_1 e_{32}\gamma_2 < 0. \tag{86}$$

According to expression (83), a zero-induction line for $g_1 = 0$ passes via $\mathbf{m}_0$ along the *z* axis (Fig. 8a). For $e_{32}\gamma_2 = 0$, a similar zero-induction line is directed along the *y* axis (Fig. 8b). If $g_1 = 0$ simultaneously with $e_{32}\gamma_2 = 0$, the two lines coexist (Fig. 8c). Finally, when $g_2 = 0$ and $g_1 e_{32}\gamma_2 < 0$, the system features an oblique cross of zero-induction lines (Fig. 8d) with the mutual orientation determined by the equation

$$\cos 2\phi = \frac{g_1 + \varepsilon_1 e_{32}\gamma_2}{g_1 - \varepsilon_1 e_{32}\gamma_2}. \tag{87}$$

6.2.2. Behavior of point singularities in response to perturbations in the material moduli. The point singularities of various types in vector fields $\mathbf{D}_\alpha(\mathbf{m})$ behave differently (shift, split, or disappear) in response to perturbations in the material moduli. An analysis of this situation, analogous to that carried out in [6], showed that singularities with $n = \pm 1$ (Fig. 5) are topologically stable and can only be displaced by such

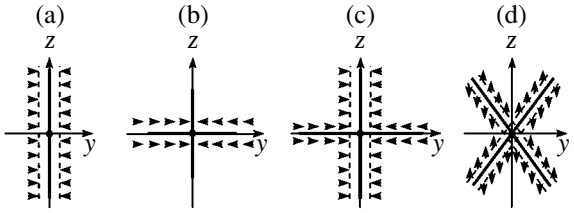

Fig. 8. Four possible types of the $\mathbf{D}_l(\mathbf{m}) = 0$ lines in a model orthorhombic crystal.

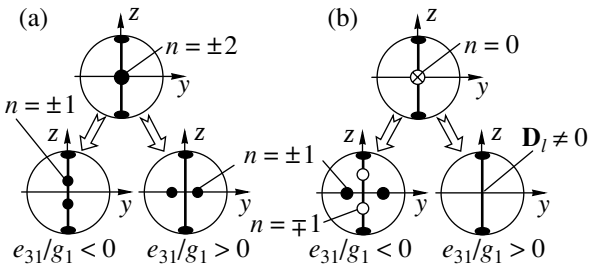

Fig. 9. Diagrams illustrating the splitting of singular points with $n = \pm 2$ (a) and 0 (b) in a model crystal of the symmetry class *mm*2.

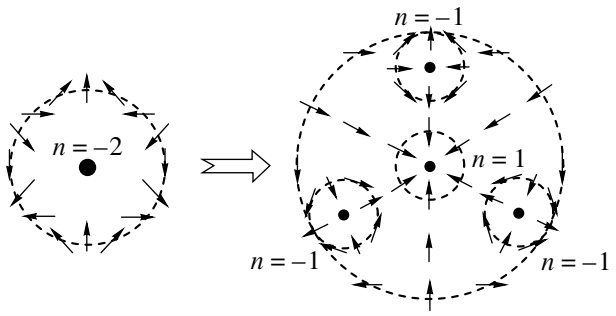

Fig. 10. Transformation of a singularity with $n = -2$ along the axis of symmetry $\bar{6}$ into four singular points with $n = \pm 1$ during the $\bar{6} \longrightarrow 3$ phase transition.

perturbations. The singular points of other types (Figs. 7a–7c) are unstable and either split or (in accordance with the law of topological charge conservation) or disappear (provided only that $n = 0$). The zero-induction lines (Fig. 7d and Fig. 8) are also unstable and disappear either completely or leaving a certain number of isolated zero points.

These general properties can be illustrated by a particular example using a model crystal of the symmetry class *mm*2 with a small modulus e_{31} described above. It should be recalled that relations (83)–(87) were obtained in the zero-order approximation for $e_{31} = 0$. In the next order with respect to the small parameter e_{31}, the initial singularity exhibits splitting along the direction $\mathbf{m}_0 = (1, 0, 0)$ so as to form two or four singular points:

$$\mathbf{m}_0 + \delta\mathbf{m} = \begin{cases} (1, \pm\mu_2, 0), & \mu_2^2 = -e_{31}/e_{32}\gamma_2, \\ (1, 0, \pm\mu_3), & \mu_3^2 = -\varepsilon_1 e_{31}/g_1. \end{cases} \quad (88)$$

As can be seen from expressions (88) with $g_1 e_{32}\gamma_2 < 0$, the zero-order approximation along $\mathbf{m}_0$ corresponds to a singularity with $n = +2$ or -2. The introduction of a small e_{31} modulus leads to a symmetric splitting of this singularity into a pair of zero-induction points with equal topological charges $n = 1$ or $n = -1$ along the y or z axis, depending on the sign of the e_{31}/g_1 ratio (Fig. 9a). For $g_1 e_{32}\gamma_2 > 0$, when the initial topological charge in the zero-order approximation is zero, the perturbed pattern comprises either four singularities with a zero total index n (for $e_{31}/g_1 < 1$) or none of them (which corresponds to the absence of zero-induction points in the vicinity of the given direction for $e_{31}/g_1 > 1$) as depicted in Fig. 9b.

It should be noted that the above splitting of unstable singularities is by no means reduced to abstract mathematical games. Perturbations in the material indices of real crystals are frequently caused by various external factors such as electric fields, mechanical stresses, or temperature fluctuations arising in the vicinity of phase transitions. For example, the phase transition from a crystal of the symmetry class $\bar{6}2m$ or $\bar{6}$ to a trigonal crystal of the symmetry class 32, 3*m*, or 3 leads to replacement of the hexad axis $\bar{6}$ ($n = -2$) by a triad axis ($n = 1$). In accordance with the law of the topological charge (index) conservation and with the final crystal symmetry, three additional zero-induction points $\mathbf{D}_l = 0$ with $n = -1$ must appear along with the central point (Fig. 10).

7. CONCLUSIONS

Two electric components, the electric field **E** and the electric induction **D**, accompanying a volume acoustic wave propagating in a piezoelectric medium exhibit significantly different properties. The electric field is always purely longitudinal, whereas the electric induction vector is, in contrast, always purely transverse. On the unit sphere ($\mathbf{m}^2 = 1$) of wave propagation directions, the directions of zero electric field ($\mathbf{E} = 0$) form lines, while the zero-induction directions ($\mathbf{D} = 0$) are usually isolated and appear as singular points of the tangential vector field $\mathbf{D}(\mathbf{m})$ orientations. The nonpiezoactive directions of both types exist practically in all (even triclinic) crystals, although the presence of crystal symmetry elements is the addition factor determining the appearance of such directions.

The topological singularities of the $\mathbf{D}_\alpha(\mathbf{m})$ vectors fields in the vicinity of zero-induction points in most crystals are characterized by the Poincaré indices $n = \pm 1$, where the sign coincides with that of the determinant of the $\hat{Q}_\alpha$ matrix (62). However, in some special cases, this tensor may vanish ($\hat{Q}_\alpha = 0$) in some special directions because of a high symmetry or as a result of vanishing of certain combinations of the material tensor components. In this case, the system has either an isolated zero-induction point $\mathbf{m}_0$ (and has the Poincaré indices $n = 0, \pm 2$) or a zero-induction line. Such special orientations are topologically unstable and, in response to any change in the anisotropy, either split into stable points with $n = \pm 1$ or disappear.

It is interesting to note that singularities of the induction vector field $\mathbf{D}_\alpha(\mathbf{m})$ in the vicinity of the zero-induction points substantially differ from analogous singularities near the acoustic axes. According to [7], stable singularities in the latter case are characterized by the Poincaré indices $n = \pm 1/2$, while the unstable ones have $n = 0, \pm 1$. The only exception to this rule is the acoustic axis along the hexad axis $\bar{6}$, for which the degenerate $\mathbf{D}_\alpha = 0$ branches also have $\hat{Q}_\alpha = 0$ and $n = -2$.

ACKNOWLEDGMENTS

The work of two authors (V.I.A. and V.N.L.) was supported by the Russian Foundation for Basic Research (project no. 05-02-16666).

This study was performed within the framework of the Agreement on Cooperation between the Shubnikov Institute of Crystallography (Russia) and Kielce Technical University (Poland). One of the authors (V.I.A.) gratefully acknowledges support from the Poland–Japan Institute of Information Technologies (grant no. PJ/MKT/02/2003); the work of A.R. was supported by Polish Grant no. 5T07A00322.

REFERENCES

1. E. Dieulesaint and D. Royer, *Elastic Waves in Solids* (Wiley, New York, 1981).
2. M. K. Balakirev and I. A. Gilinskiĭ, *Waves in Piezoelectric Crystals* (Nauka, Novosibirsk, 1982) [in Russian].
3. V. E. Lyamov, *Polarization Effects and Anisotropy of Interaction of Acoustic Waves in Crystals* (Moscow State University, Moscow, 1983) [in Russian].
4. Yu. V. Gulyaev, IEEE Trans. Ultrason. Ferroelectr. Freq. Control **45**, 935 (1998).
5. V. I. Al'shits and V. N. Lyubimov, Kristallografiya **35**, 1325 (1990) [Sov. Phys. Crystallogr. **35**, 780 (1990)].
6. V. I. Al'shits, A. V. Sarychev, and A. L. Shuvalov, Zh. Éksp. Teor. Fiz. **89**, 922 (1985) [Sov. Phys. JETP **62**, 531 (1985)].
7. V. I. Al'shits, V. N. Lyubimov, A. V. Sarychev, and A. L. Shuvalov, Zh. Éksp. Teor. Fiz. **93**, 723 (1987) [Sov. Phys. JETP **66**, 408 (1987)].
8. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media*, 3rd ed. (Nauka, Moscow, 1992; Pergamon, New York, 1984).
9. V. N. Lyubimov, Dokl. Akad. Nauk SSSR **186**, 1055 (1969) [Sov. Phys. Dokl. **14**, 567 (1969)].
10. Yu. I. Sirotin and M. P. Shaskolskaya, *Fundamentals of Crystal Physics* (Nauka, Moscow, 1975; Mir, Moscow, 1982).
11. F. I. Fedorov, *Theory of Elastic Waves in Crystals* (Nauka, Moscow, 1965; Plenum, New York, 1968).
12. V. I. Al'shits and J. Lothe, Kristallografiya **24**, 1122 (1979) [Sov. Phys. Crystallogr. **24**, 644 (1979)].
13. P. Holm, Phys. Scr. **44**, 122 (1992).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 11–21.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 17–29.

ATOMS, MOLECULES, OPTICS

Dynamics of an Extremely Short Pulse in a Stark Medium

S. O. Elyutin
Moscow Institute of Engineering Physics, Kashirskoe sh. 31, Moscow, 115409 Russia
e-mail: elyutin@mail.ru
Received October 25, 2004

Abstract—Effects of propagation of an extremely short (of one or several oscillation periods) electromagnetic pulse in a medium whose resonance transition is characterized by diagonal as well as nondiagonal matrix elements of the dipole moment operator have been studied numerically. The Maxwell–Bloch system of equations is employed without using the approximation of slowly varying envelopes. An analog of the McCall and Hahn area theorem is discussed as applied to the division of the initial extremely short pulse into subpulses. The solution is obtained in the form of a solitary stable bipolar signal with a nonzero pulse area (nonzero breather). © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The term extremely short pulses (ESPs) is applied to electromagnetic pulses with a duration of a few periods of field oscillations so that the electric field polarization cannot change its sign during the action of the so-called half-period pulse [1–9]. Such pulses are field energy flashes for which the concepts of carrier frequency and wavelength are inapplicable. The spectral composition of an ESP exhibits overbroadening [10, 11] and approaches white noise; under certain conditions, the ESP spectrum has a peak at zero frequency.

The approximations of slowly varying amplitudes for equations of electromagnetic field and rotating waves for a resonant medium, which are normally used in nonlinear optics of quasi-monochromatic pulses [12–15], should be at least refined to take into account higher orders in perturbation theory; this makes the application of these approximations rather cumbersome. In an approach developed during a recent discussion of problems in subpicosecond optics [4, 16–20], the dynamics of the pulse field is analyzed directly and not its envelope. It was also pointed out that not only the temporal, but also the spatial ESP spectrum must be overbroadened in view of the extremely small longitudinal size of the propagating light spot (on the order of the central wavelength) [21, 22]. In this case, a consistent theory of self-action of such electromagnetic formations must be nonparaxial.

The propagation of ESPs in a medium of two-level atoms was studied in a typical situation for many resonant media, when the diagonal elements of the dipole moment operator are equal to zero [5, 7, 17, 23–26]. However, the linear Stark effect is observed for some molecules. Parity violation for quantum states between which a transition takes place can be induced by external fields or by the action of surface forces in the case of surface adsorption of molecules. In these cases, the diagonal elements of the operator of dipole transitions between resonance levels are regarded as nonzero quantities. In the system of quantum dots, a constant dipole moment appears due to possible symmetry breaking in the shape of quantum dots during their formation. Analogously to Kerr media, the media with nonzero matrix elements of the dipole moment can be referred to as Stark media [27–29]. The interaction of an electromagnetic pulse with a duration of a few periods of field oscillation with a Stark medium, disregarding the effects of propagation, was considered in [30].

It should be noted that the application of the slowly varying amplitude approximation to the problem of propagation of an ESP through a Stark medium leads to generation of even harmonics; this renders such an approach complicated (if at all realizable). The rejection of the slowly varying amplitude approximation supplemented with the approximation of unidirectional waves renders the problem completely integrable and makes it possible to find steady-state solutions (solitons) [31, 32].

The goal here is to study the space–time ESP dynamics in transient processes of formation of solitons in a Stark medium. We assume that the ESP spectral half-width is smaller than the frequency of the preferred transition from the ground state to the nearest excited state. The dipole moments of resonant molecules are aligned so that all the matrix elements of the dipole moment operator are parallel to the linearly polarized electromagnetic field vector. The ESP duration is such that the characteristic times of all irreversible relaxation processes in a quantum system are much longer than the time scale of field variation. We assume that all atoms have the same transition frequency; consequently, reversible relaxation processes can be ignored in the framework under the above-formulated assumptions. We will also assume that the transverse size of the light spot, which is flat over the front, is

much larger than the characteristic wavelength; consequently, diffraction effects can be disregarded.

The main equations of the model will be formulated in Sections 2, 3, and 4, where numerical estimates of the parameters of the problem are obtained and the properties of steady-state solutions [20, 32] required for further analysis are discussed. The numerical results obtained in Sections 5 and 6 show that the space–time evolution of ESPs in the bulk of a Stark medium substantially depends on the sign of quantity μ, which is proportional to the difference of the diagonal matrix elements of the ground and excited states of a two-level atom. It was found that positive values of the Stark parameter are characterized by a division of the initial unipolar ESP into steady-state subpulses in accordance with regularities analogous to the McCall and Hahn area theorem [33, 34]. Numerical simulation for negative values of μ revealed the existence of a solitary stable bipolar ESP resembling a breather or a 0π pulse [35], but having a nonzero pulse area. The properties of the obtained solutions are associated with the extremely short duration of pulses incident on a medium of two-level atoms with nonzero matrix elements of the dipole moment operator.

2. EQUATIONS OF THE MODEL

In the approximation of two-level atoms (molecules), the Hamiltonian can be represented in the form of a 2×2 matrix [12]:

$$\hat{H} = \frac{\hbar\omega_0}{2}\left(\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} d_{11}E & d_{12}E \\ d_{21}E & d_{22}E \end{pmatrix}\right),$$

where E is the electric field of the electromagnetic wave, and $|1\rangle$ and $|2\rangle$ are the ground and excited state, respectively, differing in energy by $\hbar\omega_0$.

The polarizability p of an atom (molecule) is given by

$$p = \mathrm{tr}\hat{\rho}\hat{d} = \rho_{11}d_{11} + \rho_{22}d_{22} + \rho_{12}d_{21} + \rho_{21}d_{12}.$$

If we assume that all relaxation processes in the system of atoms can be ignored, the density matrix $\hat{\rho}$ satisfies the condition $\mathrm{tr}\hat{\rho} = \rho_{11} + \rho_{22} = 1$. In this case,

$$p = \frac{1}{2}(d_{11} + d_{22}) + \frac{1}{2}(d_{22} + d_{11})r_3 + d_{12}r_1, \qquad (1)$$

where the components

$$r_1 = \rho_{12} + \rho_{21}, \quad r_2 = -i(\rho_{12} - \rho_{21}),$$

$$r_3 = \rho_{22} - \rho_{11}, \quad r_1^2 + r_2^2 + r_3^2 = 1,$$

of Bloch vector $\mathbf{r}$ are introduced, and constant phases of the elements of density matrix $\hat{\rho}$ and of the dipole moment operator $\hat{d}$ are chosen so that $d_{12} = d_{21} = d$.

The scalar Maxwell equations lead to the wave equation for field $E(z, t)$ (not the envelope),

$$\frac{\partial^2 E}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = \frac{4\pi n_A}{c^2}\frac{\partial^2}{\partial t^2}\langle p\rangle,$$

where polarizability $p(z, t)$ is defined by relation (1) and the angle brackets denote summation over all atoms and division of the results by the density n_A of resonant particles.

If we assume that the electromagnetic field varies at a higher rate than the material variables $r_3(t, z)$ and $r_1(t, z)$, we can ignore the reflected wave [17, 25] and obtain the wave equation in the unidirectional approximation:

$$\frac{\partial E}{\partial z} + \frac{1}{c}\frac{\partial E}{\partial t} = -\frac{2\pi n_A}{c}\frac{\partial}{\partial t}\left\langle \frac{1}{2}(d_{22} - d_{11})r_3 + d_{12}r_1 \right\rangle. \quad (2)$$

The Neumann equations for the density matrix lead to the equations for the Bloch vector components:

$$\frac{\partial r_1}{\partial t} = -\left[\omega_0 + \frac{(d_{11} - d_{22})E}{\hbar}\right]r_2,$$

$$\frac{\partial r_2}{\partial t} = \left[\omega_0 + \frac{(d_{11} - d_{22})E}{\hbar}\right]r_1 + 2\frac{dE}{\hbar}r_3, \qquad (3)$$

$$\frac{\partial r_3}{\partial t} = -2\frac{dE}{\hbar}r_2.$$

Relations (2) and (3) constitute the complete system of model equations. This system differs from the self-induced transparency equations [33, 34] in that the expression in the brackets contains, instead of detuning from resonance, the transition frequency itself (this is a manifestation of the rejection from the rotating wave approximation); in addition, this expression contains the second term $(d_{11} - d_{22})E/\hbar$ defining the shift of resonance levels due to the high-frequency Stark effect.

In the dimensionless variables, system of equations (2) and (3) has a form convenient for numerical analysis, i.e.,

$$\frac{\partial e}{\partial \zeta} = -\frac{\partial}{\partial \tau}(r_1 - \mu r_3) = r_2,$$

$$\frac{\partial r_1}{\partial \tau} = -(1 + \mu e)r_2, \qquad (4)$$

$$\frac{\partial r_2}{\partial \tau} = (1 + \mu e)r_1 + er_3, \quad \frac{\partial r_3}{\partial \tau} = -er_2,$$

where $\tau = \omega_0(t - z/c)$, $\zeta = z/L_{abs}$, $e = 2dE/\hbar\omega_0$, $\mu = (d_{11} - d_{22})/2d$ is a parameter, and the reciprocal resonance absorption length is given by $L_{abs}^{-1} = 4\pi n_A d^2(\hbar c)^{-1}$.

In system of equations (4), we will define the condition of rest of the quantum system ($r_1(t, z) = r_2(t, z) = 0$, $r_3(t, z) = -1$) as the variables of asymptotic forms for polarization and the difference in population densities for $\tau \longrightarrow \pm\infty$, while the conditions for the field at remote instants are formulated as $e(\zeta, \tau) = 0$, $\partial e(\zeta, \tau)/\partial\tau = 0$. At the input to the medium, the pulse had the following form in most computations:

$$e_{in}(\tau) = e_p \operatorname{sech}\left(-\frac{\tau - \tau_0}{\delta}\right). \quad (5)$$

System of equations (4) describes the precession of vector $\mathbf{r}$ about the effective field vector $\mathbf{\Omega} = \{-e, 0, 1 + \mu e\}$ (Fig. 1):

$$\frac{\partial \mathbf{r}}{\partial \tau} = \mathbf{\Omega} \times \mathbf{r}. \quad (6)$$

3. NUMERICAL ESTIMATES

Choosing the transition dipole moment $d \approx 10^{-18}$ CGSE units, number density of resonant atoms $n_A \approx 10^{18}$ cm^{-3}, and the ESP duration $\delta \approx \omega_0^{-1} \approx 10^{-15}$ s, we can estimate the pulse amplitude $\mathscr{E}_p \approx \hbar/2dt_p \approx 5 \times 10^5$ CGSE units for a peak intensity $I_p = c\mathscr{E}_p^2/8\pi \approx 3 \times 10^{13}$ W/cm^2.

The characteristic time $t_c = \hbar/4\pi d^2 n_A \approx 8 \times 10^{-11}$ s for two-level systems is the time of the emergence of the dipole moment in the two-level system induced by the field of a passing pulse. During this time, the ESP traverses a distance $L_{abs} = ct$, which is approximately several centimeters, while the spatial pulse length $c\delta \approx 3 \times 10^{-5}$ cm. The reciprocal of time gives an estimate for the Rabi oscillation frequency in the characteristic field of reciprocal response of the medium, $\mathscr{E}_{char} \approx 2\pi n_A d = \hbar/2dt_c \approx 10$ CGSE units.

The unidirectional wave approximation is valid when $t_c\delta^{-1} \gg 1$ or, in other words, for low concentrations of resonant emitters and small values of dipole moments, which is obviously observed in our case. For larger dipole moments of $d \approx 10^{-16}$ CGSE units, which are typical of quantum dots, the above parameters have different values: $\mathscr{E}_p \approx 5 \times 10^3$ CGSE units, $\mathscr{E}_{char} \approx 6 \times 10^2$ CGSE units, $t_c \approx 8 \times 10^{-13}$ s, and $L_{abs} \approx 2 \times 10^{-2}$ cm. It can be seen that the condition of applicability of the unidirectional wave approximation holds for quantum dots as well, although the scales of variation of the field and polarization of the medium differ less significantly as compared to quantities characteristic of a medium formed by resonant atoms or molecules.

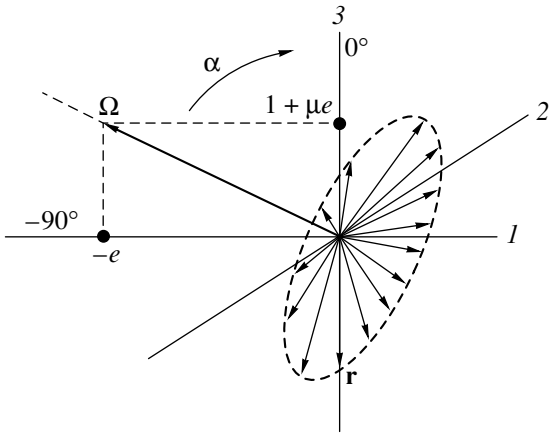

Fig. 1. Vector model.

Parameter $\mu = (d_{11} - d_{22})/2d$, which is important for subsequent analysis, may be positive or negative. The value of μ varies from 0.2 to 7.0 for some types of semiconducting quantum wells [36]; for example, $\mu \approx 0.4$ for two lower vibrational states of the ground electron state of the HeH^+ molecule [37], while $\mu \approx 1.0$ according to estimates made in [27].

4. STEADY-STATE SOLUTION

It is well known that system of equations (4), describing unidirectional propagation of an ESP in a Stark medium, is a completely integrable problem. Its steady-state solution was found in [20] in the form of a solitary wave whose velocity V obeys the inequality $V > c(1 + \tau_c^{-1})^{-1}$, which is significant for $\tau_c < 1$–10:

$$e_{st}(\tau) = 2\left\{\delta_{st}\left[\sqrt{1 + \mu^2(1 + \delta_{st}^2)}\cosh\frac{\tau - \tau_0}{\tau_{st}} + \mu\delta_{st}\right]\right\}^{-1}, \quad (7)$$

where

$$\delta_{st}^2 = \frac{\tau_c(c/V - 1)}{1 - \tau_c(c/V - 1)}$$

is the steady pulse duration.

Parameter $\tau_c = \omega_0 t_c$ has the physical meaning of the ratio of the transition energy to the energy of the

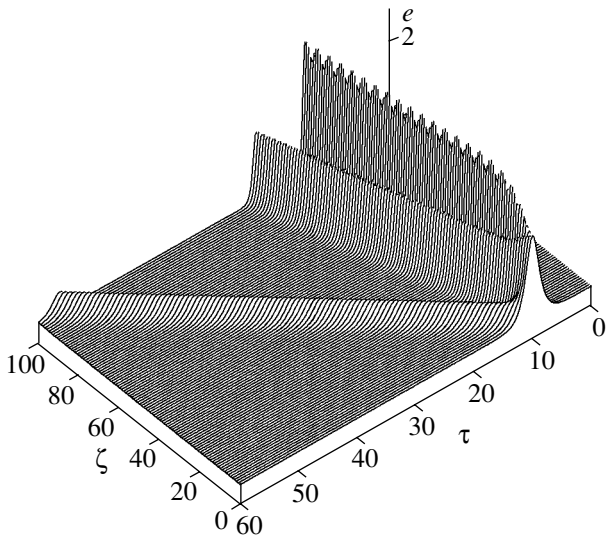

Fig. 2. Space-time field profile in the decomposition of the initial ESP into subpulses in a Stark medium with $\mu > 0$ ($\mu = 2$) with a conserved total pulse area $\Theta = 2\pi$ in the bulk of the medium.

dipole–dipole interaction between neighboring atoms. The pulse area of a steady signal Θ_{st},

$$\Theta_{st} = \int_{-\infty}^{\infty} e_{st}(\zeta, t)d\tau = \frac{8}{\sqrt{1+\mu^2}}\arctan\left\{\frac{\sqrt{1+\mu^2}}{\sqrt{1+\mu^2(1+\delta_{st}^2)}+\delta_{st}\mu}\right\}, \quad (8)$$

gives a value of 2π for parameter $\mu \longrightarrow 0$. A characteristic feature of dependence (8) is the presence of a peak in the region of small ($|\mu| < 1$) negative values of μ; e.g., $(\Theta_{st})_{max} \approx 3\pi$ for $\mu = -0.4$ and $\delta_{st} = 5$. The velocity V of propagation of a steady ESP is given by the expression

$$\frac{c}{V} = 1 + \frac{\delta_{st}^2\tau_c^{-1}}{1+\delta_{st}^2}, \quad (9)$$

which implies that the value of V as a function of δ_{st} rapidly decreases in the region of $\delta_{st} < 1$, the deceleration of the signal being the stronger, the smaller the normalized cooperative time τ_c (i.e., the smaller the response time of the medium and, hence, the stronger the reciprocal effect of the medium on the propagating pulse).

For $\mu = 0$, system (4) is transformed into a system of reduced Maxwell–Bloch equations, which are formally equivalent to the equations of self-induced transparency with a finite detuning from resonance, and solution (7) is defined by function 2πsech, which was derived by McCall and Hahn (34). It is convenient to use this property of system (4) for comparison with numerical solutions for other (nonzero) values of parameter μ.

5. TEMPORAL DYNAMICS OF ESP FOR $\mu > 0$

For a medium with a constant dipole moment ($\mu \neq 0$), the input pulse in the form of (5) is not a steady-state solution any longer; its behavior in the bulk of the medium differs from the propagation of a steady ESP. For $\mu > 0$ and for the same area of the input pulse equal to 2π, the ESP splits into subpulses (Fig. 2) propagating at their individual velocities. The steady ESPs (7) resulting from the decomposition preserve their temporal shape, which is a feature of affiliation of such objects to solitons. Analogous space–time dynamics is characteristic of the effect of self-transparency of $2\pi n$ pulses [35], in which the number of signals appearing as a result of decomposition of an input pulse multiple of 2π is determined by the area theorem. In our case, it is impossible to formulate a direct analog of the McCall and Hahn area theorem. However, in accordance with the decomposition pattern for the input pulse, which is similar in many respects to the analogous process in the self-induced transparency effect, the problem under consideration undoubtedly contains intrinsic regularities analogous to the area theorem. Here, the role of area Θ under the envelope of a coherent pulse incident on the medium, which determines the number of solitons formed in the theory of self-induced transparency, can be played by the angle Φ of rotation of the Bloch vector **r** around the effective field vector **Ω** (6) less the term emerging due to uniform precession of vector **r** about the third component of vector **Ω** (6) containing irreducible unity:

$$\Phi = \int_0^{\infty} d\tau\left(\sqrt{e(\zeta,\tau)^2 + (1+\mu e(\zeta,\tau))^2} - 1\right). \quad (10)$$

With increasing Stark parameter μ, the magnitude of the input pulse Φ (5) increases (see Fig. 10c below); accordingly, the number of decaying signals increases, which is manifested in the emergence of a larger number of spikes on the curve describing the time dependence of inversion at the entrance of an ESP in the medium (Fig. 3).

At the same time, it can be noted that the pulse area (8) of each of the increasing number of secondary steady ESPs (Fig. 4a), which are formed from the same initial pulse, decreases with increasing μ (see Fig. 4b and the explanation given in the caption to Fig. 4). The solitons formed as a result of decay of the initial pulse

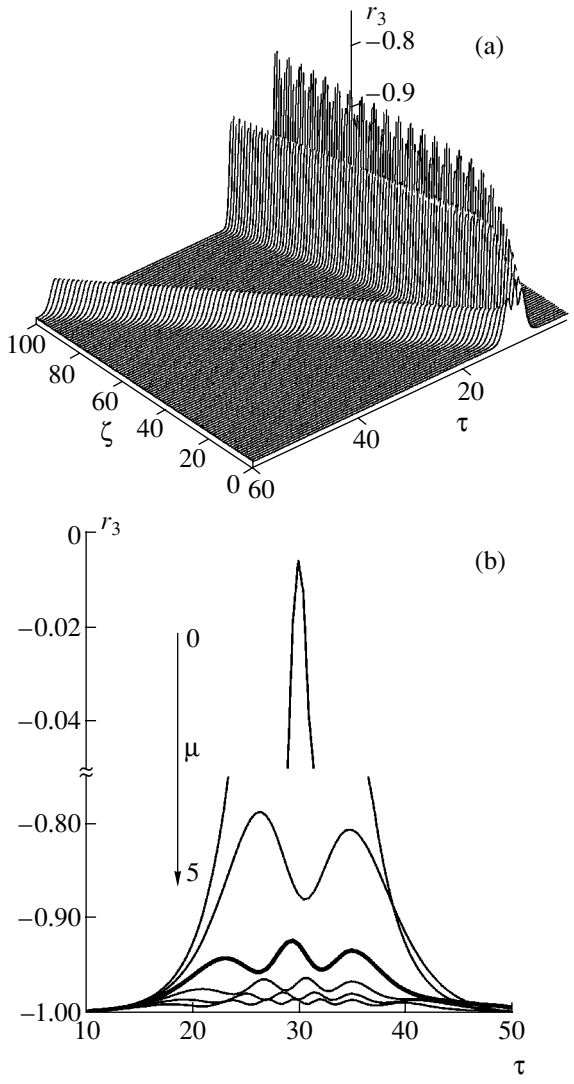

Fig. 3. (a) Space-time profile of inversion. (b) Time dependence of inversion at the entrance to the sample upon an increase in parameter μ.

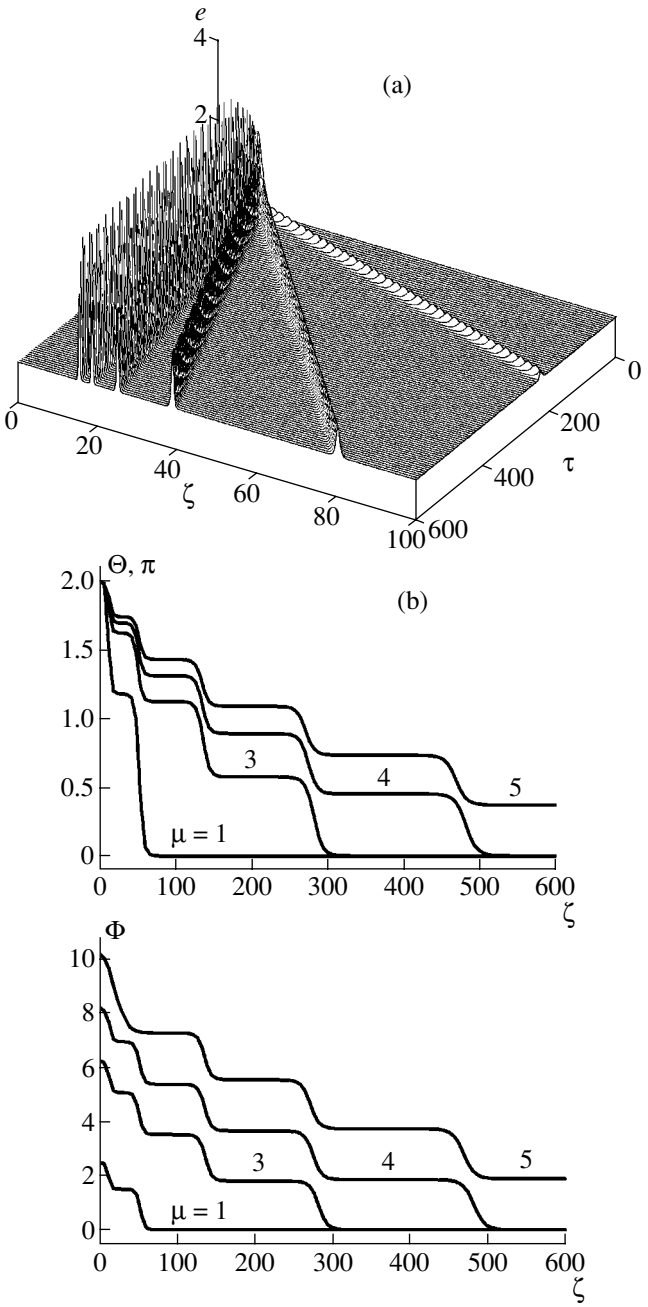

Fig. 4. (a) Decay of an ESP for large values of the Stark parameter (μ = 5). (b) Dependences of the pulse area Θ and rotation angle Φ on the spatial coordinate for large values of μ. The height of the steps is equal to the area carried away by a subpulse extending beyond the computational grid.

have different pulse areas in accordance with their steady-state durations, the shorter ESPs carrying a higher energy with a higher velocity.

In contrast to the pulse area, the rotation angle Φ in our problem can be used as a standard for measuring secondary steady ESPs since Φ_{st} is virtually independent of $\mu > 0$ (Fig. 5a) in the entire range of δ_{st} under investigation (Fig. 5b). In this case, the fractions of the initial value of Φ ($\zeta = 0$) corresponding to each steady ESP formed as a result of decay are approximately equal (see the lower curve in Fig. 4b).

Along with a 2π pulse, a breather [35] (nonlinear signal with zero pulse area, which is a bound state of two solitons with different phase velocities, but with the same group velocity) is a solitary nonstationary solution of problem (4) for $\mu = 0$.

The main analytic method for solving the Maxwell–Bloch equations is the method of inverse scattering problem (see, for example, [35]) based on the solution of the Zakharov–Shabat spectral problem [38]. Here, a breather appears as a result of emergence of two coupled ($\lambda_k^* = -\lambda_{k+1}$) values of the spectral parameter. The imaginary and real parts of parameters λ_k were interpreted as the phase velocity and amplitude of the breather components.

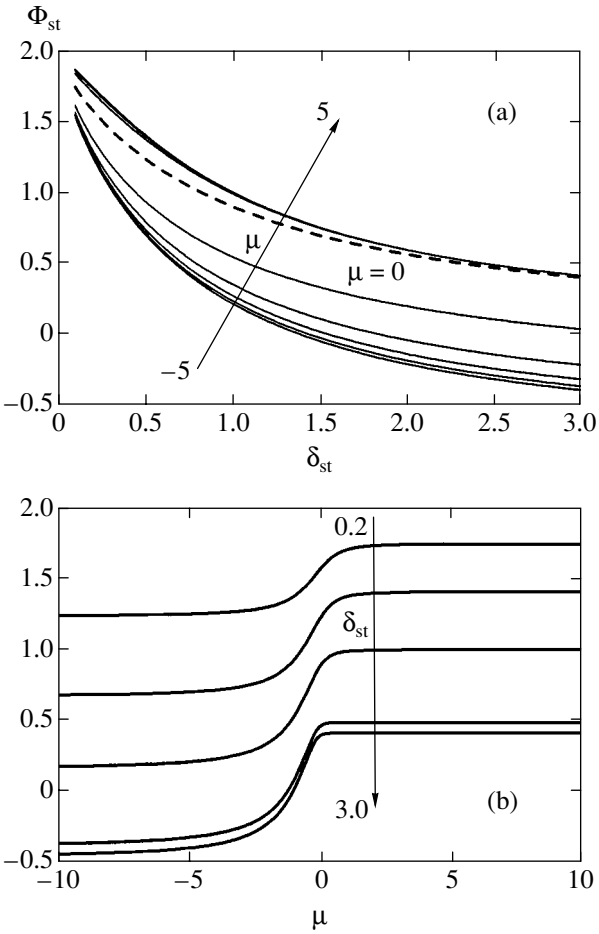

Fig. 5. (a) Dependence of the rotation angle Φ_{st} for a steady ESP on the pulse duration δ_{st}. (b) Dependence of the rotation angle Φ_{st} for a steady input ESP on the Stark parameter μ.

A breather propagating to the bulk of the medium experiences internal oscillations due to the difference in the phase velocities of the parts constituting it. This intrinsic motion leads to a periodic variation in the breather shape with increasing spatial coordinate.

However, the same breather from the problem of self-induced transparency [35], which appears at the entrance of a medium with induced polarization ($\mu \neq 0$), does not evolve in the bulk of the sample as a stable solitary ESP. Sharp variations of the polarity of the incident pulse at small depths of a resonant medium (Fig. 6a) lead to rapid rotations of the effective field vector about axis *2* (see Fig. 1), generating a nonlinear oscillatory signal (Fig. 6b) of nonzero area, which propagates at an extremely low velocity. After the emergence of this object beyond the computational grid, the only surviving pulse (see Fig. 6a) is a solitary steady unipolar ESP (7), which is gradually freed from background field oscillations.

The specific nonlinearities in equations (4) for the Bloch vector components, which are associated with Stark corrections to the transition frequency, lead to generation of spectral components supplementary to the modulation frequency in the breather, and the temporal profile of the difference in populations turns out to be modulated by the second harmonic of field oscillations in the input signal [30] (Fig. 6c).

6. NUMERICAL RESULTS FOR $\mu < 0$: NONZERO BREATHER

Sign reversal in the Stark coefficient μ leads to nontrivial consequences for the dynamics of an ESP in a resonant medium with a constant dipole moment. A unipolar ESP (5) specified at the sample surface gives rise to a nonlinear stable breather-type pulse (Fig. 7a), whose pulse area, however, differs from zero (Fig. 7b) (so-called nonzero breather). This is due to the form of the third effective field component $\Omega_3 = 1 + \mu e$ in Eqs. (4), which may vanish, or even reverse its sign, for negative values of the Stark parameter (Fig. 8a).

At small depths in the sample, when $z \sim L_{abs}(\zeta \sim 1)$, we can assume that $e(\zeta, \tau) \approx e_{in} + r_2\zeta$ (4). Essentially, the polarization component r_2 defines the field emerging in the medium as a response to the external pulse. Function $r_2(\tau, \zeta)$ at small depths is a quantity that varies at Rabi frequency Ω and generates the oscillating "tail" of the main ESP (see Fig. 7a). In the course of propagation, positive and negative spikes of the internal field distort the profile of the initial pulse, leading to the emergence of a bipolar signal (nonzero breather) in the developed phase of the process (see Fig. 7a).

At instants when the value of Ω_3 approaches zero (gray circle in Fig. 8a), vector **Ω** is directed almost against axis *1* (see Fig. 1) and forms an angle of $\alpha \approx -90°$ with axis *3* (black circle in Fig. 8b). In this position, precession of vector **r** about vector **Ω** occurs in the *32* plane (see Fig. 1), leading at the initial stage to almost complete inversion of the resonant medium (black circle in Fig. 8c) and, accordingly, to the effective interaction between the field and the medium, followed by reemission of energy back to the ESP field (as in the propagation of a 2π self-induced transparency pulse). However, over a time period following the instant marked by the light circle in Fig. 8, polarization r_2 exhibits a sharp dip of negative polarity (Fig. 8d), leading to the emergence of a negative half-period in the field profile (see Fig. 8a). In this case, in view of peculiar features of the vector model for system of equations (4), vector **Ω** makes a complete turn about axis *2* so that the angle of its inclination to axis *3* becomes positive (see Fig. 8b). For the new position of **Ω**, precession of vector **r** reverses its direction and the decrease in inversion is transformed into a new spike followed by energy absorption and reemission.

In view of nonlinear features of the problem, the swinging of the self-consistent system "quantum medium + classical pulsed field" illustrated above leads to the formation of a clearly manifested bipolar signal

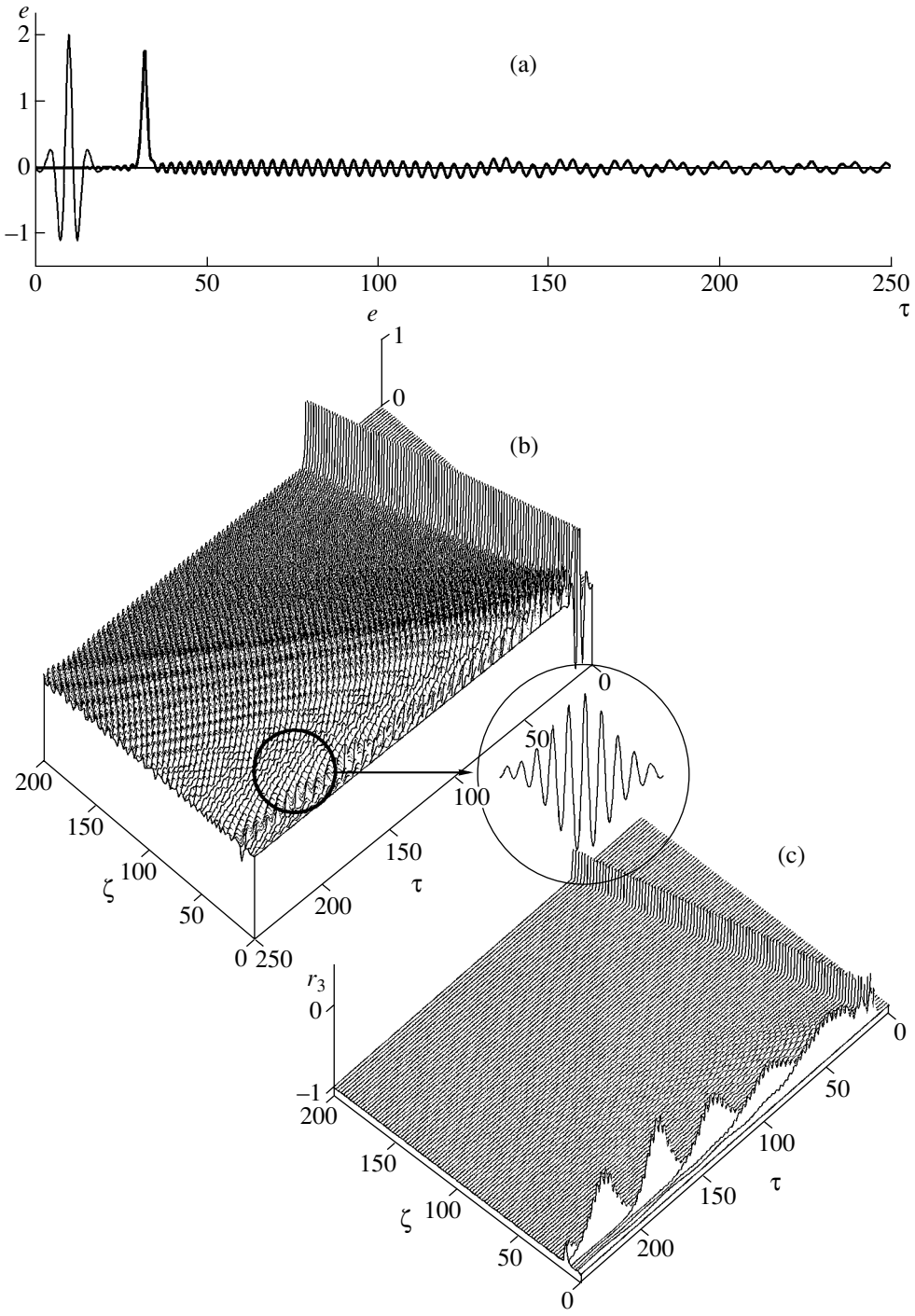

Fig. 6. (a) Formation of a solitary unipolar signal (bold curve) from the input 0π pulse (breather) (fine curve) for μ > 0 (μ = 1). (b) Temporal field profile at a depth ζ = 10 for μ = 1. (c) Generation of harmonics of an input small-period breather in the bulk of a Stark medium with μ > 0.

from the initially unipolar ESP. The increase in parameter μ in the region of negative values is accompanied by periodic variations in the decay of the initial ESP (Fig. 9). At the initial stage of each cycle, the above-mentioned nonzero breather (Fig. 9a) with a small period T_{phase} of variation of its phase is formed against the background of the oscillatory noise (Fig. 10a). This bipolar pulse propagates at its own velocity U differing from the velocity of steady signal (7); the value of U can be smaller than $0.2c$ (Fig. 10b). With increasing absolute value of μ, period T_{phase} of internal oscillations in the nonzero breather increases (see Figs. 9 and 10a),

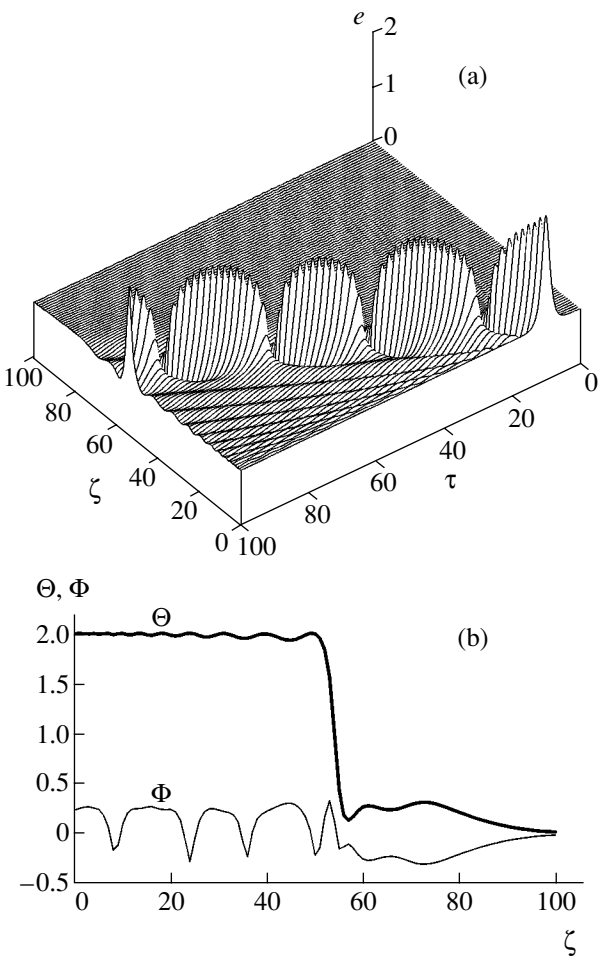

Fig. 7. (a) Formation of a nonzero breather from a unipolar input ESP in the Stark medium with $\mu < 0$ ($\mu = -1$). (b) The $\Theta(\zeta)$ curve (in units of π) shows that the pulse area of the bipolar signal (Fig. 7a) is approximately equal to 2π; spatial periodicity of the rotation angle $\Phi(\zeta)$ corresponds to penetration depths at which negative field spikes appear in Fig. 7a.

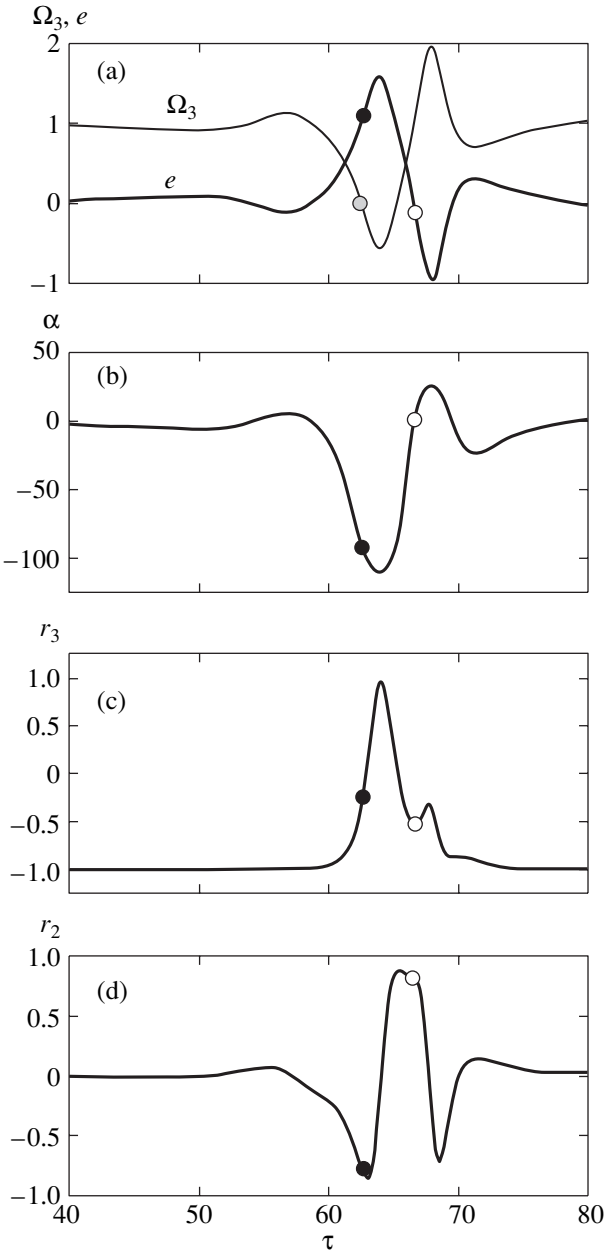

Fig. 8. (a) Time dependence of the field strength and the third component of effective field Ω_3 in the bulk of a resonant medium with $\mu = -1$. Time dependence of the angle α of rotation of effective field vector $\mathbf{\Omega}$ (b), inversion r_3 (c), and medium polarization r_2 (d) at a distance $\zeta = 35$ in the bulk of the medium.

which ultimately leads to the emergence of a steady unipolar signal (7) with $T_{\text{phase}} \longrightarrow \infty$. With a further decrease in μ, this signal gradually joins the package of steady subpulses (Fig. 9c) that have been formed at the beginning of propagation to the bulk of the medium. The number and parameters of these decay signals is determined, as in the case of $\mu > 0$, by the value of the total precession angle Φ ($\zeta = 0$) at the entrance of the medium, which increases with decreasing μ (Fig. 10c).

The oscillatory object discovered in calculations for $\mu < 0$ (see Figs. 7–9) can be represented numerically as the sum of two pulses propagating at the same group velocity, viz., a steady signal with a preset μ (7) and a bipolar signal (Fig. 11) with a pulse area equal to zero to a high degree of accuracy.

It should be noted that an adequate explanation of the ESP evolution in the bulk of a Stark medium is contained in the spectral structure of the inverse scattering problem for Eqs. (4) (the derivation and analysis of this spectrum forms an individual nontrivial problem). However, the representation of the above bipolar ESP as the sum of two signals (Fig. 11) makes it possible to qualitatively describe the processes illustrated in Figs. 7–9.

As parameter μ increases to the region of negative values, a bound state of three pulses propagating at the same group velocity (which is much smaller than the

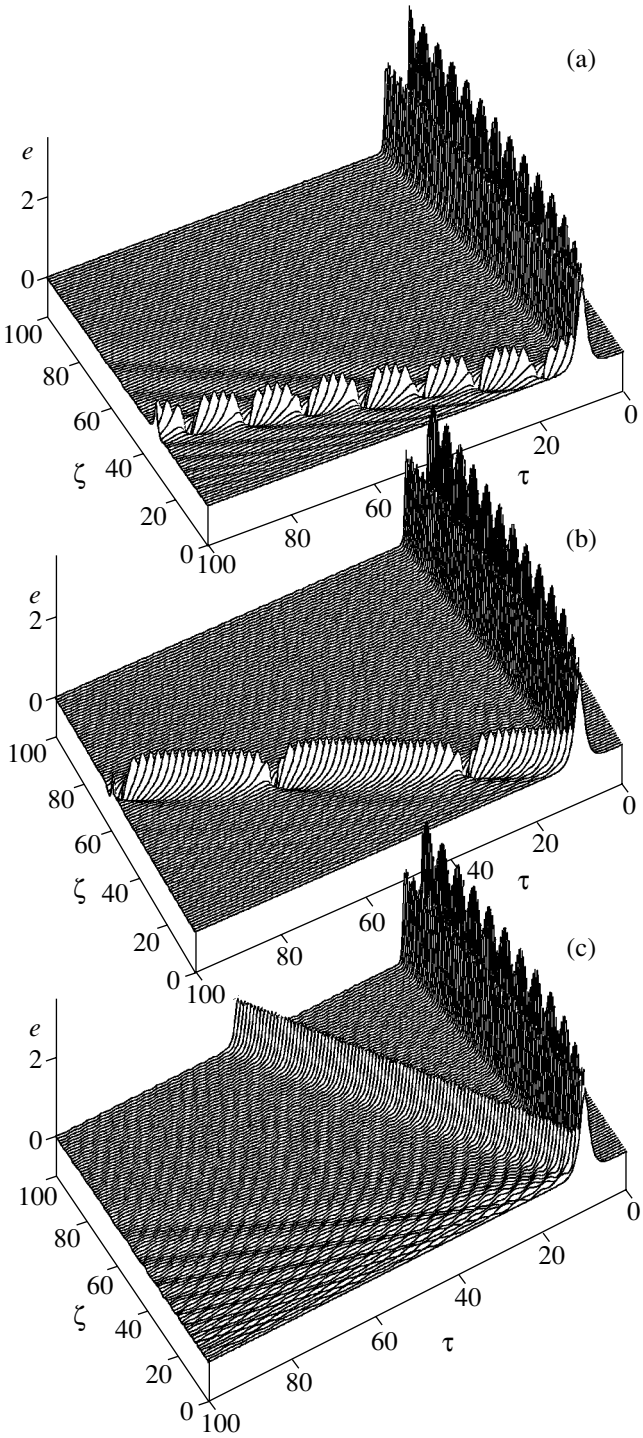

Fig. 9. Formation of a unipolar steady pulse from a bipolar nonzero breather with increasing μ in the region of negative values: μ = –3.78 (a), –3.85 (b), and –4.0 (c).

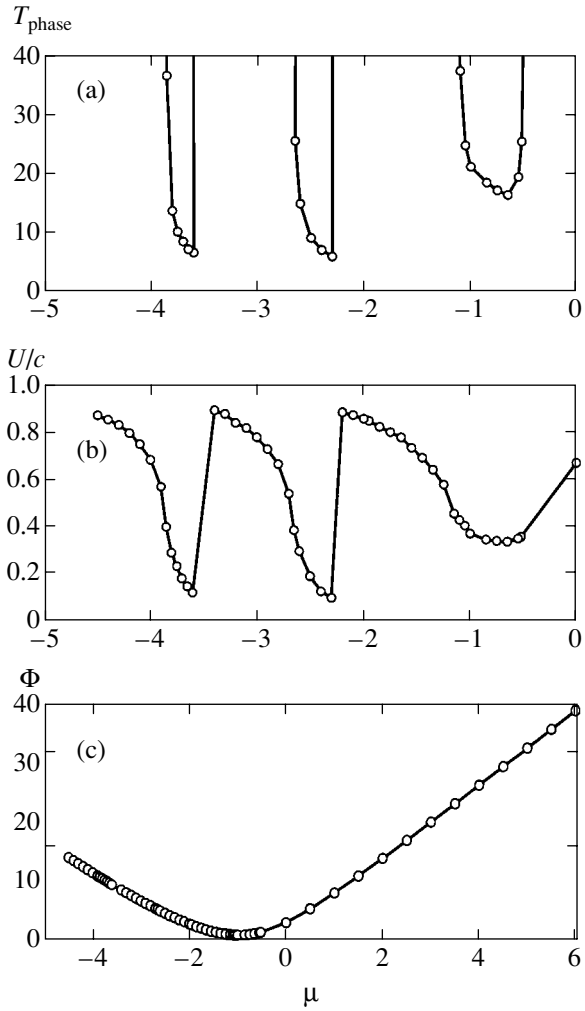

Fig. 10. Dependence (a) of period T_{phase} of phase change of a nonzero breather and (b) velocity U of its propagation on parameter μ < 0. (c) Dependence of angle Φ (ζ = 0) of rotation of the Bloch vector relative to the effective field vector for an ESP at the entrance to the medium on μ.

velocity of light) is formed from the boundary temporal profile even at small distances (see Figs. 7 and 9). Two components from this triad have equal and opposite phase velocities and form a bipolar 0π pulse (see Fig. 11). The third pulse is a steady ESP (7). As the Stark parameter decreases, the difference in the phase velocities decreases and, accordingly, the rate of phase change in the temporal profile of the observed bipolar signal slows down (see Fig. 9). In the limit of zero difference in the phase velocities, the two pulses with opposite polarities coincide on the time scale and are annihilated. As a result, we are left with a steady signal (see Fig. 9c) whose velocity is the lower, the smaller its duration. With decreasing μ, the process of generation and annihilation of the bipolar signal is repeated with increasing frequency (see Figs. 10a and 10b).

It should be noted that the above oscillatory signal with a nonzero area corresponding to negative values of the Stark parameter μ is also formed in the case when a Lamb [35] 0π signal is specified at the entrance to the medium (Fig. 12). Variation of μ and the frequency of field oscillations in the initial breather in this process gives rise to a large number of peculiar features, one of which corresponds to a negative-polarity ESP with a

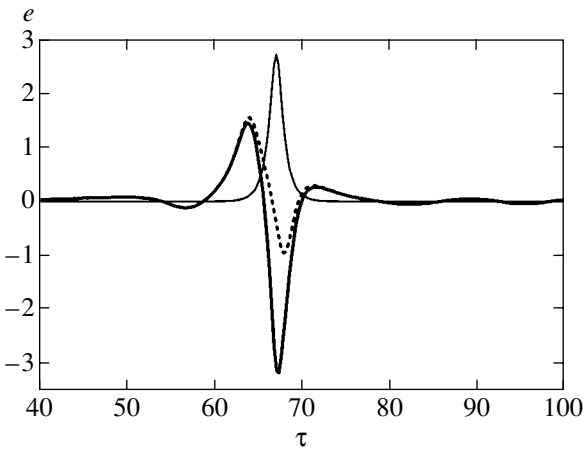

Fig. 11. Nonzero breather as the sum of steady signal (7) and 0π pulse for $\mu = -1$, $\zeta = 35$. The bold curve corresponds to $e(\zeta = 35, \tau)$, i.e., a bipolar pulse of a nonzero area; the fine curve corresponds to $e_{st}(\zeta = 35, \tau)$, i.e., a steady ESP of duration $\delta_{st} = 1$; the dashed curve corresponds to $e - e_{st}$, i.e., a bipolar pulse of zero area (0π pulse).

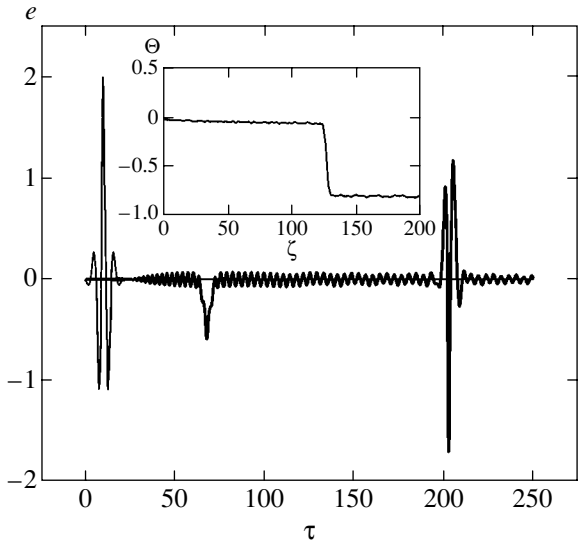

Fig. 12. Emergence of a negative-polarity ESP and a nonzero breather from the initial 0π pulse (breather) at large depths of penetration to a Stark medium with $\mu = -1$: $\zeta = 0$ (fine curve) and 100 (bold curve). The pulse area of the nonzero breather (inset) is equal to π.

pulse area compensating the area of the "nonzero" breather (see Fig. 12).

7. CONCLUSIONS

We have analyzed numerically the space–time dynamics of extremely short pulses in a medium in which the transition between two quantum states is characterized by a dipole moment operator with nonzero diagonal matrix elements (Stark medium). The assumption on extremely small pulse duration leads to the rejection of concepts of slow dependent variables for the field and the medium, their frequencies and phases, as well as the formulation of the Maxwell–Bloch model in terms of the self-consistent field itself with a definite polarity and true (not approximate) rotating waves (density matrix components). As a result, Bloch equations (4) acquire (in the region of frequency detuning) cofactors exhibiting a threshold dependence on the magnitude and sign of the field and Stark parameter μ. It is this feature of the model that determines the main aspects of time evolution of the ESP. For example, the pulse area cannot be used for predicting the number of subpulses emerging from the intense input signal as was done in the case of self-induced transparency. To formulate a rule analogous to the area theorem, the angle of rotation of the Bloch vector during its precession about the effective field vector should rather be used instead as a parameter. It should also be noted that each time the amplitude (and not the envelope) of the input field e_p tends to $1/\mu$ for $\mu < 0$, the value of effective detuning Ω_3 tends to zero and a stable solitary bipolar ESP with a nonzero pulse area is formed even for small depths of the medium. The electromagnetic object observed in Figs. 7–10 was derived numerically, and analytic properties of this solution are not known as yet; however, simulation reveals a clearly manifested stability of the "nonzero" breather. This stability is manifested in the fact that the bipolar signal withstands collisions with a steady-state solution of type (7) (with a similar signal of other periodicity) without decaying and penetrates through the region of a strong modulated perturbation as a single entity. These are features of soliton solutions, although the final answer to the questions whether or not the observed bipolar pulse is a soliton can be obtained only from an analysis of the spectral problem in the method of inverse scattering problem for the model under investigation. In this connection, it is appropriate to mention the work [39], in which the eigenvalues of the Zakharov–Shabat spectral problem, which are parametric functions of the pulse duty factor of the external field, were calculated under the conditions of two-pulse excitation typical of photon echo problem. Subsequent simulation [40, 41] proved that the eigenvalues appear in triads in the form of an anticonjugate pair giving a breather and a solitary root responsible for a 2π pulse (soliton). An increase in the pulse duty ratio led to a decrease in the difference in the phase velocities of the breather components and to the transformation of the breather to a pair of ±2π pulses. The obvious similarity of this pattern with the results of calculations represented in Figs. 7–11 suggests that similar approaches can be used for interpreting the results within an analysis of the properties of the spectral problem, which were obtained in [20, 31].

ACKNOWLEDGMENTS

The author is grateful to his colleagues A.I. Maĭmistov, A.M. Basharov, and É.A. Manykin for fruitful discussions of the problems considered in this work.

This study was supported by the Russian Foundation for Basic Research (project no. 03-02-16979).

REFERENCES

1. A. E. Kaplan and P. L. Shkolnikov, Phys. Rev. Lett. **75**, 2316 (1995).
2. A. E. Kaplan, S. F. Straub, and P. L. Shkolnikov, J. Opt. Soc. Am. B **14**, 3013 (1997).
3. L. Gilles, J. V. Moloney, and L. Vázquez, Phys. Rev. E **60**, 1051 (1999).
4. A. B. Shvartsburg, Usp. Fiz. Nauk **168**, 85 (1998) [Phys. Usp. **41**, 77 (1998)].
5. N. Bloembergen, Rev. Mod. Phys. **71**, S283 (1999).
6. Th. Brabec and F. Krausz, Rev. Mod. Phys. **72**, 545 (2000).
7. A. I. Maĭmistov, Kvantovaya Élektron. (Moscow) **24**, 963 (1997).
8. K. Akimoto, J. Phys. Soc. Jpn. **65**, 2020 (1996).
9. S. V. Sazonov and E. V. Trifonov, J. Phys. B **27**, L7 (1994).
10. S. L. Chin, A. Brodeur, S. Petit, *et al.*, J. Nonlinear Opt. Phys. Mater. **8**, 121 (1999).
11. S. A. Kozlov, in *Problems in Coherent and Nonlinear Optics*, Ed. by I. P. Gurov and S. A. Kozlov (S.-Peterb. Inst. Tochn. Mekh. Opt., St. Petersburg, 2000), p. 136 [in Russian].
12. A. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (Wiley, New York, 1975; Atomizdat, Moscow, 1978).
13. S. A. Akhmanov, V. A. Vysloukh, and A. S. Chirkin, *The Optics of Femtosecond Laser Pulses* (Nauka, Moscow, 1988) [in Russian].
14. A. I. Maimistov and A. M. Basharov, *Nonlinear Optical Waves* (Kluwer Academic, Dordrecht, 1999).
15. A. I. Maimistov, A. M. Basharov, S. O. Elyutin, and Yu. M. Sklyarov, Phys. Rep. **191**, 1 (1990).
16. V. G. Bespalov, S. A. Kozlov, Yu. A. Shpolyansky, and I. A. Walmsley, Phys. Rev. A **66**, 013811 (2002).
17. A. I. Maĭmistov, Kvantovaya Élektron. (Moscow) **30**, 287 (2000).
18. S. A. Kozlov and S. V. Sazonov, Zh. Éksp. Teor. Fiz. **111**, 404 (1997) [JETP **84**, 221 (1997)].
19. V. N. Serkin, É. M. Shmidt, T. L. Belyaeva, *et al.*, Kvantovaya Élektron. (Moscow) **24**, 923 (1997).
20. J.-G. Caputo and A. I. Maimistov, Phys. Lett. A **296**, 34 (2002).
21. A. N. Berkovskiĭ, S. A. Kozlov, and Yu. A. Shpolyanskiĭ, Opt. Zh. **69** (3), 35 (2002) [J. Opt. Technol. **69**, 163 (2002)].
22. N. N. Rozanov, in *Problems in Coherent and Nonlinear Optics*, Ed. by I. P. Gurov and S. A. Kozlov (S.-Peterb. Inst. Tochn. Mekh. Opt., St. Petersburg, 2000), p. 53 [in Russian].
23. R. K. Bullough and F. Ahmad, Phys. Rev. Lett. **27**, 330 (1971).
24. R. K. Bullough, P. M. Jack, P. W. Kitchenside, and R. Saunders, Phys. Scr. **20**, 364 (1979).
25. J. C. Eilbeck, J. D. Gibbon, P. L. Caundrey, and R. K. Bullough, J. Phys. A **6**, 1337 (1973).
26. M. Kaluža and J. T. Muckerman, Phys. Rev. A **51**, 1694 (1995).
27. R. Bavli and Y. B. Band, Phys. Rev. A **43**, 5039 (1991).
28. G. F. Thomas, Phys. Rev. A **33**, 1033 (1986).
29. A. Brown and W. J. Meath, Phys. Rev. A **63**, 013403 (2001).
30. L. W. Casperson, Phys. Rev. A **57**, 609 (1998).
31. M. Agrotis, N. M. Ercolani, S. A. Glasgow, and J. V. Moloney, Physica D (Amsterdam) **138**, 134 (2000).
32. A. I. Maĭmistov and J.-G. Caputo, Opt. Spektrosk. **94**, 271 (2003) [Opt. Spectrosc. **94**, 245 (2003)].
33. S. L. McCall and E. L. Hahn, Phys. Rev. Lett. **18**, 908 (1967).
34. S. L. McCall and E. L. Hahn, Phys. Rev. **183**, 457 (1969).
35. G. L. Lamb, Jr., Phys. Rev. A **9**, 422 (1974).
36. S. Kočinac, Z. Ikonić, and V. Milanović, Opt. Commun. **140**, 89 (1997).
37. S. Nakai and W. J. Meath, J. Chem. Phys. **96**, 4991 (1992).
38. V. E. Zakharov and A. B. Shabat, Zh. Éksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys. JETP **34**, 62 (1971)].
39. V. E. Zakharov and É. A. Manykin, Zh. Éksp. Teor. Fiz. **82**, 379 (1982) [Sov. Phys. JETP **55**, 227 (1982)].
40. S. O. Elyutin and A. I. Maĭmistov, Opt. Spektrosk. **61**, 1058 (1986) [Opt. Spectrosc. **61**, 661 (1986)].
41. S. Elyutin, A. Maimistov, E. Manykin, and Yu. Sklyarov, Phys. Lett. A **142**, 493 (1989).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 120–127.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 139–149.

ELECTRONIC PROPERTIES OF SOLIDS

Effect of Coulomb Explosion on Track Formation in Metals Irradiated by Heavy Ions

A. I. Ryazanov, S. A. Pavlov, E. V. Metelkin, and A. V. Zhemerev
Russian Research Center Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia
e-mail: ryazanoff@comail.ru
Received July 5, 2004

Abstract—The existing descriptions of the process of track formation in metals bombarded with high-energy heavy ions are usually based on the T-spike model. However, this model fails to explain high temperatures developed in some target materials (e.g., in copper). We present a comparative analysis of track formation in metals under the action of heavy ion bombardment, as described using the Coulomb explosion model and the T-spike model in various modifications. Both models are used to calculate changes in the temperatures of the electron and ion subsystems in the track region in amorphous alloy $Fe_{85}B_{15}$) and copper targets bombarded with identical high-energy ($E > 1$ MeV/nucleon) heavy ions. The results show that the Coulomb explosion model predicts stronger heating of the ion subsystem in the track region (with the possible formation of a melt-through zone) as compared to the T-spike model. The formation of point defects in copper as a result of the ionization losses has been also described using the two models and compared to the available experimental data. The Coulomb explosion model provides for a more adequate description of track formation as compared to the T-spike model. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

When heavy charged particles with energies $E >$ 1 MeV/nucleon are stopped in a target, more than 90% of the particle energy is transferred to the electron subsystem as a result of the electron drag and only about 10% is transferred directly to atoms of the crystal lattice via their elastic collisions with the projectiles. Subsequent heating of the material is related to the energy transfer from electrons to ions via processes described within the framework of the thermal spike (T-spike) model [1, 2] and the Coulomb explosion (or ion explosion) model [3]. As a result, the temperature of the ion subsystem may increase to a significant level, reaching the melting point for metals and the crystallization temperature for amorphous alloys.

The bombardment with heavy ions at a high degree of ionization losses may also lead to the production of point defects with anomalously high concentration [4]. This process is determined predominantly by the ionization and electronic stopping power, rather than by the elastic collisions of heavy impinging ions with atoms of the crystal lattice. The contribution of such elastic collisions to the point defect formation is several orders of magnitude lower than the contribution due to the ionization losses. Therefore, the physical properties of metals in the track region significantly differ from those in the regions far from tracks.

Let us briefly consider the physical mechanisms involved in the heating of electron and ion subsystems of a target penetrated by a heavy charged particle, as described using the T-spike model and the Coulomb explosion model.

The essence of the T-spike model consists in the fact that a high-energy heavy charged particle penetrating through a target material exhibits electronic stopping power and transfers more than 90% of its kinetic energy to the electron subsystem, while only a small remaining energy is transferred directly to the ion subsystem (i.e., to the crystal lattice) via the elastic collisions of host atoms with the projectile. Thus, an effective energy source localized at the track axis operates in the electron subsystem, which leads (via electron thermal conductivity) to heating of the electron subsystem at distances greater than the track size (the characteristic thermalization time for the electron subsystem in a metal is on the order of 10^{-15} s). Subsequently, electrons give part of their energy to ions via electron–phonon interactions (the characteristic time of these interactions in metals is on the order of 10^{-13} s), which results in heating of the ion subsystem. Thus, electrons are thermalized within a period of time that is much shorter than the characteristic time of heat exchange between electron and ion subsystem.

There are various modifications of the T-spike model [1, 2], which differ from one another by the mechanisms of heating of the electron subsystem and by the expressions used to describe the electron–phonon interactions (exchange terms) in the heat balance equation. According to Yasui [2], the effective energy source in the electron subsystem plays the role of the initial condition for the electron temperature, while the heat balance equation contains only an

exchange term related to the heat transfer between the electron and ion subsystems. Dufour [1] assumed that the effective energy source operates over the entire period of thermalization of the electron subsystem and introduced the corresponding additional term into the heat balance equation.

Now let us briefly consider the mechanism of heating of the ion and electron subsystems within the framework of the Coulomb explosion model, which is used in this study for description of this process. The model proper and the conditions of its applicability were described previously [3]. Here, we only consider the physical processes leading to additional heating of the ion and electron subsystems according to this model. A high-energy heavy charged particle penetrating through a target material collides with so-called δ-electrons, which travel predominantly in the radial directions relative to the heavy-ion trajectory (track axis). As a result, an effective electric field is generated, which exists over a characteristic period of time equal to the δ-electron current lifetime. The magnitude and spatial distribution of this electric field were determined in [3]. Ions of the crystal lattice and electrons of the target material occurring within the zone of action of the effective field acquire additional momenta and, hence, additional energies. Thus, the model provides a mechanism accounting for an increase in the average kinetic energies and, eventually, the initial temperatures of both electrons and ions at the expense of energy received by these particles during their short-time interaction with the effective electric field of δ-electrons.

This paper is devoted to a comparative study of heating of the electron and ion subsystems in the track region of amorphous metal ($Fe_{85}B_{15}$) and copper targets bombarded with high-energy heavy ions in the frameworks of the T-spike and Coulomb explosion models. The obtained nonlinear equations were numerically solved by methods of implicit scheme and sequential trials, which have proved to be sufficiently robust. Examples of using these methods for solving analogous problems can be found in [1, 8].

The contributions due to different mechanisms of heating of the ion and electron subsystems, which are employed in various modifications of the T-spike model, are described by introducing various exchange terms into the heat balance equation for electrons and ions [1, 2]. Below, the significance of these mechanisms is analyzed based on the results of numerical calculations carried out for amorphous alloy $Fe_{85}B_{15}$ within the framework of the corresponding models. The results obtained using various modifications of the T-spike model showed that, depending on the parameters of irradiation, the temperature of the ion subsystem in the alloy studied may increase to a significant level, reaching the temperatures of melting and crystallization. As a result, a melt-through zone can form at the heavy-ion trajectory in the alloy target (see also [1]).

The results of application of the T-spike model to the calculations of heating of the ion subsystem in copper showed that the ion temperature cannot exhibit sharp jumps and increases by no more than 100 K, never reaching the melting point (see also [2]). In contrast, numerical calculations using the Coulomb explosion model showed that the temperature of the ion subsystem may reach the melting point and that a melt-through zone can form in the track region.

2. USE OF THE T-SPIKE MODEL FOR STUDYING CHANGES IN THE TEMPERATURES OF THE ELECTRON AND ION SUBSYSTEMS IN METAL TARGETS BOMBARDED BY HIGH-ENERGY HEAVY IONS

Let us formulate a system of nonlinear equations describing the heat transfer in the electron and ion subsystems in a metal target near a track produced by a high-energy heavy charged particle, adopting a cylindrical geometry of the temperature distribution relative to the particle trajectory (track axis). Taking into account various modifications of the T-spike model developed previously [1, 2], the heat conduction equations for the electron and ion subsystems can be written as follows:

$$C_e(T_e)\frac{\partial T_e}{\partial t} = \mathrm{div}(K_e(T_e)\nabla T_e) - \gamma(T_e)[T_e - T_i] + A(r, t), \quad (1)$$

$$C_i(T_i)\frac{\partial T_i}{\partial t} = \mathrm{div}(K_i(T_i)\nabla T_i) + \gamma[T_e - T_i], \quad (2)$$

where T_e and T_i are the temperatures of the electron and ion subsystems in the metal target, respectively; C_e, K_e, C_i, K_i are the heat capacities and thermal conductivities of electrons and ions, respectively, which are nonlinear functions of the temperature; r is the distance from the track axis; $\gamma[T_e - T_i]$ is the exchange term, which takes into account heat transfer due to the electron–phonon interaction between the electron and ion subsystems in the ion-irradiated material; and $A(r, t)$ is the effective source of energy in the electron subsystem, which is related to the electronic stopping power of the high-energy heavy particle stopped in the metal target.

The nonlinear temperature dependences of the heat capacities and thermal conductivities of electrons and ions for the copper and $Fe_{85}B_{15}$ targets are presented in the Appendix.

First, let us consider a modification of the T-spike model proposed in [2]. In this variant, the coefficient γ in the exchange term, which describes the electron–phonon interaction in the heat balance equation, appears as $\gamma = C_e/\tau$, where τ is the characteristic time of the electron–phonon interaction ($\tau = 10^{-13}$ s).

The boundary conditions for the electron and ion subsystems in the regions far from the track reflect the absence of irradiation-induced heating in these regions:

$$T_e|_{r\to\infty} = T_i|_{r\to\infty} = T_{matr},$$

where T_{matr} = 10 K is the temperature of the metal matrix.

The initial condition for the temperature distribution in the electron subsystem reflects the electronic stopping power for heavy ions [2]:

$$E_e(t=0) = \frac{Q}{4\pi\sigma^2}\exp\left(-\frac{r^2}{4\sigma^2}\right) + \int_0^{10\text{ K}} C_e dT, \quad (3)$$

$$\sigma^2 = \left(\frac{L}{2\ln 2}\right), \quad (4)$$

where E_e is the volume electron energy density, Q are the electronic stopping power per unit ion range for a high-energy heavy charged particle; L is the halfwidth of the spatial distribution of energy deposited in the electron subsystem (we adopt L = 2 nm [2]). On the other hand, the volume electron energy density at the initial moment can be expressed as

$$E_e(t=0) = \int_0^{T_e(t=0)} C_e dT, \quad (5)$$

where $T_e(t = 0)$ is the initial temperature of the electron subsystem. Equating the electron energy densities given by expressions (3) and (5), we obtain a relation for determining the initial temperature of the electron subsystem in a target penetrated by a high-energy heavy charged particle:

$$\int_0^{T_e(t=0)} C_e dT = \frac{Q}{4\pi\sigma^2}\exp\left(-\frac{r^2}{4\sigma^2}\right) + \int_0^{10\text{ K}} C_e dT.$$

The solution of this equation gives the initial condition for the temperature of the electron subsystem. The initial temperature of the ion subsystem coincides with the target material temperature T_{matr} in the absence of ion irradiation:

$$T_i(t=0) = T_{matr}.$$

There are two additional boundary conditions in both variants of the T-spike model employed in this study:

$$\left.\frac{\partial T_e}{\partial t}\right|_{r=0} = \left.\frac{\partial T_i}{\partial t}\right|_{r=0} = 0,$$

which correspond to the absence of heat fluxes at the track axis.

Now let us consider another modification of the T-spike model proposed in [1]. In this variant, the coefficient γ in the exchange term is assumed to be constant, $\gamma = g$, and is not varied in the course of calculations. In our calculation, $g = 5 \times 10^{12}$ W/(K cm^3) (see [1, 8]). Therefore, the term describing the electron–phonon interaction is proportional to the difference between the temperatures of the electron and ion subsystems and does not take into account the nonlinear temperature dependence of the amount of heat transferred between these subsystems. In the variant proposed in [2], this nonlinear dependence is taken into account via the temperature dependence of the coefficient g in the electron–phonon exchange term.

The boundary and initial conditions for the temperatures of the electron and ion subsystems in the regions far from the track are written as follows:

$$T_e|_{r\to\infty} = T_i|_{r\to\infty} = 100\text{ K},$$

$$T_e(t=0) = T_i(t=0) = 100\text{ K},$$

where T_{matr} = 100 K is the matrix temperature adopted in this model.

Heating of the electron and, accordingly, of the ion subsystems is provided by the presence of the effective energy source $A(r, t)$, related to the electron-drag losses of a high-energy heavy charged particle propagating through the matrix, in the equation for the electron temperature distribution (1). The main difference of this T-spike model variant from that proposed in [2] is that the electron subsystem is assumed to be gradually heated over the time $t_0 = 10^{-15}$ s (the characteristic time of thermalization of the electron subsystem), followed by cooling. Thus, despite the fact that the electron thermalization time is significantly shorter than the characteristic electron–phonon interaction time, the model proposed in [1] takes into account the fact that the electron subsystem is also gradually heated together with the ion subsystem. The effective energy source $A(r, t)$ in this model is as follows [1]:

$$A(r,t) = \begin{cases} C_1\left(\frac{dE}{dz}\right)_e \exp\left(-\frac{r}{r_0} - \frac{(t-t_0)^2}{2\sigma_t^2}\right) \text{ at } t < 2t_0, \\ 0 \text{ at } t > 2t_0, \end{cases} \quad (6)$$

where $(dE/dz)_e$ are the electronic stopping power of the high-energy heavy charged particle; $\sigma_t = t_0$; r_0 = 2.5 nm is the characteristic distance determining the size of the region of energy evolution in the electron subsystem; and C_1 is a constant determined from the following normalization condition [1]:

$$2\pi C_1 \int_0^\infty r dr \int_0^{2t_0} dt \exp\left(-\frac{r}{r_0} - \frac{(t-t_0)^2}{2\sigma_t^2}\right) = 1.$$

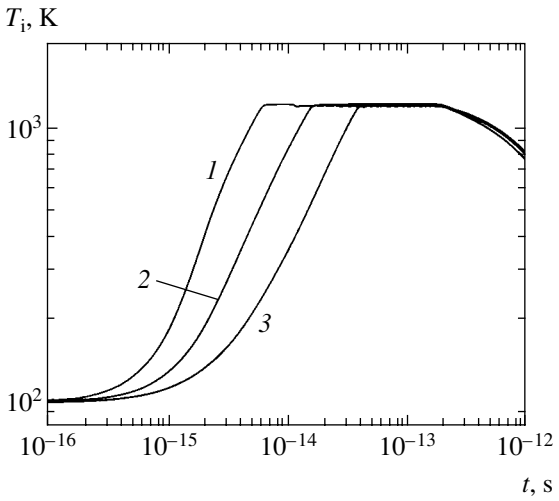

Fig. 1. Time variation of the ion subsystem temperature in the track region of a $Fe_{85}B_{15}$ amorphous metal alloy target irradiated by high-energy heavy charged particles. The calculations were performed using the modified T-spike model [1] for various distances from the track axis $r = 0$ (*1*), 5 (*2*), and 10 nm (*3*).

Let us consider whether it is expedient to use a T-spike model with allowance for the gradual heating of the electron subsystem and for the nonlinear exchange term in the heat balance equation.

Figure1 presents the results of numerical calculations of the time variation of the ion subsystem temperature for the $Fe_{85}B_{15}$ amorphous alloy, which were performed using the T-spike model [1] with (i) the electronic stopping power $(dE/dz)_e = 40$ keV/nm and (ii) the electron subsystem heating in Eq. (1) described by the effective source $A(r, t)$ according to relation (6). As can be seen from these data, the T-spike model according to [1] predicts the formation of local melt-through zones in the $Fe_{85}B_{15}$ amorphous alloy.

Since some particular mechanisms involved in various modifications [1, 2] of the T-spike model do not coincide, it was of interest to elucidate the influence of such mechanisms on the ion subsystem temperature distribution in the track region at various moments of time. For this purpose, we have performed special numerical calculations for the $Fe_{85}B_{15}$ amorphous alloy using the T-spike model according to [2].

First, we have modified the mechanism of the electron subsystem heating (see [1]), while the form of the exchange term remained unchanged (see [2]). The results obtained in this case are presented in Fig. 2. As can be seen, use of the second variant of the electron subsystem heating [1] leads to stronger heating of the ion subsystem, which is manifested by a longer period of melting. This result is by no means surprising, since the first model of the electron subsystem heating [2] included the energy splash only at the initial moment, which was accompanied by elongation of a high-temperature period of time in the electron subsystem. In the

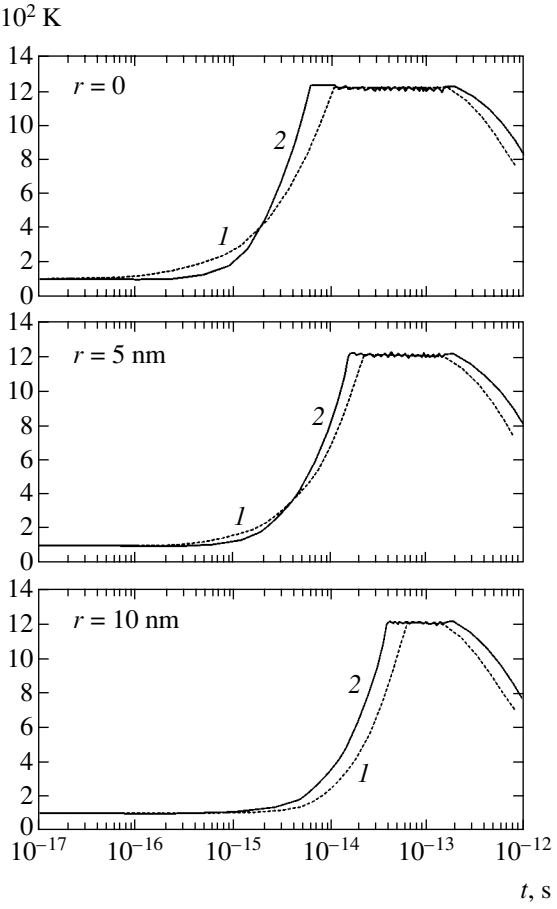

Fig. 2. Time variation of the ion subsystem temperature in the track region of a $Fe_{85}B_{15}$ amorphous metal alloy target irradiated by high-energy heavy charged particles. The calculations were performed for various distances from the track axis $r = 0$, 5, and 10 nm using a modified T-spike model with the exchange term in the thermal balance equation taken from [2] and the electron subsystem heating described according to (*1*) model [2] and (*2*) model [1].

second model, the energy source was also effectively operating in the electron subsystem during the entire thermalization period [1]. However, the obtained differences in the extent of heating of the ion subsystem are not very significant, which is explained by the fact that the characteristic time of the electron–phonon interaction is much greater than that of the electron thermalization. Thus, the process of heat exchange between the electron and ion subsystems in the target depends rather insignificantly on the mechanism (gradual [1] against instantaneous [2]) of the electron subsystem heating.

Now let us consider the second variant, whereby the method of heating was selected according to the first model [2] and the exchange term was borrowed from the second model [1]. As can be seen from the results presented in Fig. 3, use of the exchange terms of different types leads to a difference in the rates of heating of the ion subsystem. Since the time of heating is approximately the same, the amounts of energy transferred

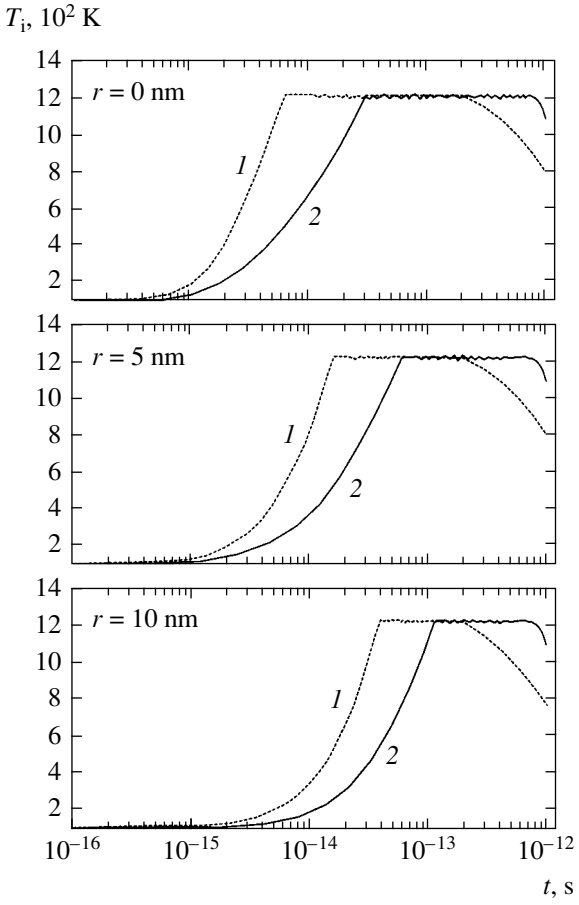

Fig. 3. Time variation of the ion subsystem temperature in the track region of a $Fe_{85}B_{15}$ amorphous metal alloy target irradiated by high-energy heavy charged particles. The calculations were performed for various distances from the track axis r = 0, 5, and 10 nm using a modified T-spike model with the electron subsystem heating described according to the model [2] and the exchange term in the heat balance equation taken from (*1*) model [1] and (*2*) model [2].

from the electron to the ion subsystem are approximately the same. The difference in the rates of heating of the ion subsystem is explained by the fact that the two models [1, 2] use different forms of the temperature dependence of the exchange term in the heat balance equation. In selecting this dependence, special attention has to be paid to the behavior of the heat capacity of the electron subsystem in the temperature interval under consideration. In our case, an analysis of the temperature dependence of C_e for $Fe_{85}B_{15}$ shows that it is expedient to use the exchange term in the form proposed in [2] (which takes into account the nonlinear temperature dependence of the heat exchange between the electron and ion subsystems).

Analysis of obtained results leads to the conclusion that no significant differences in the results of calculations of the temperature distribution for the ion subsystem in the track region is observed for various modifications of the T-spike model [1, 2]. Some differences are related only to the duration of this process.

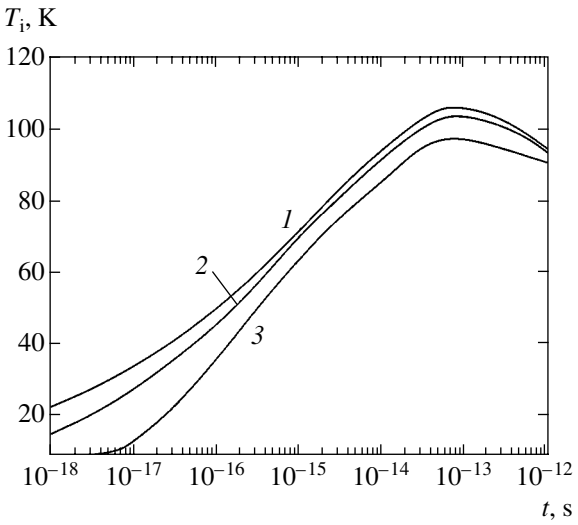

Fig. 4. Time variation of the ion subsystem temperature in the track region of a copper target irradiated by high-energy heavy charged particles. The calculations were performed using the modified T-spike model [2] for various distances from the track axis r = 0 (*1*), 5 (*2*), and 10 nm (*3*).

Figure 4 presents the results of calculations of the ion subsystem temperature of a copper target using a T-spike model [2] with the effective energy source in the electron subsystem $A(r, t) = 0$ (see Eq. (1)), Q = 100 keV/nm, and the initial electron heating described by Eq. (5). As can be seen from these data, the ion subsystem heating according to the T-spike model is low and is not accompanied by the formation of melt through zone in the track region.

3. USE OF THE COULOMB EXPLOSION MODEL FOR ANALYSIS OF CHANGES IN THE TEMPERATURES OF THE ELECTRON AND ION SUBSYSTEMS AND THE FORMATION OF POINT DEFECTS IN THE TRACK REGIONS FOR COPPER BOMBARDED BY HIGH-ENERGY HEAVY IONS

As was demonstrated above, use of the T-spike model for description of the process of ion subsystem heating in copper leads to a relatively small increase in the ion temperature in the track region and is not accompanied by the formation of a melt-through zone in this region. Thus, it would be of interest to study the ion subsystem heating in copper caused by the stopping of high-energy heavy ions as described in terms of the Coulomb explosion model. In this case, it would be easy to perform a comparative analysis of the results obtained using the two models (T-spike versus Coulomb explosion) with the same electron drag losses and other parameters, and to determine how the various physical mechanisms involved in these models affect the results of calculations.

Previously, we used the Coulomb explosion model for calculating the spatial distribution of the radial electric field $E_\rho(\rho)$ generated by a high-energy heavy particle stopped in the target material, where ρ is the distance from the track axis in the plane perpendicular to the heavy particle trajectory) [3]. We have also evaluated the kinetic energy Δε acquired by ions in this field over a characteristic lifetime t_r ($t_r \sim 10^{-15}$–10^{-16} s), which has the form [3]

$$\Delta\varepsilon(\rho) = \frac{(\Delta p)^2}{2M} = \frac{(eZE_\rho t_r)^2}{2M}, \quad (7)$$

where M and eZ are the host ion mass and charge, respectively; Δp is the momentum acquired by host ions in the effective electric field $E_\rho(\rho)$ for the time t_r.

For a period of time on the order of t_r, ions in the lattice are not thermalized, and we may assume that their initial temperature is determined by the average kinetic energy in the electric field. Therefore, the initial temperature of the ion subsystem, which is established quite rapidly under the action of a short-lived electric field $E_\rho(\rho)$, can be expressed as

$$k_B T_i(\rho, t = 0) = \Delta\varepsilon(\rho), \quad (8)$$

where k_B is the Boltzmann constant.

Figure 5 shows the results of calculations of the spatial profiles of the electric field generated in copper at $t = t_r$ by various singly charged ions with $Z_1 = 8$ (oxygen), 36 (krypton); 54 (xenon), and 92 (uranium) incident with an energy of 10 MeV/nucleon. The calculations were carried out using a procedure described in [3]. As can be seen, the electric field strength sharply increases with the atomic number of incident particles and reaches ~200 V/Å for uranium ions.

Figure 6 presents the time variation of a kinetic energy (temperature increment) gained by ions of the copper lattice under the action of the effective electric field, which was calculated using relation (7) and the field strength determined for krypton ions (Fig. 5).

Now let us consider the most "unfavorable" case for ion subsystem temperature variation, whereby the Coulomb explosion model is used to calculate only the ion subsystem heating for the initial ion temperature determined by formula (8) and the initial electron temperature, assumed to be constant and equal to the matrix temperature:

$$T_e(t = 0) = T_{matr} = 100 \text{ K}. \quad (9)$$

The results of calculations of the ion temperature variation in the track region of a copper target irradiated by high-energy heavy (krypton) ions ($Z_1 = 36$, $E = 10$ MeV/nucleon) are presented in Fig. 7. A comparison between Figs. 6 and 7 shows that the Coulomb explosion model predicts a stronger and faster increase in the ion temperature than does the T-spike model. The achieved level of heating of the ion subsystem may lead to the formation of a melt-through zone in the metal target.

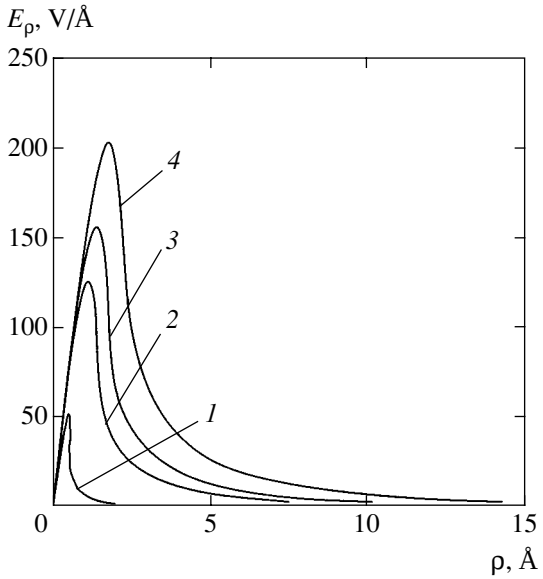

Fig. 5. Spatial profiles of the electric field generated in copper at $t = t_r$ by various ions with $Z_1 = 8$ (*1*), 36 (*2*), 54 (*3*), and 92 (*4*) incident with an energy of 10 MeV/nucleon.

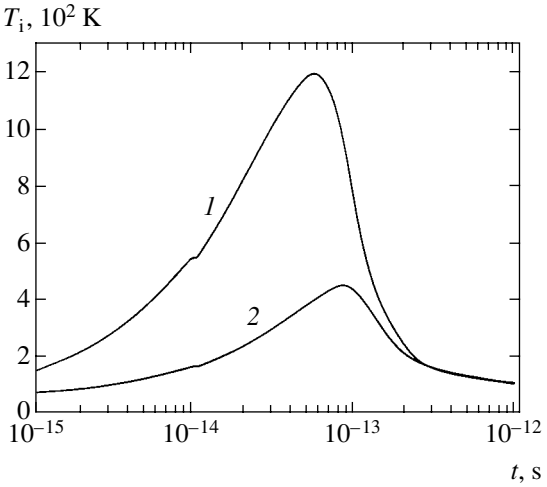

Fig. 6. Time variation of the ion subsystem temperature in the track region of a copper target irradiated by high-energy heavy (Kr) ions. The calculations were performed for r = 5 (*1*) and 10 Å (*2*) using the Coulomb explosion model, with the electron subsystem heating caused by the electron drag losses (dE/dr = 100 keV/nm) calculated using the T-spike model.

Now let us consider the influence of the Coulomb explosion on the formation of point defects in copper bombarded by heavy ions. As was noted above, an anomalous number density of point defects was observed [4] in Cu and Ag targets irradiated with heavy ions at a level of the electronic stopping power S_e from 0.1 to 100 keV/nm. The concentration of point defects

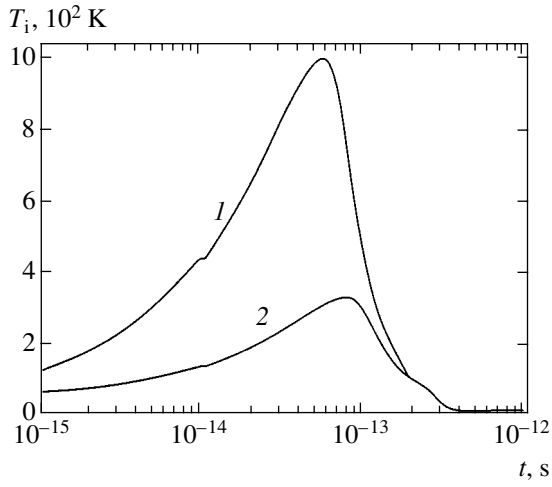

Fig. 7. Time variation of the ion subsystem temperature in the track region of a copper target irradiated by high-energy heavy (Kr) ions. The calculations were performed for r = 5 (*1*) and 10 Å (*2*) using the Coulomb explosion model, with the electron subsystem temperature assumed to be equal to the matrix temperature (100 K).

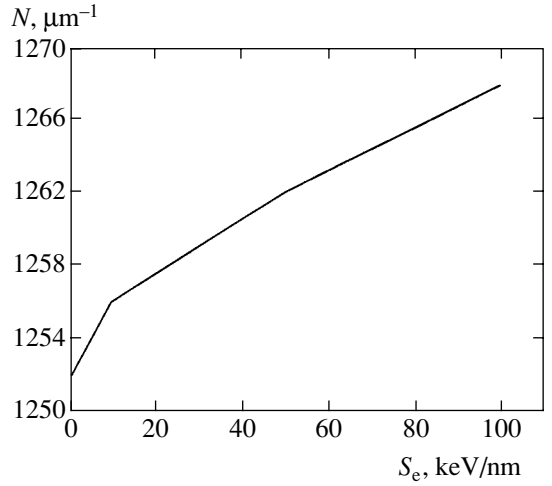

Fig. 8. A plot of the total number N of point defects generated by a heavy ion per unit ion range due to temperature fluctuations $T(\rho, t)$ for a characteristic track cooling time versus electron drag losses in the track region of a copper target. Calculations were performed using the Coulomb explosion model.

was significantly higher than that expected for their formation as a result of the elastic collisions of heavy projectiles with ions of the crystal lattice. For example, in copper irradiated with heavy ions at $S_e \approx 100$ KeV/nm, the number of point defects per unit heavy ion range (micron) reached $N \sim 10^3$–10^4 μm^{-1}. In this context, it would be of interest to calculate the point defect concentration as a function of the electronic stopping power of heavy ions using the Coulomb explosion and T-spike models and to compare these theoretical results to experimental data.

The total number N of point defects generated by a heavy ion per unit ion range due to the temperature fluctuations $T(\rho, t)$ for a characteristic track cooling time τ in a cylindrical region near the ion trajectory is

$$N = N_0 \nu_0 2\pi \int_0^{\rho_0} \rho d\rho \int_0^{\tau} \exp\left(-\frac{E_f}{kT(\rho, t)}\right) dt. \qquad (10)$$

Here, N_0 is the number density of host atoms in the matrix, ν_0 is the Debye frequency, and E_f is the energy of vacancy formation in the matrix lattice. Our calculations of the number of point defects formed due to thermal fluctuations in the track region in copper were performed for the following parameters: $N_0 = 8.5 \times 10^{28}$ m^{-3}, $\nu_0 = 10^{13}$ s^{-1}; $\rho_0 \approx 10^{-7}$ m; $\tau \approx 10^{-10}$ s; $E_f = 3$ eV. Figure 8 shows a plot of N versus S_e calculated using Eq. (10) within the framework of the Coulomb explosion model. As can be seen, the number of point defects produced by one ion per unit ion range (i.e., per unit track length) at $S_e \approx 100$ keV/nm is about $N \approx 1.3 \times 10^3$ μm^{-1}, which is close to the experimental data for Cu targets [4]. Analogous calculations were performed using Eq. (10) and the ion temperature profiles determined previously using the T-spike model, according to which the maximum ion temperature in the track region is about 110 K (Fig. 4). These calculations gave point defect densities about three orders of magnitude lower as compared to the values obtained using the Coulomb explosion model (where the maximum ion temperature is about $T_i \approx 10^3$ K, see Fig. 6). From this we infer that the Coulomb explosion model shows a better agreement with experiment than does the T-spike model for the number of point defects formed in the track region. Thus, the Coulomb explosion model provides a more adequate description of track formation in metal targets irradiated by heavy ions.

4. CONCLUSIONS

We have analyzed the process of heating of the electron and ion subsystems in the track region in copper and $Fe_{85}B_{15}$ amorphous alloy targets bombarded by high-energy heavy ions, as described within the framework of two theoretical models of track formation: Coulomb explosion and T-spike. Using these models, we calculated the spatial distributions of the ion and electron subsystem temperatures in the track regions. Various modifications of the T-spike model were considered, which differed in the mechanisms of electron subsystem heating and in the form of exchange terms describing the electron–phonon interaction in the heat balance equation. The results of numerical calculations of the ion subsystem heating in a $Fe_{85}B_{15}$ amorphous alloy were compared for two variants of the T-spike

model. It was found that modifications of this model do not describe significant changes observed in the alloy irradiated by heavy ions.

The ion subsystem heating in a copper target was calculated using a combination of the Coulomb explosion and T-spike models. The Coulomb explosion model predicts a stronger heating of the ion subsystem than does the T-spike model, with possible formation of melt-through zones in the track region. The point defect concentration in a copper target as a function of the electronic stopping power was calculated within the framework of the Coulomb explosion and T-spike models. Comparison to the available experimental data showed that the Coulomb explosion model (predicting higher ion temperatures in the track region: $T \sim 10^3$ K, Fig. 6) gives point defect densities three orders of magnitude higher than the values obtained using the T-spike model and are close to the experimental values [4]. The results showed that the Coulomb explosion more adequately describes the track formation in metals than does the T-spike model.

APPENDIX

1. Constants used in the numerical calculations for Cu (taken from [2, 5, 6, 7]). $N_e = 1.4 \times 10^5$ mol/m^3; $N_i = 1.4 \times 10^5$ mol/m^3; $\Delta H_{fus} = 13.01 \times 10^3$ J/mol;

T_e, K	$T_e < 1.7945 \times 10^4$ K	$T_e > 1.7945 \times 10^4$ K
C_e, J/(mol K)	$6.95 \times 10^{-4}\, T_e$	12.47 ($3R/2$)
T_e, K	$10 < T_e < 100$ K	$T_e > 100$ K
K_e, W/(m K)	$\dfrac{3.5 \times 10^6}{T_e T_i}$	$3.5 \times 10^2 \dfrac{T_e}{T_i}$
T_i, K	$10 < T_i < 80$ K	$T_i > 80$ K
C_i, J/(mol K)	$4.74 \times 10^{-5} T_i^3$	25 ($3R$)
K_i, W/(m K)	$\dfrac{1.8 \times 10^4}{T_i}$	$\dfrac{1.8 \times 10^4}{T_i}$

Note: It is assumed that 10 K < T_i < 1357.65 K.

2. Constants used in the numerical calculations for $F_{85}B_{15}$ amorphous metal alloy (taken from [1, 8]). $\rho_{sol} = 7.5$ g/cm^3; $\rho_{liq} = 7.3$ g/cm^3; $N_i = 8.47 \times 10^{22}$ cm^{-3}; $N_e = 1.73 \times 10^{23}$ cm^{-3}; $\zeta = 8.97 \times 10^{-5}$ J/(cm^3 K^2); $T_{fus} = 1220$ K; $T_0 = 44 \times 10^3$ K; $D_{e\,min} = 8$ cm^2/s; $H_{fus} = 1410$ J/cm^3; $T_L = 5625$ K; $k_1 = 135 \times 10^5$ (cm^2 K^2)/s; $k_2 = 45 \times 10^3$ (cm^2 K)/s;

$$C_i\ [\mathrm{J/g\ K}] = \begin{cases} 0.67 + 1.1 \times 10^{-4} T_i - \dfrac{4.6}{\sqrt{T_i}} & \text{(solid phase)}, \\ 0.95 & \text{(liquid phase)}, \end{cases}$$

$$K_i\ [\mathrm{W/cm\ K}] = \begin{cases} 0.021 & \text{(solid phase)}, \\ 0.83 & \text{(liquid phase)}. \end{cases}$$

T_e	0–300 K	300 K–T_L	T_L–T_0	$>T_0$
C_e, J/(cm^3 K)	ζT_e	ζT_e	ζT_e	$\dfrac{3}{2} N_e k_B$
K_e, W/(cm K)	$\zeta \dfrac{k_1}{T_e}$	ζk_2	$\zeta D_{emin} T_e$	$\dfrac{3}{2} N_e k_B D_{emin}$

REFERENCES

1. C. Dufour, *Commissariat L'energie atomique, Service de documentation et D'edition multimedia* (France), CEA-R-5638.
2. K. Yasui, Nucl. Instrum. Methods Phys. Res. B **90**, 409 (1994).
3. E. V. Metelkin and A. I. Ryazanov, Zh. Éksp. Teor. Fiz. **117**, 420 (2000) [JETP **90**, 370 (2000)].
4. A. Iwase, T. Iwata, and T. Nihira, J. Phys. Soc. Jpn. **61**, 3878 (1992).
5. J. M. Ziman, *Electrons and Phonons* (Clarendon, Oxford, 1960).
6. Y. S. Touloukian and E. H. Buyco, *Thermophysical Properties of Matter*, Vol. 4: *Specific Heat* (IFI/Plenum, New York, 1970).
7. Y. S. Touloukian, R. W. Powell, C. Y. Ho, and P. G. Klemens, *Thermophysical Properties of Matter*, Vol. 1: *Thermal Conductivity* (IFI/Plenum, New York, 1970).
8. M. Toulemonde and C. Dufour, Phys. Rev. B **46**, 14362 (1992).
9. Yu. V. Martynenko and Yu. N. Yavlinskiĭ, Dokl. Akad. Nauk SSSR **270**, 88 (1983) [Sov. Phys. Dokl. **28**, 391 (1983)].
10. I. M. Lifshits, M. I. Kaganov, and L. V. Tanatarov, Zh. Éksp. Teor. Fiz. **31**, 232 (1956) [Sov. Phys. JETP **4**, 173 (1956)].

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 128–144.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 150–167.

ELECTRONIC PROPERTIES OF SOLIDS

Effect of a Cluster Structure on the Interaction of Electron and Magnetic Subsystems in LaCa(Sr)MnO Epilayers

V. D. Okunev[a], Z. A. Samoilenko[a], R. Szymczak[b], and S. J. Lewandowski[b]

[a]*Donetsk Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 83114 Ukraine*
[b]*Institute of Physics, Polish Academy of Sciences, 02-668 Warsaw, Poland*
e-mail: okunev@mail.fti.ac.donetsk.ua

Received July 23, 2004

Abstract—The structure, electrical, and magnetic properties of epitaxial LaCa(Sr)MnO single crystal films with a clustered structure have been studied. In films with a "metallic" phase content $C_m^0 \le 0.15$, the electric conductivity is determined by the spin-dependent tunneling of charge carriers between "metallic" clusters, and the magnetoresistance is maximum at $T = 4.2$ K. The correlated motion of carriers over the system of tunneling-linked clusters leads to the formation of a window in the Coulomb blockade. The interactions between atomic, magnetic, and electron subsystems increase in the vicinity of the dielectric–metal percolation transition (T = 200–210 K), where the metal phase content C_m in the samples with $C_m^0 \ge 0.2$ reaches a maximum ($C_m^{crit} = 0.5$) due to an increase in the cluster size upon cooling. In this case, the magnetoresistance exhibits a maximum at $T = 260$ K, on the dielectric side of the percolation transition. Due to the presence of space charge regions at the periphery of the clusters, the content of a ferromagnetic phase is 1.5–2 times that of the "metallic" phase. For this reason, the calculations are performed using a model combining the tunneling conductivity mechanism with the percolation approximation for the description of magnetization. Allowance for the Coulomb interaction between charge carriers and clusters improves the agreement of theory and experiment. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

As is known, the colossal magnetoresistance in manganites of the LaCa(Sr)MnO type is due to the interaction between electron and magnetic subsystems, which results in the ordering of the magnetic subsystem upon cooling of a sample in the magnetic field. An important role in this phenomenon is played by structural inhomogeneities. The electron mechanisms of phase separation [1, 2] make possible the formation of inhomogeneities even in a single crystal structure, but this process is usually not related to qualitative changes in the atomic order. We believe that these mechanisms are operative at film growth temperatures (~700°C), when the magnetic interactions are weak but the possibility of phase separation between dielectric and "metallic" regions is retained. This is favored by a high density of holes ($p_0 \sim 10^{21}$ cm^{-3}), which is several orders of magnitude higher than the density of intrinsic carriers in $LaMnO_3$,

$$n_i = p_i \sim \exp(-E_g/2kT) \approx 10^{15}\ \text{cm}^{-3} \ll p_0,$$

at $E_g = 1.5$ eV and $T = 700$°C. High temperatures favor the formation of clusters with linear dimensions within 70–160 Å [3].

An integral characteristic of manganites, which characterizes the interaction between electron and magnetic subsystems, is offered by the temperature dependence of the resistance. This dependence exhibits a maximum near the Curie temperature (T_C), but factors determining the exact position (T_{max}) of this maximum were not studied previously [4–15]. The dielectric state of manganites at $T > T_C$ is related to a disorder in the magnetic subsystem and to the presence of a random potential whose nature has never been discussed [1, 2, 6, 9, 14, 16]. A decrease in the resistance at $T < T_C$ is usually explained by the magnetic ordering and the delocalization of electron states. We believe that such states cannot appear without changes in the atomic order of manganites.

The resistance usually exhibits a monotonic decrease when the temperature is varied within 200–300 K, although several cases have been reported [13, 14] when $R(T)$ showed abrupt changes in the region of 20–30 K. As will be demonstrated below, the rate of decrease in the resistance of epitaxial LaCa(Sr)MnO layers with a single-crystal structure depends on the local atomic order in a clustered structure. Slow $R(T)$ variations at $T < T_{max}$ are related to the spin-dependent tunneling of charge carriers between clusters with metallic conductivity. Switch-on of a percolation mechanism leads to a significant increase in the efficiency of interactions between electron and magnetic subsystems and is accompanied by a rapid drop in the resistance

with a large slope dR/dT. The results of calculations performed for particular samples show satisfactory agreement with experiment.

2. SAMPLES AND EXPERIMENTAL METHODS

We have studied a series of $La_{0.6}Sr_{0.2}Mn_{1.2}O_3$ (LaSrMnO) and $La_{2/3}Ca_{1/3}MnO_3$ (LaCaMnO) films with properties variable within broad limits, which were prepared by pulsed laser sputtering of the corresponding targets [17, 18]. The process was carried out using a KrF excimer laser with a pulse width of τ = 25 ns and an output beam intensity of Φ = 1.5–3.0 J/cm^2. The films were grown at an oxygen pressure of 300 mTorr on various single crystal substrates, including $SrLaGaO_4$, $Nd_3Ga_5O_{12}$, $Gd_3Ga_5O_{12}$, SiO_2/Si, $(LaAlO_3)_{0.3}(Sr_2AlTaO_6)_{0.7}$ (referred to below as SAT-30), and $(LaAlO_3)_{0.22}(Sr_2AlTaO_6)_{0.78}$ (SAT-22). The substrate temperature during growth was T_s = 600–730°C.

The magnetization of samples in the temperature range from 4.2 to 300 K was measured using a SQUID magnetometer. The electrical characteristics were studied using the standard four-point-probe technique. The structure was studied by means of X-ray diffraction using CrK_α radiation, which also provided information about the features of diffuse scattering in the clustered manganite films.

3. EXPERIMENTAL RESULTS

Below we present the results of our experimental investigation of the properties of epitaxial LaCa(Sr)MnO single crystal films. We will sequentially consider the structure, electrical, and magnetic properties of these films. The main attention will be devoted to the analysis of interactions between atomic, magnetic, and electron subsystems in the films studied. The experimental data are interpreted using the results of our previous investigations into the optical properties of such films [19]. In particular, the optical absorption data are used for obtaining quantitative estimates of the metal phase content in the samples.

3.1. The Structure of Samples

Although the mechanisms of phase separation in manganites are well known, experimental data for the structures with perfect atomic order are frequently interpreted within the framework of models assuming a homogeneous single-phase state of the material. It was found that data available on the phase composition of our samples are insufficient for solving the problems posed in this study, and we have additionally studied a fine structure of obtained manganite films. This section presents a detailed analysis of the local structure of our films and shows differences in the local atomic order of samples, which are related to the formation of mesoscopic inhomogeneities (clusters). The nature of clusters influencing the structure-sensitive properties of manganites is related to various lengths of Mn–O bonds for the different charged states of manganese.

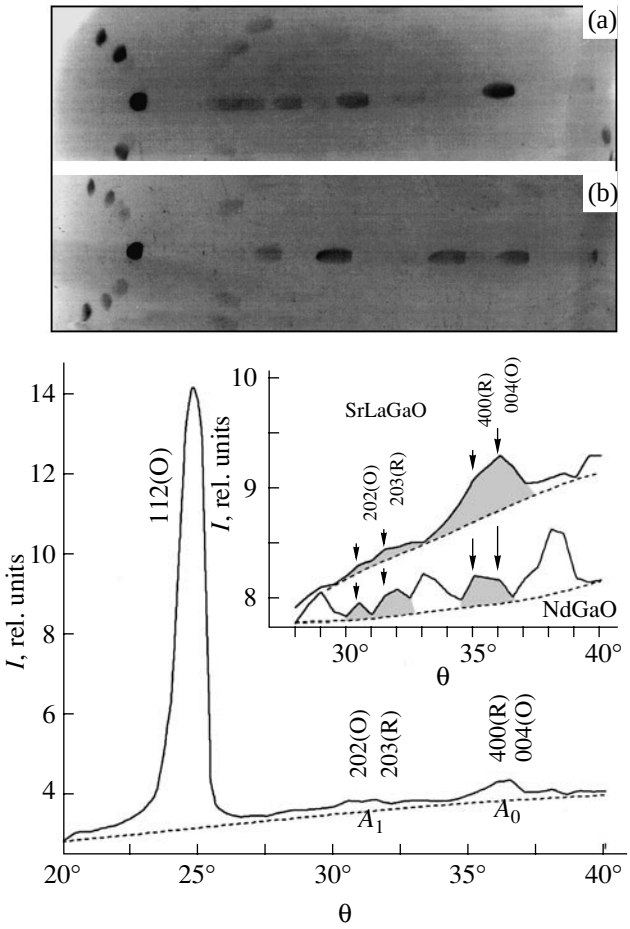

Fig. 1. The typical X-ray diffraction patters for LaSrMnO film on $SrLaGaO_4$ substrate. The bottom inset shows a difference between the diffractograms of LaSrMnO films on $SrLaGaO_4$ and $Nd_3Ga_5O_{12}$ substrates. The top inset shows two Debye patterns with Laue reflections for (a) LaCaMnO and (b) LaSrMnO films.

The multiscale character and variability of the local atomic order. The results of X-ray diffraction analysis of the films confirmed the absence of foreign phases in LaCa(Sr)MnO single crystal films. The diffraction patterns of various samples were much alike (Fig. 1), displaying a single pronounced peak (θ_{Cr} ~ 25°), which corresponds to the reflection from the main group of planes, and several weak diffuse maxima (θ_{Cr} ~ 27–38°) due to heterogeneous plane atomic groups (clusters) [3, 19–23]. The Debye patterns with Laue reflections showed that the films represented single crystal layers in which a long-range matrix structure coexisted with a mesoscopic order in the local regions (clusters) without violation of coherent binding with the matrix. The absence of sharp boundaries facilitates rearrangement of the atomic order in response to varia-

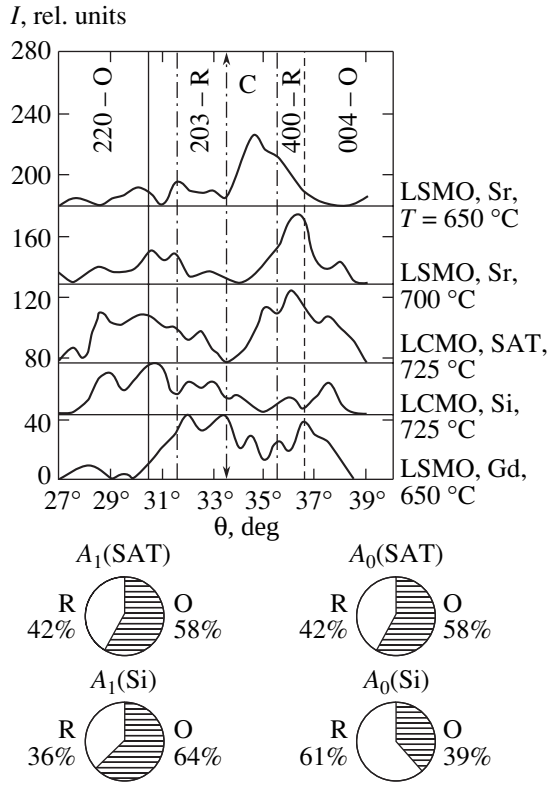

Fig. 2. Fragments of the X-ray diffraction patterns in the region of diffuse maxima for LaCaMnO (LCMO) films on SAT-30 and SiO_2/Si (T_s = 725°C) and LaSrMnO (LSMO) films on $SrLaGaO_4$ (T_s = 650 and 700°C) and $Gd_3Ga_5O_{12}$ (T_s = 650°C). Regions A_0 and A_1 correspond to $\theta > 33.5°$ and $\theta < 33.5°$, respectively.

tions of the temperature or applied magnetic field, not excluding the possible influence of defects present in peripheral regions of the clusters and the localization of charge carriers on the related states.

The main structural differences are revealed by the analysis of diffuse scattering from the clusters containing fragments of the Mn–O planes. The character of the diffraction intensity distribution in the region of the diffuse maxima (Fig. 2) is indicative of variability of the local order in the clustered regions. Depending on the θ value, there are three types of local order representing the cubic, orthorhombic, and rhombohedral phases [3, 21, 22]. The former two phases may exhibit metallic conductivity [24], while the rhombohedral phase forms are dielectric clusters [3, 19, 20]. For this reason, the conductivity σ follows variations of the phase composition of the clustered structure, which is characterized by the ratio C_m/C_Σ = (cubic + ortho)/(cubic + ortho + rhombic). At a large mismatch between the crystal lattice parameters of films and substrates (e.g., for the layers grown on $Gd_3Ga_5O_{12}$ and $Nd_3Ga_5O_{12}$), the concentration of defects and the width of transition regions in the clusters are much greater than those in well-matched layers (e.g., on $SrLaGaO_4$). This results in significantly (2–10 times) lower σ values (for the same C_m/C_Σ) and to a considerable scatter in $\sigma(C_m/C_\Sigma)$ for the films on $Gd_3Ga_5O_{12}$ and $Nd_3Ga_5O_{12}$ substrates.

The samples with high σ and dR/dT values (at $T < T_{max}$) are characterized by a symmetric (relative to the center at $\theta_c = 33.5°$) intensity distribution with two maxima (see the curve for a film on SAT-30 in Fig. 2). For this reason, we have studied the fine structure of diffraction patterns showing no such symmetry (Fig. 2). The observed diffraction peaks are due to the coherent diffuse X-ray scattering from the clustered fragments of planes in the (220)-O, (203)-R, (200)-C, and (400)-R, (004)-O phases (R, C, and O refer to the rhombohedral, cubic, and orthorhombic phases, respectively), where the (200)-C peak observed at $\theta_c = 33.5°$ corresponds to the center C). For a lateral size of clusters within D = 100–150 Å, a thickness of $D/3 \approx 30$–50 Å, and an interplanar spacing of $d \approx 1.8$–2 Å, clusters consist, on average, of 10–30 layers representing identical crystal planes.

As is known, the introduction of strontium ions Sr^{2+} (or calcium ions Ca^{2+}) into the composition of $LaMnO_3$ leads to the appearance of Mn^{4+} and (in the presence of local stresses) Mn^{2+} ions via the reaction $2Mn^{3+} \longrightarrow Mn^{4+} + Mn^{2+}$ [25, 26]. Taking into account the dependence of the Mn–O bond length on the charged state of manganese, $Mn^{(2-4)+}$ [25, 26], and the experimental values of d (corresponding to the observed θ), we conclude that the observed diffuse regions provide information about the structural components with different sets of manganese ions (Mn^{2+}, Mn^{3+}, and Mn^{4+}) and the corresponding Mn–O distances [19, 20]. In particular, the diffraction patterns reveal the points corresponding to large ($\theta < \theta_c$) and small ($\theta > \theta_c$) d values corresponding to the Mn^{2+}–O and Mn^{4+}–O bonds, respectively. The center (C in Fig. 2) corresponds to the Mn^{3+}–O bonds.

The distribution of planes between groups of clusters. Using the measured distribution of the diffuse scattering intensity, it is possible to estimate the fraction of scattering planes for regions A_1 and A_0 in Fig. 1 relative to the middle of each maximum ($\theta_1 = 30.5°$ and $\theta_0 = 36°$). These estimates are obtained using the I_+/I_- ratios, where I_+ corresponds to $d > d_{max}$, while I_- is taken on the side of $d < d_{max}$. The discrete values of intensities were determined for Δd = 0.02 Å. Thus, by measuring the I_+/I_- values in symmetric points with respect to θ_1 and θ_0, we obtain information about less and more closely spaced planes (Fig. 3). At the middle of each maximum, $I_+/I_- = 1$. When the planes with smaller d predominate, the I_+/I_- curve falls within the region of $I_+/I_- < 1$. As the proportion of planes with greater d grows, the curve shifts toward the region where $I_+/I_- > 1$.

For LaCaMnO films on SAT-30 (and the like films of LaSrMnO on $SrLaGaO_4$), the deviation of the A_0 curve in Fig. 3 toward $I_+/I_- < 1$ implies that the manga-

nese ions forming shorter Mn–O bonds (i.e., Mn^{4+}) predominate over Mn^{3+}. At the same time, the shift of the A_1 curve toward $I_+/I_- > 1$ is evidence for the predominance of Mn^{2+} ions forming longer Mn–O contacts.

The films characterized by high stresses at the film–substrate interface (such as LaCaMnO on SiO_2/Si and LaSrMnO on $Gd_3Ga_5O_{12}$) are described by sine-shaped I_+/I_- curves (Fig. 3). These curves cross the level of $I_+/I_- = 1$ and occur for the most part in the region of $I_+/I_- < 1$. Their behavior corresponds to alternating local density of clusters, whereby the more or less closely spaced planes sequentially predominate. The planes with less ionized manganese ($Mn^{(2–3)+}$ in A_1) exhibit weak (~20% relative to the average level) fluctuations of the intensity (Fig. 3). At the same time, the planes with more ionized manganese ($Mn^{(3–4)+}$ in A_0) show much greater intensity variations (~170%) of about the average level (Fig. 3). This nonmonotonic behavior of the relative intensity reflects a greater structural diversity related to the mutual misorientation of the fragments of different Mn–O planes containing differently charged manganese ions (Mn^{2+}, Mn^{3+}, and Mn^{4+}) in the clusters. The misorientation leads to an increase in the density of local states and a decrease in the density of free holes inside the clusters, thus influencing the conductivity of samples.

This situation is illustrated in the table, which presents data for LaCaMnO films on SiO_2/Si and SAT-30. As can be seen from these data, the ρ values in the films grown at the same temperature on different substrates may differ by one to two orders of magnitude.

Phase ratios in the clustered regions A_1 and A_0. It can be expected that a difference in the mismatch stresses for the clustered structures grown on different substrates will also affect the phase ratios in these mesoscopic components. It is natural to suggest that, in films on SiO_2/Si substrates, the concentration of elements of the rhombohedral (R) phase with a more distorted structure ($a = b = c$, $\alpha = \beta \neq \gamma$) will be greater than the concentration of elements of the orthorhombic (O) phase with a less distorted lattice ($a \neq b \neq c$, $\alpha = \beta = \gamma$). In order to estimate the relative content of these phases, let us use the ratio of the intensities observed at discrete angles θ for the corresponding *hkl* values (Fig. 2). In region A_1 of the LaCaMnO films on SAT, the R/O phase ratio is 0.7, while the same film of SiO_2/Si has R/O = 0.56. For A_0 in the film on SAT, we also obtain R/O = 0.7, while the ratio for SiO_2/Si is R/O = 1.56. In accordance with the diagrams in Fig. 2, the maximum difference is observed in the A_0 region for the structures with $Mn^{(3–4)+}$ ions on SiO_2/Si.

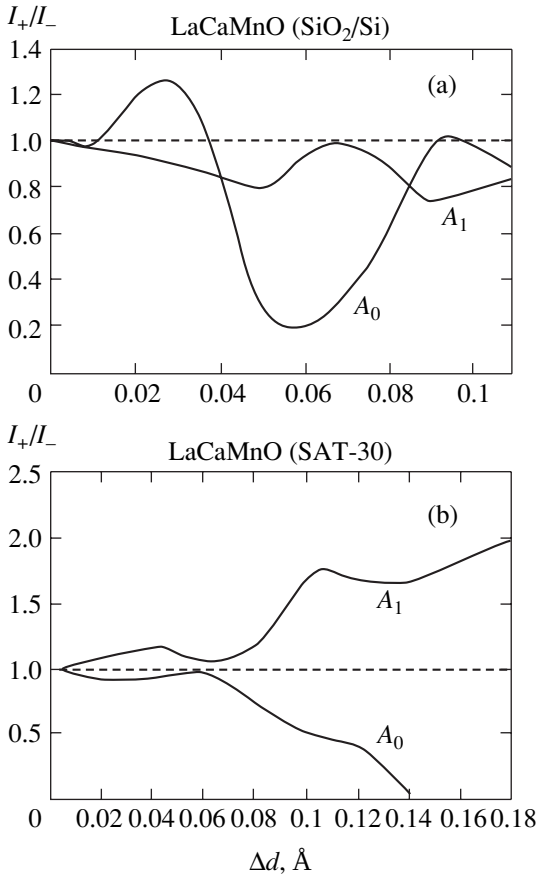

Fig. 3. Plots of the ratio I_+/I_- of the intensities of diffuse scattering in regions A_0 and A_1 (in symmetric points $d_>/d_<$ relative to d_{max}) for LaCaMnO films on (a) SiO_2/Si and (b) SAT-30 substrates.

3.2. The Minimum of Metallic Conductivity and the Electrical State of Manganite Films

The nature of a decrease in the resistance at $T < T_C$ is usually explained in terms of a metallic conductivity component or "metal-like" behavior of the conductivity. However, analysis of the electrical properties of manganite films (which exhibit characteristic $R(T)$ curves) shows that the dielectric–metal transition usually does not take place in the vicinity of the Curie temperature. The electrical properties of samples are deter-

Comparative characteristics of LaCaMnO films grown under identical conditions on SiO_2/Si and SAT-30 substrates

Substrate	T_s, °C	T_{max}, K	T_{min}, K	$\rho_{4.2\ K}$, ω cm	ρ_{min}, ω cm	ρ_{max}, ω cm	$\rho_{80\ K}$, ω cm	$\rho_{290\ K}$, Ω cm
SiO_2/Si	725	200	30	0.0667	0.0575	1.52	0.0828	0.23
SAT-30	725	268	–	0.00052	–	0.026	0.00104	0.0245

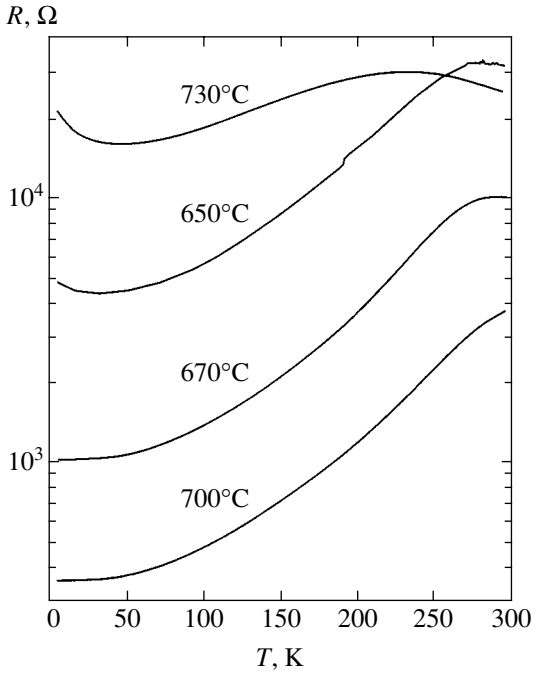

Fig. 4. Temperature dependences of the resistance of LaSrMnO films grown on $SiLaGaO_4$ substrates at various temperatures (T_s = 650–730°C).

mined by the tunneling of charge carriers between "metallic" clusters. We have established that the transition (of the percolation type) to a state with metallic conductivity (if it takes place at all) proceeds at lower temperatures (well below T_{max}), where the sample resistance decreases to less than one-tenth of the value at T_{max}. Thus, we can speak only of a formal similarity of the $R(T)$ curves for manganites and metals.

Indeed, let us compare the specific conductivity of manganite samples to the minimum metallic conductivity given by the formula [27, 28]

$$\sigma_{min} = \frac{\pi e^2}{4z\hbar a}\left[\frac{B}{V_0}\right]^2_{crit}, \quad (1)$$

where z is the coordination number, a is the distance between the impurity centers, V_0 is the random potential amplitude, and B is the bandgap width. For LaCa(Sr)MnO, the σ_{min} value according to formula (1) must not be less than 100 Ω^{-1} cm^{-1} [3, 29]. The conductivities of our samples fall within $\sigma = 9.5 \times 10^{-3}$–$4.1 \times 10^1$ Ω^{-1} cm^{-1} at T = 290 K and vary in a somewhat broader range at T = 4.2 K, where $\sigma = 7.1 \times 10^{-4}$–$2 \times 10^3$ Ω^{-1} cm^{-1}. In the region of $R(T)$ maximum, the conductivity is $\sigma = 3.4 \times 10^{-3}$–$3.85 \times 10^1$ Ω^{-1} cm^{-1}, which is significantly lower than σ_{min}.

The $R(T)$ curves with maxima of the usual type [30, 31] are observed even for σ = 0.25–0.4 Ω^{-1} cm^{-1}, which is two to three orders of magnitude smaller than σ_{min} [23]. Within the interval of T_{max}–T = 60–270 K, the samples still occur in the dielectric state. The values of $\sigma < 10$ Ω^{-1} cm^{-1} correspond the metallic phase content $C_m < 12\%$, which is below the percolation threshold for a lattice of any type and an arbitrary shape of grains [32, 33]. The increase in σ at $T < T_{max}$ is stimulated by the process of magnetic ordering. When the level of $\sigma = \sigma_{min}$ is attained and the material passes to a metallic state (e.g., for LaSrMnO grown on $SrLaGaO_4$ at T_s = 700°C (Fig. 4), this transition is observed at T = 160–170 K, while for LaCaMnO grown on SAT-30 and SAT-22, it takes place at T = 190–200 K), the monotonic character of $R(T)$ is retained.

In the dielectric state, the electric conductivity can be explained by several mechanisms, including hoping transport, injection of charge carriers, and tunneling. The measurements of σ as a function of the temperature T and the electric field strength F showed that the hopping transport does not take place even in high-ohmic ($\rho > 10^3$ Ω cm) samples [3]. The mechanisms involving charge carrier injection must exhibit characteristic features related to nonlinearity of the conductivity. However, we observed a weak nonlinearity (manifested by a 20–50% increase in σ in response to the field intensity growth by two orders of magnitude) only in high-ohmic ($\rho > 200$ Ω cm) samples. The I–V curves of low-ohmic samples were linear, whereby the resistance remained unchanged when the field F was varied within three to four orders of magnitude. Thus, the most probable mechanism of conductivity is that based on the tunneling between metallic clusters.

The influence of magnetic order on the charge transfer makes the tunneling spin-dependent [34–39]. There are two characteristics features in this effect. First, the absolute value of the negative magnetoresistance (MR) monotonically decreases with increasing temperature (Fig. 5, curve *1*), in contrast to the MR value for the sample whose state is closer to the dielectric–metal transition (curve *2*). Second, the MR exhibits maximum changes in weak magnetic fields (Fig. 6). Accordingly, the derivative dR/dH significantly varies only in the region of weak magnetic fields (H < 3000 Oe). The dR/dH value grows with decreasing temperature, exhibits a maximum at T = 50 K (which corresponds to the minimum in $R(T)$, Fig. 4), and then further decreases with the temperature.

In the case of electron tunneling, the resistivity of the films exponentially grows with increasing average distance L between clusters:

$$\rho = \rho_0 \exp\left(\frac{L}{L_0}\right), \quad (2)$$

where ρ_0 includes a factor on the order of $\exp(\Phi^{1/2})$ and Φ is the effective barrier height that is assumed to be constant. The main parameters of this process can be determined from experiments, considering a model of

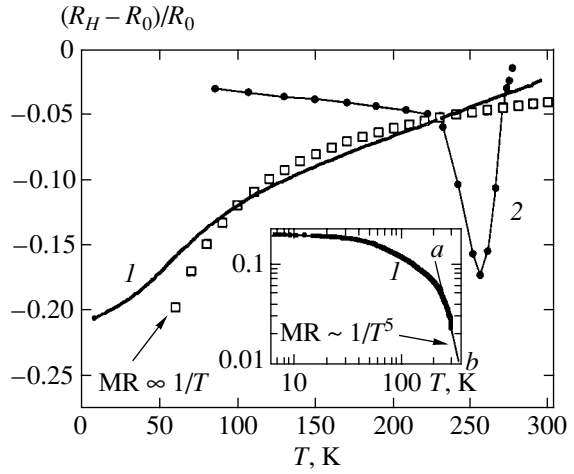

Fig. 5. Temperature dependences of the magnetoresistance for (*1*) LaSrMnO on $SiLaGaO_4$ (T_s = 730°C) measured at 0.376 T and (*2*) LaCaMnO on SAT-30 (T_s = 725°C) measured at 0.4 T. The inset shows agreement between the high-temperature region of the MR curve *1* and the MR ∝ $1/T^5$ approximation.

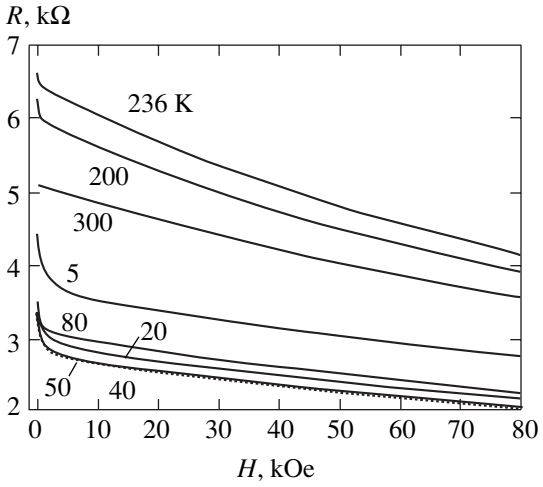

Fig. 6. Isothermal $R(H)$ curves measured at various temperatures, showing the influence of a magnetic field on the resistance of LaSrMnO film grown on $SrLaGaO_4$ at T_s = 730°C.

quasi-two-dimensional clusters in the form of plates parallel to the film surface. Let the plates have a lateral size of D and a thickness of $D/3$ (in accordance with the X-ray diffraction data) and form a regular lattice. Then, there are $1/(L + D)^3$ clusters per unit volume and the metal phase concentration is

$$C_\mathrm{m} = \frac{1}{3}D^3\frac{1}{(L+D)^3},$$

which yields

$$L = D[(3C_\mathrm{m})^{-1/3} - 1],$$

and

$$\rho = \rho_0\exp\left\{\frac{D}{L_0}[(3C_\mathrm{m})^{-1/3} - 1]\right\}. \quad (3)$$

The C_m value was determined using the optical absorption data [17, 18, 40–42] and taking into account the mechanisms of optical transitions [43–45] and the results of X-ray diffraction measurements. In some of the films, $C_\mathrm{m} \approx 15$–18% (which is below the percolation threshold) and metallic conductivity in the region of T_max is not observed. A shift in the percolation threshold to C_m^crit = 0.5–0.6 is inherent only in thin-film objects [32, 33]. Metallic conductivity was observed (for the most part, in LaCaMnO films on SAT-30 and SAT-22) at low temperatures.

Figure 7 shows the plot of $\rho = f(L)$ according to relation (3). A high value of the preexponential factor (ρ_0 = 0.0081 Ω cm) is related to reduced density of holes (p), which are localized in the peripheral regions of clusters. At $p \approx 3 \times 10^{21}$ cm^{-3}, the thickness of the space charge layer depleted of holes is

$$h = (2\varepsilon\varepsilon_0 V_D/ep)^{1/2}, \quad (4)$$

where V_D is the barrier height [46]. According to (4), $h \sim 10$ Å, which is a significant value in the case of small clusters.

3.3 Magnetic Properties

Manganites are known to represent two-phase magnetic systems [1, 2, 47]. At the same time, the samples

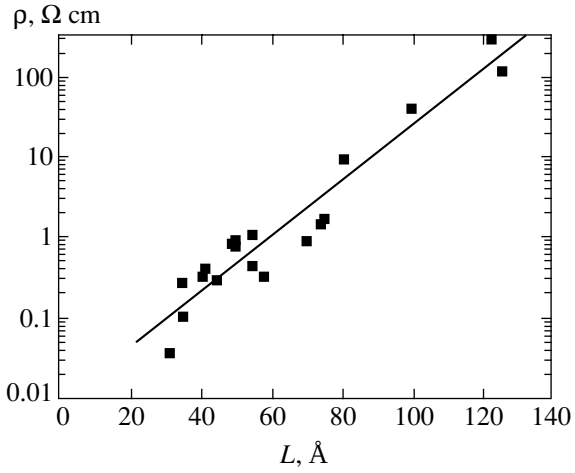

Fig. 7. A plot of the resistivity of LaSrMnO films versus average distance $L = D[(3C_\mathrm{m})^{-1/3} - 1]$ between "metallic" clusters (D = 100 Å, ρ_0 = 0.0081 Ω cm).

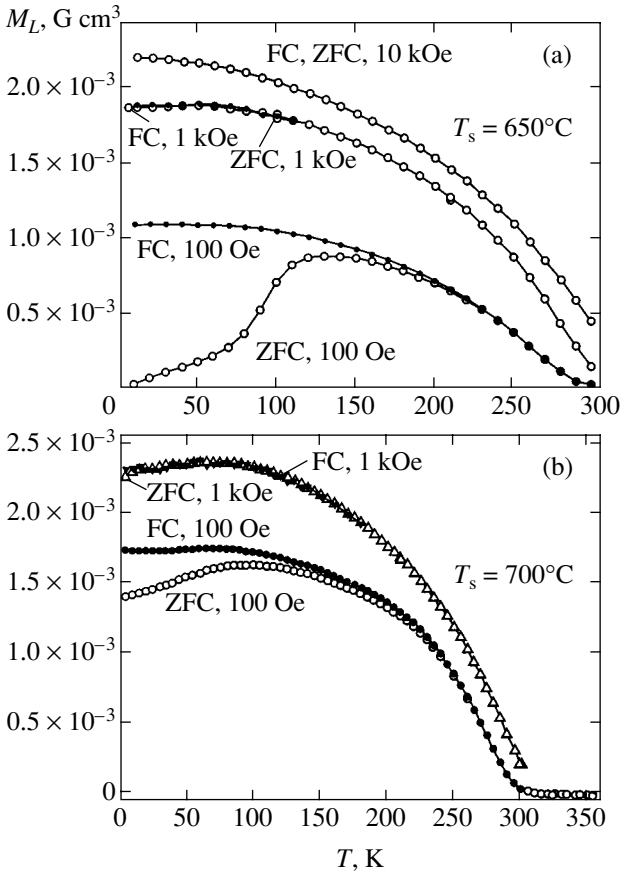

Fig. 8. Temperature dependences of magnetization of the FC and ZFC samples of LaSrMnO films grown on $SrLaGaO_4$ at T_s = 650 (a) and (b) 700°C.

are usually characterized by X-ray diffraction as single-phase. Proceeding from the notion of a two-phase magnetic state of LaCa(Sr)MnO single crystal films and taking into account their local heterogeneous and heterophase character, we will show below that the clustered structure of these films is responsible for both their electrical properties (charge carrier tunneling between clusters and, eventually, the dielectric–metal percolation transition) and magnetic properties (spin-dependent tunneling coupling of clusters and, eventually, the transition to a ferromagnetic state).

For example, let us consider the behavior of two LaSrMnO films (grown on $SrLaGaO_4$ at T_s = 650 and 700°C), which are characterized by similar $R(T)$ curves (Fig. 4) at a tenfold difference in the values of conductivity σ. Figure 8 shows the temperature dependences of the magnetization measured using samples cooled in the presence (field-cooled, FC) and in the absence (zero-field-cooled, ZFC) of magnetic field. The most pronounced difference between $M(T)$ curves of the FC and ZFC samples, as well as between the MR curves (see Figs. 5 and 6) is observed at low temperatures and weak (about 100 Oe) magnetic fields. Indeed, the $M(T)$ curves of the FC samples exhibit saturation at low temperatures, whereas the curves of the ZFC samples have a domelike shape characteristic of clustered spin glasses [47]. As the field is increased, the difference between FC and ZFC measurements decreases and virtually disappears at H = 1 kOe. As can be seen in Fig. 8, the difference between FC and ZFC samples is pronounced for the low-conductivity (high resistance) film grown at T_s = 650°C and is rather weak for the film grown at T_s = 700°C.

The trapping of free holes on shallow surface states in the peripheral regions of clusters in the course of cooling cannot be the main reason for the observed frustration, since the difference between FC and ZFC samples (Fig. 8) is not related to the cluster size D ($D \approx$ 130 Å for both samples in Fig. 8). Estimates show that this difference is most pronounced for the samples with large distances L between clusters. The magnetic dipole interactions between ferromagnetic clusters, which create a collective magnetic moment and determine the long-range magnetic order, grow with decreasing L and hinder the freezing of local magnetic moments of separate clusters. For example, the intercluster distance in the LaSrMnO films grown at T_s = 650°C (C_m = 0.085) is L = 75 Å. In the films grown at T_s = 700°C (C_m = 0.15), this distance decreases to L = 40 Å, which accounts, in accordance with formula (3), for the differences in the behavior of $R(T)$ (Fig. 4) and $M(T)$ (Fig. 8) curves.

For the ZFC sample of a film grown at T_s = 700°C, the value of M(10 K) at H = 100 Oe is 20 times that for an analogous film grown at T_s = 650°C. For the FC samples, this difference is much lower: the ratio decreases to 1.6 and 1.3 at 100 Oe and 1 kOe, respectively. Thus, the $M(T)$ value for the FC samples weakly depends on L in the region of small intercluster distances. As this distance grows, the dependence increases. Indeed, as the L value changes from 75 Å (T_s = 650°C) to 180 Å (T_s = 600°C), the magnetization M(10 K) exhibits a fivefold decrease at H = 10 kOe, a sevenfold decrease at H = 1 kOe, and a 14-fold drop at H = 100 Oe (Figs. 8a and 9). The films grown at T_s = 600°C exhibit a very pronounced difference between the results of measurements on the FC and ZFC samples [19]. Weak magnetic interactions between clusters (evidenced by the magnetization curves of ZFC samples) spaced by L = 180 Å in these films lead to vanishing of the maximum in the $R(T)$ curves. The absence of a spontaneous magnetic order (favoring a decrease in the resistance upon cooling) lead to the appearance of a temperature interval where $R(T)$ = const [19].

For the FC samples, the M value increases with decreasing temperatures (Figs. 8 and 9). There are intervals with $M(T)$ = const, but their lengths vary. Indeed, a film grown at T_s = 650°C is characterized by an interval with $M(T)$ = const (at T < 60 K) for the samples measured at 100 Oe and 1 kOe, but no such region is observed at H = 10 kOe. For the films grown at

T_s = 600°C, a small interval (at T < 40 K) with $M(T)$ = const is observed only at H = 100 Oe. A decrease in length of this interval with increasing field (for T_s = 650°C) and the absence of such interval at $H \geq 1$ kOe (for T_s = 600°C) is evidence for an increase in the average size of ferromagnetic clusters on cooling. Apparently, this trend is favored by smearing of the peripheral regions of these clusters.

3.4. Electrical Properties and the State of a Magnetic Subsystem

As is known, the electrical properties of manganites depend on the state of their magnetic subsystem, which accounts both for the appearance of a maximum in the $R(T)$ curve near the Curie temperature (T_C) and for the phenomenon of colossal magnetoresistance. However, the exact position of this maximum (T_{max}) and the factors responsible for a difference in the slopes of the $R(T)$ curves observed when the samples are cooled below T_{max} are still not established. We will demonstrate that this behavior is determined by various threshold states of the magnetic subsystem and by features of the cluster structures of manganites.

We have studied two series of LaCa(Sr)MnO films, which showed different rates of variation of the resistance R and magnetization M as functions of the temperature. Figure 10 shows the $R(T)$ and $M(T)$ curves measured at H = 1 kOe and matched on the temperature scale (note that the curves of FC and ZFC samples for H = 1 kOe coincide).

In samples of the first group, the resistance slowly varies with decreasing temperature (Figs. 10a and 10b) and the metallic conductivity state is not attained: the resistance exhibits a minimum at $T \approx 30$ K and then somewhat increases. The relative variation rate, $(dR/dT)/R$, typically does not exceed 0.7%/K. The maximum of $R(T)$ occurs on the sloping part of the $M(T)$ curve. For illustration, Fig. 11 shows the $M(T)$ and $R(T)$ curves matched with respect to the position of the $R(T)$ maximum (indicated by the vertical line in Fig. 11b for all samples). The maximum slopes dR/dT and dM/dT (indicated by arrows R and M, respectively, in Figs. 11a and 11b) do not coincide on the temperature scale, the shift reaching 75 and 47 K, respectively (curves *1* and *2*). This difference increases with the ratio $M(T)/M(5\text{ K})$ at $T = T_{max}$, showing a weak sensitivity of the electron subsystem to the state of the magnetic subsystem. There are critical (threshold) states of the magnetic subsystem: above these thresholds, the electrical properties of the samples studied begin to change.

In samples of the second group, the $R(T)$ curves exhibit no low-temperature minimum, while the maximum of $R(T)$ coincides with the onset of $M(T)$ growth. For these films, the resistance rapidly decreases within a narrow temperature interval (Figs. 10c and 10d). The relative variation rate, $(dR/dT)/R$, reaches 7%/K (which is 5–6 times the maximum value in the first group),

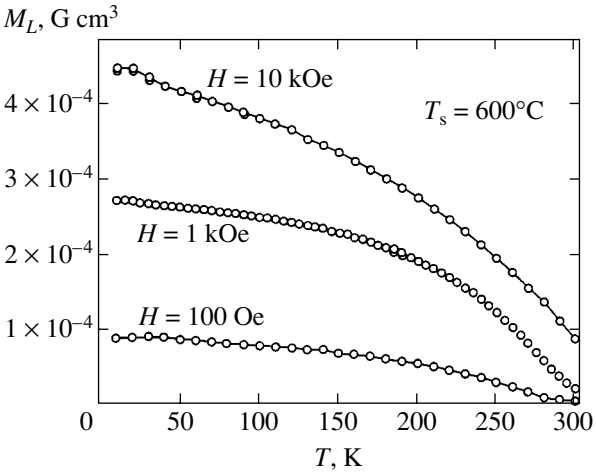

Fig. 9. Temperature dependences of magnetization of the FC samples of LaSrMnO films grown on $SrLaGaO_4$ at T_s = 600°C.

which is evidence for a high sensitivity of the electrical properties of these films to the state of the magnetic subsystem.

A temperature shift between the points of maximum slopes dR/dT and dM/dT is small, amounting to 3.5 and 1.8 K, respectively (Fig. 11b, curves *3* and *4*). At T < 100 K, the specific conductivity is $\sigma > 10^3\ \Omega^{-1}\ \text{cm}^{-1}$, which is much greater than σ_{min}. According to the estimates of σ_{min}, the dielectric–metal transition in these films takes place within the temperature interval 200–210 K.

In the region of metallic conductivity, the LaCaMnO films on SAT-22 and SAT-30 exhibited a nearly quadratic temperature dependence at T < 180 K

$$\rho(T) = \rho_0 + AT^{2.27},$$

where $\rho_0 = 5.37 \times 10^{-4}\ \Omega$ cm and $A = 10^{-8}\ \Omega\ \text{cm K}^{-2}$ (in agreement with the values reported for manganites [48]). At low temperatures, the $\rho(T)$ curve can be approximated by a linear dependence, $\rho = \rho_0(1 + \alpha T)$, with the temperature coefficients of resistance $\alpha = 7.1 \times 10^{-3}\ \text{K}^{-1}$ (LaCaMnO on SAT-22) and $3.7 \times 10^{-3}\ \text{K}^{-1}$ (LaCaMnO on SAT-30), which are close to the values of these coefficients for metals (Co, Al).

By analysis of the $R(T)$ and $M(T)$ curves, we can study the possible relationship between the electric conductivity and magnetization of manganites. We have compared the values of σ (normalized to σ_{max}) to the values of M at H =1 kOe for the same temperatures. We did not observe a correlation of the $\sigma \propto M^2$ type reported for strongly inhomogeneous media in the vicinity of the percolation threshold [49, 50]. Instead, the samples of LaCaMnO on SAT-30 and SAT-22 exhibited an exponential relation between σ and M

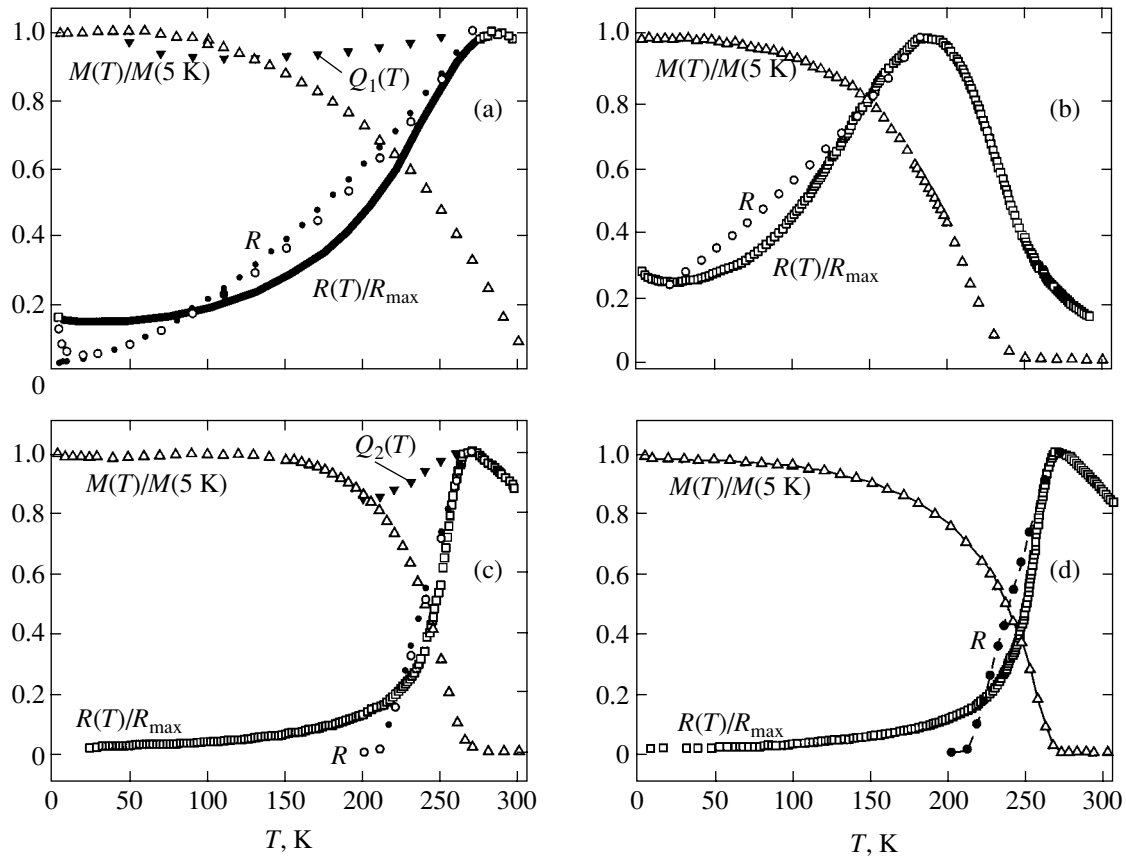

Fig. 10. Temperature dependences of the magnetization and resistance of LaCa(Sr)MnO manganite films on various substrates: (a) LaSrMnO films grown on $SrLaGaO_4$ (T_s = 650°C); (●, ○) $R(T)$ calculated for $C_m^0 \approx 0.1$ without (●) and with (○) the Coulomb blockade taken into account; (b) LaCaMnO on SiO_2/Si (T_s = 725°C); (o) $R(T)$ calculated for $C_m^0 \approx 0.1$ disregarding the Coulomb blockade; (c) LaCaMnO on SAT-30 (T_s = 725°C); (●, ○) $R(T)$ calculated for $C_m^0 \approx 0.2$ without (●) and with (○) the Coulomb blockade taken into account; (d) LaCaMnO on SAT-22 (T_s = 725°C); (●) $R(T)$ calculated for $C_m^0 \approx 0.2$ disregarding the Coulomb interaction of charge carriers with clusters.

(Fig. 12) in a broad range of relative magnetizations $M(T)/M(5\ K)$ (from 0.05–0.1 to 0.80–0.85). In Fig. 10, this interval corresponds to variations in the $R(T)$ within one order of magnitude.

The data plotted in Fig. 12 clearly reveal various threshold values of $M(T)/M(5\ K)$, which characterize the sensitivity of the electron subsystem to the magnetic state of the samples of both groups. For the films of LaCaMnO on SAT-30 and SAT-22, this threshold corresponds to $M(T)/M(5\ K) = 0.02–0.04$ (Fig. 12). For the samples of LaSrMnO on $SrLaGaO_4$, the magnetic threshold corresponds to $M(T)/M(5\ K) = 0.3$, and for the LaCaMnO films on SiO_2/Si this threshold increases to 0.6, in agreement with Figs. 10 and 11.

4. RELATIONSHIPS BETWEEN PROPERTIES OF MANGANITE FILMS

The temperature dependences of the resistance and magnetoresistance of manganites are usually described within the framework of a model based on the notion of an increase in the radius of ferromagnetic droplets in the course of magnetic ordering on cooling or under the action of an external magnetic field. However, this model does not take into account the local microstructure of the samples and, hence, cannot provide for an adequate description of their behavior. Below, we will demonstrate that the obtained experimental data on the electrical and magnetic properties of LaCa(Sr)MnO films can be described using a model taking into account a rearrangement of the cluster structure. According to this model, the size of clusters increases when the local atomic order in their peripheral regions becomes more perfect as a result of their magnetic ordering. The increase in the size of small clusters to a critical level is equivalent to the generation of new metallic clusters. In this context, we will consider the transition from spin-dependent tunneling to percolation, with allowance for the fact that concentrations of the metallic phase (C_m) and the ferromagnetic phase

(x^F) may not coincide because of the presence of a space charge layer depleted of holes in the peripheral regions of clusters.

4.1. Relationship between Parameters of the R(T) Curve for Small Metallic Phase Concentrations C_m

Recently, we established a linear relationship between T_{max} and T_{min} values [23],

$$T_{min} = \alpha T_{max} + \gamma, \quad (5)$$

where $\alpha = -0.46$ and $\gamma = 154.32$. The temperature interval $\Delta T = T_{max} - T_{min}$ is the region of active interaction between the atomic, magnetic, and electron subsystems. As the degree of atomic disorder decreases as a result of the magnetic ordering within ΔT, the amplitude of the random potential decreases, which leads to an increase in the electric conductivity. The nature of the $R(T)$ minimum at T_{min} can be related [51–53] to charge ordering, although this hypothesis is not free of drawbacks.

In considering the tunneling mechanism of conductivity in LaCa(Sr)MnO films, we will neglect the Coulomb interaction between charge carriers and clusters [54–56], since this factor is insignificant under the assumptions made. In the case of large concentrations of the "metallic" and ferromagnetic phases, a system of a tunneling-linked clusters is formed, which also feature a strong ferromagnetic interaction. The collective effects [57, 58] make possible a correlated motion of charge carriers over this system of clusters, with the formation of a window in the Coulomb blockade [59]. At a small (1–7%) concentration of the ferromagnetic phase and a high charge energy,

$$W_D \approx [e^2/\varepsilon D] \approx 3600 \text{ K } (\sim 0.3 \text{ eV}) \gg kT,$$

manganites exhibit a clearly pronounced dielectric behavior [60].

Assuming that the energy w of tunneling coupling between clusters in the region of the $R(T)$ maximum corresponds to their thermal energy, we obtain a relation

$$kT_{max} = w = w_0 \exp\left[-\left(\frac{L}{L_1}\right)\right]. \quad (6)$$

From a comparison of the two exponential relations, $w = w_0\exp[-(L/L_1)]$ and $\rho = \rho_0\exp(L/L_0)$, we infer that

$$\rho_{max} \approx (kT_{max})^{-\beta}, \quad (7)$$

where $\beta = L_1/L_0$. Such a relation between ρ_{max} and T_{max} with $\beta = 10.7$ was actually observed in our experiments (Fig. 13).

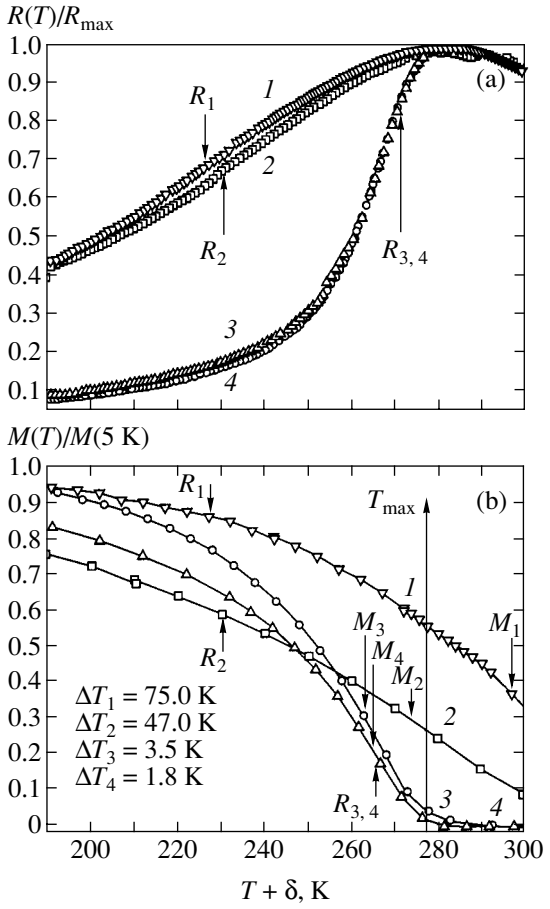

Fig. 11. Temperature dependences of the (a) resistance and (b) magnetization matched with respect to T_{max}: (*1*) LaCaMnO on SiO_2/Si (T_s = 725°C), δ = 91.81 K; (*2*) LaSrMnO films grown on $SrLaGaO_4$ (T_s = 650°C), δ = 0 K; (*3*) LaCaMnO on SAT-30 (T_s = 725°C), δ = 12.7 K; (*4*) LaCaMnO on SAT-22 (T_s = 725°C), δ = 11.7 K.

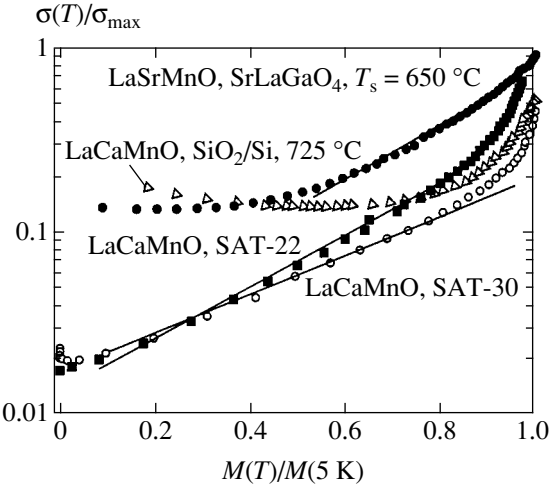

Fig. 12. Relationship between conductivity σ/σ_{max} and magnetization $M(T)/M(5\text{ K})$ for LaCa(Sr)MnO films on various substrates.

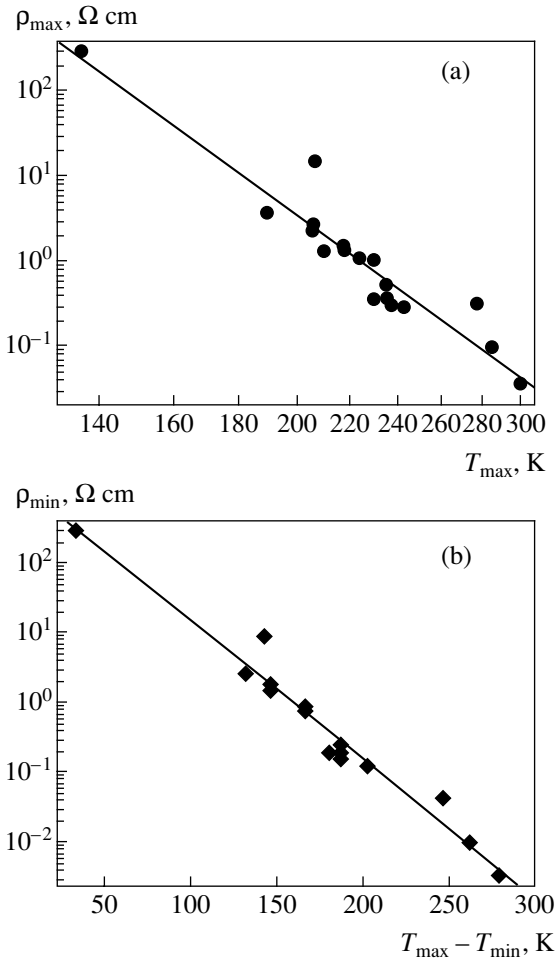

Fig. 13. Relationship between the extremal parameters for LaSrMnO films on various substrates: (a) ρ_{max} versus T_{max}; (b) ρ_{min} versus $T_{max} - T_{min}$.

A relation between ρ values in the minimum and maximum of $R(T)$ curves can be obtained assuming that C_m linearly increases with the T_{max}–T interval:

$$C_m(T) = C_m^0 + B(T_{max} - T), \tag{8}$$

where C_m^0 is the metallic phase concentration at $T = T_{max}$. Under these conditions, the samples belong to the class of systems with variable potential relief [61].Taking into account formulas (3) and (8), we obtain a relation valid at $T \le T_{max}$:

$$\rho(T) = \rho_0 \times \exp\left\{\frac{D(T)}{L_0}[(3C_m^0 + 3B(T_{max} - T))^{-1/3} - 1]\right\}. \tag{9}$$

This relation is valid for C_m values below the percolation threshold, $C_m < C_m^{crit} = 0.5$. In relation (9), the cluster size is assumed to be temperature-dependent: $D = D(T)$. Taking into account the presence of small clusters ($D \sim 2h$, where h is determined by formula (4)), which do not contribute to C_m, a relation between D and T should be considered as independent of $C_m(T)$. As the temperature decreases, the size of small clusters increases to a critical level, $D_{crit} \approx 2h$, which is equivalent to the generation of new metallic clusters. In this case, it is expedient to use the results known from percolation theory. Assuming that the variation of D with the temperature T is analogous to a change in the correlation radius,

$$L^*(C_m) \approx \left|C_m - C_m^{crit}\right|^{-\nu}$$

(where $\nu = 0.85$ is a critical index [32]), we eventually obtain

$$D(T) \approx D^0\left|[C_m^0 + B(T_{max} - T)] - C_m^{crit}\right|^{-\nu}. \tag{10}$$

Using this relation, the expression for the resistivity at $T = T_{min}$ can be written as

$$\rho_{min} = \rho_0 \times \exp\left\{\frac{D^0}{L_0}\left|[C_m^0 + B(T_{max} - T_{min})] - C_m^{crit}\right|^{-\nu} \times [3^{-1/3}[C_m^0 + B(T_{max} - T_{min})]^{-1/3} - 1]\right\}. \tag{11}$$

According to this relation, ρ_{min} must exponentially decrease with increasing interval $\Delta T = T_{max} - T_{min}$, which agrees with experiment (Fig. 13). For $T = T_{max}$, formula (3) yields

$$\rho_{max} = \rho_0 \exp\left\{\frac{D}{L_0}[(3C_m^0)^{-1/3} - 1]\right\}. \tag{12}$$

Using formulas (11) and (12), we eventually obtain

$$\rho_{min} = (\rho_{max})^{\xi}, \tag{13}$$

where $\xi = 1.25$. This relation also agrees with experiment for $B = 1 \times 10^{-3}$ K^{-1}, $D^0 = 44.2$ Å, $C_m^{crit} = 0.5$; C_m^0 values were taken from the optical data, and T_{max} and T_{min} values, from the experimental $R(T)$ curves.

Let us consider the behavior of $R(T)$ at $T < T_{max}$ and small C_m in the case of samples exhibiting slow variations of the resistance R (occurring far from the dielectric–metal transition). Writing expression (11) for the

current ρ and T values, denoting $C_m^0 + BT_{max} = X$ and $C_m^0 + BT_{max} - C_m^{crit} = Y$, taking into account that $BT/X < 1$, and using the approximation

$$\left(1 - \frac{B}{X}T\right)^{-1/3} \approx \left(1 + \frac{B}{3X}T\right),$$

we eventually obtain

$$\rho(T) \approx \rho_0 \times \exp\left\{\frac{D^0}{L_0}\frac{\left[(3X)^{-1/3}\left(1 + \frac{B}{3X}T\right) - 1\right]}{|Y|^\nu\left|1 - \frac{B}{Y}T\right|^\nu}\right\}. \quad (14)$$

Since the samples at $T < T_{max}$ occur in the dielectric state ($Y < 0$), the denominator in this formula is always positive and decreases slower than the numerator, so that ρ decreases on cooling of the sample. Figures 10a and 10b show the results of $R(T)$ calculations for the films with various T_{max} for $C_m^0 \approx 0.1$. The agreement with experiment is provided for $B = 1 \times 10^{-3}$ K^{-1} and $C_m^{crit} = 0.5$.

4.2. Relationship between Parameters of the R(T) Curve for C_m in the Vicinity of the Percolation Threshold

For the samples with steep $R(T)$ curves, which are close to the dielectric–metal transition, the agreement with experiment is observed for a nonlinear relation between C_m and T:

$$C_m = C_m^0 + (G(T_{max} - T))^t, \quad (15)$$

where $t = 1.75$ is the critical index used for the description of conductivity in strongly inhomogeneous media close to the percolation threshold [32] and G is a constant factor. Retaining the exponential dependence (inherent in the tunneling conductivity mechanism) of the resistance on the intercluster distance on the dielectric side of the transition, we can rewrite relation (14) as

$$\rho(T) = \rho_0 \exp\left\{\frac{D^0}{L_0}\left|[C_m^0 + (G(T_{max} - T))^t] - C_m^{crit}\right|^{-\nu} \cdot\times [3^{-1/3}[C_m^0 + (G(T_{max} - T))^t]^{-1/3} - 1]\right\}. \quad (16)$$

In this case, the agreement with experiment for a LaCaMnO film on SAT-30 ($C_m^0 = 0.2$, $\rho_{max} = 0.026$ Ω cm) is achieved at $G = 7.7 \times 10^{-3}$ K^{-1}; $\rho_0 = 0.0081$ Ω cm, $D^0/L_0 = 2.11$, and $C_m^{crit} = 0.5$ (Fig. 10c). It was assumed that the percolation threshold $C_m^{crit} = 0.5$ was independent of features of the process of magnetic ordering. The validity of this assumption is confirmed by an analysis of any particular $R(T)$ curve. Indeed, for a LaCaMnO film on SAT-30, the percolation transition takes place at 200–210 K (Fig. 10c). Using formula (15), we obtain $C_m = 0.45 < C_m^{crit} = 0.5$ for $T = 210$ K and $C_m = 0.52 > C_m^{crit} = 0.5$ for $T = 200$ K.

4.3. Relationship between Electrical and Magnetic Properties

An analysis of the relationship between electrical and magnetic properties of the clustered films under consideration requires taking into consideration the presence of a space charge layer depleted of holes (with a thickness of $h \sim 10$ Å according to formula (4)) in the peripheral regions of clusters. For a spherical cluster with $D = 100$ Å, the metallic core has a diameter of 80 Å. In planar clusters with a thickness of about several tens of ångströms, the conducting core extends over one-third of the layer. Thus, the real volume of the metallic phase involved in the conductivity amounts from 1/3 to 1/2 of the total volume of clusters. If the cluster has a ferromagnetic core, the ferromagnetism (according to the phase diagrams of manganites [24]) is probably also retained in the peripheral regions with dielectric properties. The difference in volumes of the metallic and ferromagnetic phases makes it necessary to combine the tunneling approximation used for the description of the electrical properties of samples and the percolation approximation used to describe their magnetic characteristics.

The LaCaMnO films with $C_m^0 = 0.2$ on SAT-22 and SAT-30 in the region of magnetizations $M(T)/M(5\text{ K}) = 0.1$–0.85 exhibit an exponential relation between σ and M, while the LaSrMnO films with $C_m^0 = 0.1$ on $SrLaGaO_4$ obey such a relation in a somewhat narrower region of $M(T)/M(5\text{ K}) = 0.6$–0.9 (Fig. 12).

If the content of the ferromagnetic phase x^F is close to the percolation threshold x_c^F, then

$$M \propto \chi \sim \left|x^F - x_c^F\right|^{-\nu}, \quad (17)$$

where χ is the magnetic susceptibility [62, 63]. In the range of concentrations $x^F > x_c^F$, we have

$$M = \xi(x_c^F - x^F)^{-\nu}, \quad (18)$$

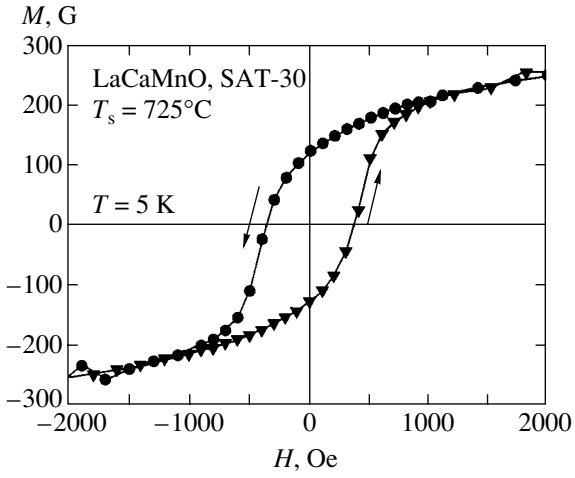

Fig. 14. Magnetization curves $M(H)$ with hysteresis for a LaCaMnO film on SAT measured at $T = 5$ K.

where ξ = const. This behavior agrees with a model according to which the volume of the ferromagnetic phase increases in the course of magnetic ordering [47]. Assuming that the x^F and C_m values differ by a constant factor β, so that $C_m = \beta x^F$, we obtain the relation

$$C_m = \beta\left[x_c^F - \left(\frac{M}{\xi}\right)^{-1/\nu}\right]. \tag{19}$$

The sample is still in the dielectric state and its conductivity (and resistance) is determined by the tunneling of charge carriers between clusters. According to formula (3),

$$\rho = \rho_0 \exp\left\{\frac{D}{L_0}\left[(3\beta x_c^F)^{-1/3} \times\left[1 - \frac{1}{x_c^F}\left(\frac{M}{\xi}\right)^{-1/\nu}\right]^{-1/3} - 1\right]\right\}. \tag{20}$$

Using the approximation

$$\left[1 - \frac{1}{x_c^F}\left(\frac{M}{\xi}\right)^{-1/\nu}\right]^{-1/3} \approx \left[1 + \frac{1}{3}\frac{1}{x_c^F}\left(\frac{M}{\xi}\right)^{-1/\nu}\right],$$

we arrive at

$$\rho \approx \rho_0 \exp\left\{\frac{D}{L_0} \times\left[(3\beta x_c^F)^{-1/3}\left[1 + \frac{1}{3}\frac{1}{x_c^F}\left(\frac{M}{\xi}\right)^{-1/\nu}\right] - 1\right]\right\}, \tag{21}$$

$$\sigma \approx \sigma_{01} \exp\left\{\left[\frac{M}{\xi}\right]^{1/\nu}\right\}. \tag{22}$$

Thus, in the vicinity of the percolation threshold, the conductivity σ and magnetization M obey an exponential relation, in agreement with experiment (Fig. 12).

4.4. Magnetoresistance

Combining the tunneling approximation used to describe the conductivity of samples and the percolation approximation used to characterize their magnetic state, we can also obtain an expression for the magnetoresistance. Taking into account relation (22), the conductivity in weak fields obeys the relation

$$\sigma \approx \sigma_0 \times \exp\left\{-Z(T) + \Delta(T, H) + \left[\frac{M(T, H)}{\xi}\right]^{1/\nu}\right\}, \tag{23}$$

where $Z(T)$ replaces the expression in braces in Eq. (16), which combines the terms dependent on the internal state of the magnetic subsystem and on its spontaneous variation with the temperature; $\Delta(T, H) > 0$ is the additive term describing the effect of the external magnetic field. In accordance with the magnetization curves $M(H)$ (Fig. 14), we have $\Delta(T, H) > 0$ for $M = 0$. Representing the magnetization as $M(T, H) = M_0 H^\eta (T_{max} - T)^t$ and using the above relation, we obtain an expression for the magnetic-field-induced change in the conductivity:

$$\frac{\Delta\sigma}{\sigma} = \frac{\sigma_H - \sigma_{H=0}}{\sigma_{H=0}} = \exp\left\{\left[\frac{M_0 H^\eta (T_{max} - T)^t}{\xi}\right]^{1/\nu} + \Delta(T, H)\right\} - 1. \tag{24}$$

The experimental data for such samples (Fig. 10c) show that $T_{max} \approx T_c$. According to formula (24), the maximum of $\Delta\sigma/\sigma$ takes place at a temperature T_{max} obeying the relation

$$\left[\frac{\partial}{\partial T}\Delta(T, H) - \frac{t}{\nu}\left(\frac{M_0 H^\eta}{\xi}\right)^{1/\nu}(T_{max} - T)^{t/\nu - 1}\right] \times \exp\left\{\left[\frac{M_0 H^\eta (T_{max} - T)^t}{\xi}\right]^{1/\nu} + \Delta(T, H)\right\} = 0. \tag{25}$$

The maximum of $\Delta(T, H)$ can be expected in the vicinity of the transition to the ferromagnetic state ($T_c \approx T_{max}$), where the system is characterized by maximum disorder. The value of Δ, being maximum at the onset of magnetic ordering, must decrease when the sample is cooled below T_{max}. If this condition is valid, we have $\partial\Delta(T, H)/\partial T > 0$ and, hence, $T_{max} - T > 0$. This implies that the maximum of the magnetoresistance must occur

to the left of T_{max}, falling in the region of the maximum slope of $R(T)$, in agreement with experiment (cf. T_{MR} = 256 K in Fig. 5, curve *2* and T_{max} = 270 K in Fig. 10c for the LaCaMnO films on SAT). For the films of LaSrMnO on $SrLaGaO_4$, which remain in the dielectric state, the experimental curve (Fig. 5, curve *1*) for the magnetoresistance in the interval $T_{max} - T_{min}$ is satisfactorily described by the dependence MR ~ $1/T$ (Fig. 5, squares) characteristic of spin-dependent tunneling [34] (the effect of temperature on the spin polarization was ignored). At $T > T_{max}$, a good approximation is offered by the dependence MR ~ $1/T^5$ (Fig. 5, line *ab* in the inset) reported for manganites in the dielectric state [60].

5. DISCUSSION

Our experiments showed that the presence of a maximum on the $R(T)$ curves and the subsequent decrease in the resistance on cooling, which reflect a decrease in the degree of disorder in the magnetic subsystem at $T < T_C$, are not related to the transition to metallic conductivity. It was established that the electrical properties of samples are more or less sensitive to changes in their magnetic state. This sensitivity is manifested by a shift of the position of maximum in the $R(T)$ curve relative to the $M(T)$ curve (and by the presence of threshold magnetization values in Figs. 10–12), which is indicative of the onset of an active influence of the magnetic ordering on the electric conductivity. An increase in the threshold $M(T)$ values is accompanied by a decrease in the dR/dT slope at $T < T_{max}$.

The magnetic properties of manganite films are determined by the structure and interactions of clusters [20, 23]. There are two topologically different processes leading to an increase in the magnetic order (i) inside the cluster and (ii) between clusters. In the former case, the size of a ferromagnetic cluster increases due to the attachment of atomic groups from the peripheral regions. In this process, the main factor is the local atomic order of the clustered structure (ion composition, set of crystallographic planes, relative content of clusters, and their spatial arrangement) and the distribution of charge carrier density in the peripheral regions of clusters. In the second case, the main role is played by the shape of clusters, their mutual orientation, and the intercluster distance, which determine the threshold magnetization. An increase in the concentration of a "metallic" phase and a decrease in the resistance at $T < T_{max} \approx T_C$ is related to the fact that an increase in the "metallic" cluster size leads to the delocalization of holes in the peripheral regions of such clusters.

For the electron mechanism of phase separation, the size R of ferromagnetic drops possessing metallic conductivity can be evaluated using the formula [64]

$$R^3 = \frac{135}{32}\frac{\pi^2}{m^2}\frac{\varepsilon\varepsilon_0}{e^2}\frac{(1+X)^{1/3}}{E_F f(X)},$$

where E_F is the Fermi energy, $f(X) = 2X + 3 - 3(1 + X)^{2/3}$, $X = V_A/V_F$ is the ratio of volumes of the antiferromagnetic and ferromagnetic phases (in fact, $X \approx 1/C_m$). For C_m = 0.12–0.25, this formula yields R = 50–70 Å. If the drops are oblate, so that the thickness is $(1/3)R$, the maximum dimensions of clustered formations increase to 90–120 Å, which agrees with the X-ray data and the results of magnetic measurements described in [65].

For films grown under the optimum conditions, C_m = 15–20% (at room temperature). These values agree with the metal phase content anticipated if all holes generated as a result of Sr^{2+} (Ca^{2+}) introduction are involved in the formation of the "hole" drops. This agreement confirms the assumption that the primary factor responsible for cluster formation is the electron mechanism of phase separation. The real shapes of clusters formed on the basis of hole drops are determined by the anisotropy of atomic order in the crystal.

The X-ray diffraction patterns represent diffuse reflections from the fragments of planes with Mn–O bonds coherently built into a single crystal structure of LaCa(Sr)MnO films. The clustered fragments of Mn–O planes alternate with nonclustered planes (of the (112) type in Fig. 1) containing La, Sr(Ca), and O atoms, which are characterized by long-range order and fixed bond lengths. These planes retain the long-range order and maintain the orientation of clustered fragments in the single crystal structure, thus facilitating the process of atomic and magnetic ordering in the peripheral regions of clusters.

The clusters are formed as a result of the relaxation of elastic stresses related to a redistribution of the charged states of Mn ions. According to this process, the matrix must contain predominantly Mn^{3+} ions, while the Mn^{3+} and Mn^{4+} ions predominantly enter into clusters. According to the stoichiometric composition of the material, the amount of such manganese ions is sufficient to provide for metallic conductivity. However, in a structure with clusters, only about 1/3 of their total volume can be involved in the conductivity and, hence, in most cases $\sigma < \sigma_{min}$. The X-ray diffraction patterns (Fig. 2) show that the number of Mn^{3+} ions is small, but Mn^{2+} ions appear due to the local stresses accompanying the cluster formation (with a maximum in the stage of hole drop formation) and stimulating the reaction $2Mn^{3+} \longrightarrow Mn^{2+} + Mn^{4+}$ [25].

The dimensions of intermediate zones, which provide for a smooth transition from order to disorder, are comparable to the size of small (<30 Å) clusters. These regions are the main sources of localized states.

Considering the influence of the phase composition of clusters on the conductivity, it is necessary to take into account the fact that crystallographic distortions of the rhombohedral type lead to the formation of a gap with $E_{g1} \approx 0.25$ eV [19] in the density of e_g states responsible for the hole conductivity of manganites. This gap appears upon removal of the orbital degeneracy as a result of the Jahn–Teller distortions [19, 26, 43]. In the case of an ortho ⟶ rhombo phase transition, the presence of a gap with $E_{g1} \approx 0.25$ eV decreases the σ value by ten orders of magnitude [3]. A large content of the rhombohedral phase in the clusters (e.g., for LaCaMnO films on SiO_2/Si substrates, Fig. 2) hinders realization of the percolation mechanism of conductivity.

In films with an orthorhombic structure, the distortions are small, the splitting between $x^2 - y^2$ and z^2 states is not observed and $E_{g1} = 0$· in agreement with [24]. If the orthorhombic phase predominates in the cluster structure (as in LaCaMnO films on SAT-30), the samples exhibit a dielectric–metal phase transition with a large dR/dT slope (Fig. 10c).

The phase composition of the cluster structure determines the difference (from 10–15 up to 3000%) in magnetization of the FC and ZFC samples. This difference is more pronounced in high-ohmic samples grown at $T_s \le 600$°C (in which the stress relaxation is hindered) and in the films grown on substrates (e.g., on SiO_2/Si) with a large lattice mismatch (where a considerable number of dielectric rhombohedral clusters are formed). Falling between the ferromagnetic orthorhombic clusters, these dielectric clusters produce effective screening of the ferromagnetic phase). Taking into account that the position of the maximum observed on the $M(T)$ curves of ZFC samples (Fig. 8a) corresponds to the Néel temperature (135–137 K) of LaSrMnO [24], this maximum is probably related to the antiferromagnetic ordering in rhombohedral clusters.

Changes in the magnetic order (induced by cooling below T_C or by an external magnetic field) initiate rearrangement of the cluster structure, whereby C_m increases at the expense of peripheral regions. If the resulting content of metallic clusters approaches the percolation threshold $C_m^{crit} = 50\%$, the system exhibits a dielectric–metal transition of the percolation type. The metallic phase content above the percolation threshold is reached only in the structures characterized by minimum stresses, in which orthorhombic clusters are predominantly formed. This process is facilitated in the films grown on SAT substrates, where the maximum of $R(T)$ coincides with the onset of the magnetization growth (Figs. 10c and 10d), at a high rate of temperature variation (large slow dR/dT) in the region of $T < T_{max} \approx T_C$. If the rhombohedral clusters (hindering the structure rearrangement into the infinite percolation cluster) predominate, the dR/dT values are low, the state of metallic conductivity is not attained, and the conductivity mostly proceeds via the mechanism of spin-dependent tunneling with the participation of metallic clusters.

In considering the tunneling mechanism, we ignored the Coulomb interaction of charge carriers with clusters. Allowance for this interaction would require introducing additional factors (dependent on the charge energy W_a) into Eqs. (14) and (16):

$$\rho \sim Q = \exp(W_a/kT),$$

where $W_a \approx e^2/\varepsilon D[1 - (C_m/C_m^{crit})^{1/3}]$ [56]. For Eq. (14), this factor is

$$Q_1(T) = \exp\left\{\frac{e^2}{\varepsilon}\frac{1}{k}\frac{1}{D_0}\frac{1}{T}\left|[C_m^0 + B(T_{max} - T)] - C_m^{crit}\right|^\nu \times \left[1 - \left(\frac{C_m^0 + B(T_{max} - T)}{C_m^{crit}}\right)^{1/3}\right]\right\}.$$

For Eq. (16), we have

$$Q_2(T) = \exp\left\{\frac{e^2}{\varepsilon}\frac{1}{k}\frac{1}{D_0}\frac{1}{T}\left|[C_m^0 + G(T_{max} - T)^t] - C_m^{crit}\right|^\nu \times \left[1 - \left(\frac{C_m^0 + G(T_{max} - T)^t}{C_m^{crit}}\right)^{1/3}\right]\right\}.$$

Under the assumptions made in this study (as expressed by formulas (8) and (15)), the functions Q_1 and Q_2 are slowly varying with the temperature. Allowance for these terms in Eqs. (14) and (16) improves the fit of the calculated curves to experiment (Figs. 10a and 10c). It should be noted that the introduction of the Q_1 term in Eq. (14) for the films with small C_m leads (in agreement with experiment) to the appearance of a minimum on the calculated $R(T)$ curve (Fig. 10a).

6. CONCLUSIONS

We have demonstrated that LaCa(Sr)MnO manganite films are characterized by the coexistence and interaction of structural elements on different scale levels, including microscopic, mesoscopic, and long-range order. On the microscopic scale, the main factor is the ratio of differently charged manganese ions (Mn^{2+}, Mn^{3+}, and Mn^{4+}) in structural clusters. The mesoscopic order (limiting the dimensions of clusters) determines the main electrical, optical, and magnetic properties of the manganite films. The nature of clusters is related to the electron mechanisms of phase separation operating

in the course of film growth and to differently charged manganese ions, which act as sources of local stresses. The phase composition of clusters determines the efficiency of the interaction between the electron and magnetic subsystems. The long-range order in the matrix planes retains a single crystal structure of the samples and favors the correlated motion of charge carriers over the cluster system.

At a low ($C_m \le 15\%$) concentration of the metallic phase, the characteristics of manganite films can be described within the framework of the model of charge carrier tunneling between clusters with metallic conductivity. For $C_m \ge 20\%$, it is necessary to take into account conductivity according to the percolation mechanism. The presence of regions depleted of holes in the transition layers between clusters and the matrix leads to the inequality of concentrations of the metallic and ferromagnetic phases, which makes it necessary to combine the tunneling approximation used for description of the electrical properties of samples and the percolation approximation used to describe their magnetic state. The maximum efficiency of interaction between the electron and magnetic subsystem is observed in the vicinity of the dielectric–metal percolation transition.

ACKNOWLEDGMENTS

The authors are grateful to Yu.V Medvedev, A.É. Filippov, and H. Szymczak for fruitful discussions and useful advice.

This study was supported in part by the Polish Government (grant no. 2P03B04423).

REFERENCES

1. É. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996) [Phys. Usp. **39**, 781 (1996)].
2. L. P. Gor'kov, Usp. Fiz. Nauk **168**, 665 (1998) [Phys. Usp. **41**, 589 (1998)].
3. Z. A. Samoĭlenko, V. D. Okunev, E. I. Pushenko, *et al.*, Zh. Tekh. Fiz. **73** (2), 118 (2003) [Tech. Phys. **48**, 250 (2003)].
4. A. J. Millis, P. B. Littlewood, and B. I. Shraiman, Phys. Rev. Lett. **74**, 5144 (1995).
5. A. Urushibara, Y. Moritomo, T. Arima, *et al.*, Phys. Rev. B **51**, 14 103 (1995).
6. J. M. D. Coey, M. Viret, and S. von Molnar, Adv. Phys. **48**, 167 (1999).
7. A. Biswas, A. K. Raychaudhuri, and R. Mahendiran, J. Phys.: Condens. Matter **9**, L355 (1997).
8. A. Biswas, M. Rajeswari, R. C. Srivastava, *et al.*, Phys. Rev. B **63**, 184424 (2001).
9. M. B. Salamon and M. Jaime, Rev. Mod. Phys. **73**, 583 (2001).
10. M. F. Hundley and J. J. Neumeier, Phys. Rev. B **55**, 11511 (1997).
11. S. V. Trukhanov, I. O. Troyanchuk, N. V. Pushkarev, and H. Szymczak, Zh. Éksp. Teor. Fiz. **122**, 356 (2002) [JETP **95**, 308 (2002)].
12. H. L. Ju and Hyunchul Sohn, J. Magn. Magn. Mater. **167**, 200 (1997).
13. S. H. Chun, Y. Lyanda-Geller, M. B. Salamon, *et al.*, J. Appl. Phys. **90**, 6307 (2001).
14. N. G. Bebenin, R. I. Zaĭnullina, V. V. Mashkautsan, *et al.*, Zh. Éksp. Teor. Fiz. **117**, 1181 (2000) [JETP **90**, 1027 (2000)].
15. G. Zhao, V. Smolyaninova, W. Prellier, and H. Keller, Phys. Rev. Lett. **84**, 6086 (2000).
16. V. M. Loktev and Yu. G. Pogorelov, Fiz. Nizk. Temp. **26**, 231 (2000) [Low Temp. Phys. **26**, 171 (2000)].
17. V. D. Okunev, Z. A. Samoilenko, A. Abal'oshev, *et al.*, Appl. Phys. Lett. **75**, 1949 (1999).
18. V. D. Okunev, Z. A. Samoilenko, V. M. Svistunov, *et al.*, J. Appl. Phys. **85**, 7282 (1999).
19. V. D. Okunev, Z. A. Samoĭlenko, T. A. D'yachenko, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **46**, 1829 (2004) [Phys. Solid State **46**, 1895 (2004)].
20. Z. A. Samoilenko, V. D. Okunev, E. I. Pushenko, *et al.*, Acta Phys. Pol. A **105**, 93 (2004).
21. Z. A. Samoĭlenko, V. D. Okunev, T. A. D'yachenko, *et al.*, Zh. Tekh. Fiz. **74** (5), 50 (2004) [Tech. Phys. **49**, 572 (2004)].
22. Z. A. Samoĭlenko, V. D. Okunev, E. I. Pushenko, *et al.*, Zh. Tekh. Fiz. **74** (4), 58 (2004) [Tech. Phys. **49**, 435 (2004)].
23. V. D. Okunev, Z. A. Samoilenko, A. Abal'oshev, *et al.*, Phys. Lett. A **325**, 79 (2004).
24. V. E. Naĭsh, Fiz. Met. Metalloved. **92** (5), 16 (2001) [Phys. Met. Metallogr. **92**, 437 (2001)].
25. M. D. Coey, M. Viret, L. Ranno, and K. Ounagjela, Phys. Rev. Lett. **75**, 3910 (1995).
26. J. B. Goodenough, J.-S. Zhou, F. Rivadulla, and E. Winkler, J. Solid State Chem. **175**, 116 (2003).
27. N. F. Mott, *Metal–Insulator Transitions* (Taylor and Francis, London, 1974).
28. N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials*, 2nd ed. (Clarendon, Oxford, 1979).
29. A. E. Kar'kin, D. A. Shulyatev, A. A. Arsenov, *et al.*, Zh. Éksp. Teor. Fiz. **116**, 671 (1999) [JETP **89**, 358 (1999)].
30. A. J. Millis, P. B. Littlewood, and B. I. Shraiman, Phys. Rev. Lett. **74**, 5144 (1995).
31. M. B. Salamon and M. Jaime, Rev. Mod. Phys. **73**, 583 (2001).
32. B. I. Shklovskiĭ and A. L. Éfros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer, New York, 1984).
33. A. B. Khanikaev, A. B. Granovskiĭ, and J. P. Clerc, Fiz. Tverd. Tela (St. Petersburg) **44**, 1537 (2002) [Phys. Solid State **44**, 1611 (2002)].
34. J. S. Helman and B. Abeles, Phys. Rev. Lett. **37**, 1429 (1976).

35. S. Lee, H. Y. Hwang, B. I. Shraiman, *et al.*, Phys. Rev. Lett. **82**, 4508 (1999).
36. N. Zhang, W. Ding, W. Zhong, *et al.*, Phys. Rev. B **56**, 8138 (1997).
37. C. H. Shang, J. Nowak, R. Jansen, and J. S. Moodera, Phys. Rev. B **58**, R2917 (1998).
38. P. Lyu, D. Y. Xing, and J. Dong, J. Magn. Magn. Mater. **202**, 405 (1999).
39. A. Gupta and J. Z. Sun, J. Magn. Magn. Mater. **200**, 24 (1999).
40. R. A. Smith, *Semiconductors* (Cambridge Univ. Press, Cambridge, 1978).
41. V. D. Okunev, Z. A. Samoilenko, A. Abal'oshev, *et al.*, Phys. Rev. B **62**, 696 (2000).
42. V. D. Okunev, N. N. Pafomov, V. A. Isaev, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **44**, 150 (2002) [Phys. Solid State **44**, 157 (2002)].
43. N. N. Loshkareva, Yu. P. Sukhorukov, V. E. Arkhipov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 475 (1999) [Phys. Solid State **41**, 426 (1999)].
44. A. S. Moskvin, E. V. Zenkov, Yu. D. Panov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **44**, 1452 (2002) [Phys. Solid State **44**, 1519 (2002)].
45. N. N. Loshkareva, Yu. P. Sukhorukov, E. V. Mostovshchikova, *et al.*, Zh. Éksp. Teor. Fiz. **121**, 412 (2002) [JETP **94**, 350 (2002)].
46. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley, New York, 1981; Mir, Moscow, 1984).
47. R. V. Demin, L. I. Koroleva, R. Szymszak, and H. Szymczak, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 402 (2002) [JETP Lett. **75**, 331 (2002)].
48. R. I. Zaĭnullina, N. G. Bebenin, V. V. Mashkautsan, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **45**, 1671 (2003) [Phys. Solid State **45**, 1754 (2003)].
49. M. O. Dzero, L. P. Gor'kov, and V. Z. Kresin, Eur. Phys. J. B **14**, 459 (2000).
50. S. de Brion, F. Ciorcas, G. Chouteau, *et al.*, Phys. Rev. B **59**, 1304 (1999).
51. U. Staub, G. I. Meijer, F. Fauth, *et al.*, Phys. Rev. Lett. **88**, 126402 (2002).
52. É. A. Neĭfel'd, V. E. Arkhipov, N. A. Tumalevich, and Ya. M. Mukovskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 630 (2001) [JETP Lett. **74**, 556 (2001)].
53. S. F. Dubinin, V. E. Arkhipov, Ya. M. Mukovskiĭ, *et al.*, Fiz. Met. Metalloved. **93** (3), 60 (2002) [Phys. Met. Metallogr. **93**, 248 (2002)].
54. P. Sheng, Philos. Mag. B **65**, 357 (1992).
55. P. Sheng, B. Abeles, and Y. Arie, Phys. Rev. Lett. **31**, 44 (1973).
56. E. Z. Meĭlikhov, Zh. Éksp. Teor. Fiz. **115**, 1484 (1999) [JETP **88**, 819 (1999)].
57. L. I. Glazman and M. É. Raĭkh, Pis'ma Zh. Éksp. Teor. Fiz. **47**, 378 (1988) [JETP Lett. **47**, 452 (1988)].
58. T. K. Ng and H. F. Lee, Phys. Rev. Lett. **61**, 1768 (1988).
59. K. Kikoin and Y. Avishai, Phys. Rev. Lett. **86**, 2090 (2001).
60. K. I. Kugel', A. L. Rakhmanov, A. O. Sboĭchikov, *et al.*, Zh. Éksp. Teor. Fiz. **125**, 648 (2004) [JETP **98**, 572 (2004)].
61. V. D. Okunev and N. N. Pafomov, Zh. Éksp. Teor. Fiz. **116**, 276 (1999) [JETP **89**, 151 (1999)].
62. V. V. Runov, D. Yu. Chernyshev, A. I. Kurbakov, *et al.*, Zh. Éksp. Teor. Fiz. **118**, 1174 (2000) [JETP **91**, 1017 (2000)].
63. I. Ya. Korenblit and E. F. Shender, Usp. Fiz. Nauk **126**, 233 (1978) [Sov. Phys. Usp. **21**, 832 (1978)].
64. É. L. Nagaev, Fiz. Tverd. Tela (St. Petersburg) **40**, 2069 (1998) [Phys. Solid State **40**, 1873 (1998)].
65. J. I. Gittleman, Y. Goldstein, and S. Bozowic, Phys. Rev. B **5**, 3609 (1972).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 145–155.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 168–178.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

On the Structure of the Two-Dimensional Spatially Periodic Inner Transition Layers in a Gas-Discharge Plasma

A. Ya. Voronov

Russian Federal Nuclear Center, All-Russia Scientific-Research Institute of Experimental Physics, Sarov, Nizhni Novgorod oblast, 607190 Russia

e-mail: otd4@expd.vniief.ru

Received November 25, 2003

Abstract—We investigate the structure of the spatially periodic inner boundary layers in the plasma of a positive glow-discharge column produced in a long cylindrical tube with an electropositive gas inside. Asymptotic methods, namely, the method of boundary functions, are used to analyze the initial mathematical model. We consider the formation of contrast burst-type structures. We have found all principal terms of the boundary-layer asymptotics of the solution. The results obtained are compared with the available probe measurements of basic physical parameters of ionization waves (strata) in neon at low pressures. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION. FORMULATION OF THE PROBLEM

(1) The separation of a gas-discharge plasma column into layers (strata) called ionization waves is familiar to specialists who use electric-discharge lasers and other gas-discharge devices in practice or in research. Being one of the main types of gas-discharge plasma instability, the phenomenon of stratification has long attracted the attention of researchers. Its properties have been studied in quite a few experimental and theoretical works (see, e.g., the reviews [1, 2] and the papers [3, 4]). There are several reasons for the heightened interest of researchers in ionization waves. From a scientific standpoint, this phenomenon attracts attention primarily as a classical example of the self-organization of a gas-discharge plasma, while from a practical standpoint, it can be used in reactors to separate gas mixtures [5] and to produce plasma-dust crystals [6] or in electric traps to collect dust particles with the goal of their subsequent removal from a gaseous medium [7]. Finally, strata severely degrade the performance of all gas-discharge devices, appreciably narrowing the range of their applications. Many properties of strata have been studied experimentally.

The goal of this paper is to theoretically study the structure of the stationary large-amplitude strata that emerge after the stratification of the plasma of a positive glow-discharge column at low pressures.

(2) Let a discharge take place in a cylindrical tube of radius R and length $L \gg R$ with circular plane-parallel electrodes. We introduce a cylindrical (r, ϕ, z) coordinate system whose origin coincides with the center of the anode and whose z axis is directed along the discharge axis from the anode to the cathode. The mathematical model of an infinitely long steady discharge in an electropositive gas with axial symmetry can be written in dimensionless form [8] as

$$\begin{gathered}
h\nabla^2 N_e + \nu\nabla(N_e\mathbf{V}) + h\gamma(I - \theta N_p N_e) = 0,\\
h\nabla^2 N_p - \nabla(N_e\mathbf{V}) + h(I - \theta N_p N_e) = 0,\\
h\nabla\mathbf{V} = N_p - N_e, \quad \nabla\times\mathbf{V} = 0,\\
\frac{\partial N_e}{\partial x} = 0, \quad N_p = q(y, h), \quad V_x = 0, \quad x = 0,\\
N_a = b_a(y, h), \quad V_x = V_x^{(1)}, \quad x = 1,\\
a = e, p.
\end{gathered} \tag{1.1}$$

In problem (1.1) and below, we use the following notation:

$$0 \le x = \frac{r}{R} \le 1, \quad 0 \le y = \frac{z}{R} \le \frac{L}{R},$$

$$\frac{D_p}{D_e} = \gamma, \quad \frac{T}{T_e} = \nu,$$

$$\frac{c_a}{c_p^{(0)}} = N_a, \quad R\frac{k_p}{D_p}\mathbf{E} = \boldsymbol{\varepsilon} = -\nabla\Phi,$$

$$\alpha c_p^{(0)}\frac{R^2}{D_p} = \theta, \quad \mathbf{V} = h\boldsymbol{\varepsilon}, \quad h = \frac{r_D}{R},$$

$$\nabla^2 = \frac{1}{x}\frac{\partial}{\partial x}\left(x\frac{\partial}{\partial x}\right) + \frac{\partial^2}{\partial y^2},$$

1063-7761/05/10101-0145$26.00

c_e and c_p are the concentrations of the free plasma electrons and singly charged positive ions, respectively, D_a and k_a are their coefficients of free diffusion and mobility, $\mathbf{E}$ is the electric field strength, Φ is the dimensionless potential of this field, $c_p^{(0)}$ is some characteristic concentration of the positive ions, r_D is the Debye length for these ions, and α is the bulk electron–ion recombination coefficient. The function $I = I(x, y; N_e, h)$ defines the ionization rate, i.e., the number of pairs of oppositely charged particles emerging in a unit plasma volume per unit time. The coefficients D_a and k_a as well as the temperatures of the free plasma electrons, T_e, and the neutral gaseous medium, T, are assumed to be constant throughout the plasma volume. The equations of system (1.1) were written by assuming that there was no convective motion and magnetic field.

Note that a small parameter, $0 < h \ll 1$, appears before the high derivatives in the equations of system (1.1) [8]. Consequently, asymptotic methods of the theory of singularly perturbed differential equations, for example, the method of boundary functions [9], can be used to analyze problem (1.1).

We restrict our analysis to a long tube and assume the positive glow-discharge column to be infinitely long ($-\infty < y < \infty$). In this case, the effect of the electrodes on the object under study can be disregarded.

2. ASYMPTOTICS OF THE SOLUTION IN PARAMETER h. A REGULAR SERIES

2.1. Let the discharge column be stratified and stationary ionization waves of length l_{st}, which is assumed to be known, appear in it. This length can be measured experimentally or estimated theoretically [3]. The function $I(x, y; N_e, h)$ is also assumed to be known and l-periodic in variable y:

$$I(x, y; N_e, h) = I(x, y + l; N_e, h), \quad l = l_{st}/R. \quad (2.1)$$

This function is written in explicit form below. We assume the functions $N_a(x, y, h)$ and $\mathbf{V}(x, y, h)$, i.e., the concentrations of the charged plasma components and the electric field vector, to be unknown in problem (1.1). Note that the values of these functions on the tube wall ($x = 1$), i.e., b_a and $V_x^{(1)}$ (the radial electric field component) as well as $q(y, h)$ (the ion concentration on the discharge axis), are assumed to be known.

Denote the entire set of sough-for quantities by $W(x, y, h)$. The asymptotics of the solution of problem (1.1) can then be written as [8, 9]

$$W(x, y, h) = w(x, y, h) + \Pi w(x_1, y, h) + Rw(x_2, y, h), \quad (2.2)$$

where $x_1 = x/h$ and $x_2 = (x - 1)/h$ are fast variables.

Let us require that all variable quantities be l-periodic, i.e., that the following condition be satisfied:

$$W(x, y, h) = W(x, y + l, h). \quad (2.3)$$

We seek the regular part of asymptotics (2.2) in the form

$$w(x, y, h) = \sum_{s=0}^{\infty} h^s w_s(x, y) = w_0(x, y) + hw_1(x, y) + \dots . \quad (2.4)$$

The expansions of the boundary functions $\Pi w(x_1, y, h)$ and $Rw(x_2, y, h)$ have a similar form; the former and the latter can contribute appreciably to the value of the principal function only near $x = 0$ and $x = 1$, respectively. The coefficients of these expansions, i.e., the functions $w_s(x, y)$, $\Pi_s w(x_1, y)$, and $R_s w(x_2, y)$, $s = 0, 1, 2, \dots$, can be determined in a standard way. Thus, for example, when determining the coefficients $w_s(x, y)$ of the regular series (2.4), at the first step ($s = 0$), we obtain a system that is degenerate ($h = 0$) with respect to the system of equations of problem (1.1) to determine the unknown functions $\mathbf{v}_0(x, y)$, $n_0^{(e)}(x, y)$, and $n_0^{(p)}(x, y)$. The solution of this system is obvious [8]:

$$\mathbf{v}_0(x, y) \equiv 0, \quad n_0^{(e)}(x, y) \equiv n_0^{(p)}(x, y) = A_0(x, y).$$

Here, $A_0(x, y)$ is an arbitrary function.

At the next step ($s = 1$), we obtain the following problem to determine the unknown functions $A_0(x, y)$ and $\mathbf{v}_1(x, y)$:

$$\begin{gathered} \nabla^2 A_0 + \nu \nabla(A_0 \mathbf{v}_1) + \gamma(i_0 - \theta A_0^2) = 0, \\ \nabla^2 A_0 - \nabla(A_0 \mathbf{v}_1) + i_0 - \theta A_0^2 = 0, \quad \mathbf{v}_1 = -\nabla\varphi_0, \\ A_0 = q_0(y), \quad \frac{\partial \varphi_0}{\partial x} = 0, \quad x = 0; \\ A_0 = b_0(y), \quad \frac{\partial \varphi_0}{\partial x} = -\nu_{1x}^{(1)}, \quad x = 1. \end{gathered} \quad (2.5)$$

Here, $\nu_{1x}^{(1)} = \nu_{1x}(1, y)$, $i_0 = i_0(x, y)$, and $\varphi_0 = \varphi_0(x, y)$ are the principal nonzero terms in expansions (2.4) for the functions $V_x(1, y, h)$, $I(x, y, h)$, and $\Phi(x, y, h)$, respectively; the first two functions are assumed to be known. Having eliminated the terms with the vector $\mathbf{v}_1(x, y)$ from the equations of this problem, we obtain the following problem to determine the unknown function $A_0(x, y, \mu)$:

$$\begin{gathered} \mu^2 \nabla^2 A_0 = A_0^2 - I_0, \\ (x, y) \in D = [(0 < x < 1) \times (-\infty < y < \infty)], \\ A_0(0, y, \mu) = q_0(y), \quad A_0(1, y, \mu) = b_0(y), \end{gathered} \quad (2.6)$$

where

$$\mu^2 = \frac{1+\nu}{\nu\theta(1+\sigma)}, \quad \frac{k_p}{k_e} = \sigma, \quad \theta I_0 = i_0, \quad q_0(y) = q(y, 0).$$

Since, in general, one function A_0 cannot satisfy two boundary values, $b_0^{(e)}$ and $b_0^{(p)}$, we introduce a new unknown parameter, $b_0(y)$, in the additional conditions of the boundary-value problems (2.5) and (2.6) that can depend on μ and must be l-periodic. The relationship between the parameters $b_0(y)$, $b_0^{(e)}(y)$, and $b_0^{(p)}(y)$ is established when determining the boundary functions $Rw(x_2, y, h)$.

Let us consider the case where $0 < \mu \ll 1$ (at R = 1 cm, p = 1 Torr, $\nu = 10^{-2}$, $\sigma = 10^{-3}$, $\alpha = 2 \times 10^{-7}$ cm^3 s^{-1}, and $c_p^{(0)} > 10^{12}$ cm^{-3}, we have $\mu \le 0.3$).

The equation of problem (2.6) is a reaction–diffusion equation. As was shown in [10], when certain conditions are satisfied, the equations of this type have l-periodic (in variable y) solutions with a burst (contrast burst-type structures) whose vertex lies on some smooth l-periodic curve C (the burst curve). Let us denote

$$F(x, y; A_0, \mu) = A_0^2 - I_0$$

and write the expression defining the function $I_0(x, y; A_0, \mu)$ as

$$I_0(x, y; A_0, \mu) = g(x, y, \mu)A_0 + \xi(x, y, \mu)A_0^3,$$

where $g(x, y, \mu)$ is the ionization frequency of neutrals by a direct electron impact and $\xi(x, y, \mu)$ is the step ionization frequency. These quantities are assumed to be known and l-periodic in variable y.

Thus, we seek an l-periodic (in variable y) solution of problem (2.6) that at small μ is close to some solution $\alpha_{00}(x, y)$ of the degenerate ($\mu = 0$) equation

$$\begin{aligned} F(x, y; \alpha_{00}, 0) &= f(x, y; \alpha_{00}, 0) \\ &= \alpha_{00}^2 - g(x, y, 0)\alpha_{00} - \xi(x, y, 0)\alpha_{00}^3 = 0 \end{aligned} \tag{2.7}$$

everywhere inside a closed domain $\bar{D}$, except the small neighborhood of the l-periodic curve C where the burst of the solution occurs.

We construct the asymptotics of the solution of problem (2.6) in parameter μ as the sum of four series:

$$\begin{aligned} A_0(x, y, \mu) &= \alpha_0(x, y, \mu) + T\alpha_0(\tau_0, y, \mu) \\ &+ Q\alpha_0(\tau, \zeta, \mu) + S\alpha_0(\tau_1, y, \mu), \end{aligned} \tag{2.8}$$

where

$$\alpha_0(x, y, \mu) = \alpha_{00}(x, y) + \mu\alpha_{01}(x, y) + \ldots$$

is the regular series,

$$T\alpha_0(\tau_0, y, \mu) = T_0\alpha_0(\tau_0, y) + \mu T_1\alpha_0(\tau_0, y) + \ldots,$$

$$\tau_0 = \frac{x}{\mu},$$

$$S\alpha_0(\tau_1, y, \mu) = S_0\alpha_0(\tau_1, y) + \mu S_1\alpha_0(\tau_1, y) + \ldots,$$

$$\tau_1 = \frac{x-1}{\mu}$$

are the series that describe the boundary layers in the neighborhood of the boundaries of domain D, i.e., at $x = 0$ and $x = 1$, respectively, and

$$Q\alpha_0(\tau, \zeta, \mu) = Q_0\alpha_0(\tau, \zeta) + \mu Q_1\alpha_0(\tau, \zeta) + \ldots$$

is the series that describes the burst of the solution near curve C; the variables τ and ζ are described below. The unknown burst curve is determined when constructing asymptotics (2.8).

2.2. Let us determine the terms of the regular series and check whether a solution with a burst exists for problem (2.7). As was noted above, $\alpha_{00}(x, y)$ is a root of Eq. (2.7). In total, this equation has three roots:

$$\alpha_{00}^{(1)}(x, y) = \frac{1}{2\xi_0}[1 - \sqrt{1 - 4g_0\xi_0}] = \varphi_1(x, y),$$

$$\alpha_{00}^{(2)}(x, y) = \frac{1}{2\xi_0}[1 + \sqrt{1 - 4g_0\xi_0}] = \varphi_2(x, y),$$

$$\alpha_{00}^{(3)}(x, y) \equiv 0, \quad \xi_0 = \xi(x, y, 0), \quad g_0 = g(x, y, 0).$$

As α_{00}, we choose φ_1; i.e., we assume that

$$\alpha_{00}(x, y) \equiv \varphi_1(x, y) > 0,$$

$$(x, y) \in \bar{D} = [(0 \le x \le 1) \times (-\infty < y < \infty)].$$

It follows from the form of the function $\varphi_1(x, y)$ that the inequality

$$g_0\xi_0 < 0.25$$

must hold. Clearly, in this case,

$$\varphi_1(x, y) < \varphi_2(x, y), \quad g_0 \le \varphi_1(x, y) < 2g_0,$$

and it follows from (2.7) that the partial derivatives of the function $f(x, y; \alpha_{00}, 0)$ with respect to the argument α_{00} when this argument is substituted with $\varphi_1(x, y)$ and $\varphi_2(x, y)$ satisfies the conditions

$$f'_\alpha(x, y; \varphi_1, 0) > 0, \quad f'_\alpha(x, y; \varphi_2, 0) < 0, \quad x, y \in \bar{D}.$$

In addition, there exists such a function $\psi(x, y)$ that

$$\int_{\varphi_1(x,y)}^{\psi(x,y)} f(x, y; \alpha_{00}, 0)\,d\alpha_{00} = \frac{1}{3}\psi^3 - \frac{1}{2}g_0\psi^2 - \frac{1}{4}\xi_0\psi^4 - d = 0, \quad x, y \in \bar{D}, \tag{2.9}$$

where

$$d = -\frac{g_0}{2}\varphi_1^2 + \frac{1}{3}\varphi_1^3 - \frac{1}{4}\xi_0\varphi_1^4,$$

and Eq. (2.7) has no roots other than φ_1 and φ_2 in the integration segment $[\varphi_1, \psi]$. We have $f(x, y; \alpha, 0) < 0$ for $\varphi_2 < \alpha < \psi$. Thus, we can assert that problem (2.6) has an l-periodic solution with a burst that is close to $\varphi_1(x, y)$ in $\bar{D}$ at small μ everywhere, except the small neighborhood of the burst curve C [10].

It can be easily verified that Eq. (2.9) has three different real roots $\delta(x, y) < \varphi_1(x, y) < \psi(x, y)$, one of which (the function $\varphi_1(x, y)$) is known and has a double multiplicity, while the other two can be found by solving a quadratic equation [8].

The succeeding terms of the regular series α_{0s}, $s \geq 1$, can be easily determined. Thus, for example,

$$\alpha_{01}(x, y) = \frac{f'_\mu(x, y; \varphi_1(x, y), 0)}{f'_\alpha(x, y; \varphi_1(x, y), 0)} = \frac{g_1(x, y)\varphi_1 + \xi_1(x, y)\varphi_1^3}{2\varphi_1 - g_0 - 3\xi_0\varphi_1^2},$$

where

$$g_1(x, y) = g'_\mu(x, y, 0), \quad \xi_1(x, y) = \xi'_\mu(x, y, 0).$$

3. THE BOUNDARY FUNCTIONS OF ASYMPTOTICS (2.8)

3.1. Let us determine the terms of the series $T\alpha_0(\tau_0, y, \mu)$, $\tau_0 = x/\mu$. The problem defining the principal term $T_0\alpha_0(\tau_0, y)$ of this series is

$$\frac{1}{\tau_0}\frac{\partial}{\partial\tau_0}\left(\tau_0\frac{\partial T_0\alpha_0}{\partial\tau_0}\right) = F(0, y; \alpha_{00}(0, y) + T_0\alpha_0, 0) - f(0, y; \alpha_{00}(0, y), 0) = \omega_0^2 - g_0\omega_0 - \xi_0\omega_0^3, \quad 0 < \tau_0 < \infty, \tag{3.1}$$

$$T_0\alpha_0(0, y) = q_0(y) - \varphi_1(0, y), \quad T_0\alpha_0(\infty, y) = 0, \quad \omega_0 = \omega_0(\tau_0, y) = \varphi_1(0, y) + T_0\alpha_0(\tau_0, y). \tag{3.2}$$

Equation (3.1) is an ordinary nonlinear differential equation in which y is a parameter of the problem.

Let us substitute problem (3.1), (3.2) with an equivalent integral equation:

$$\begin{aligned} Z(t, y) = {} & \int_0^\infty \exp(-mx^2)[B_2\exp(-2mx^2) - B_1\exp(-mx^2) + 4m^2x^2 - 4m - 1]G(t, x)\,dx \\ & + \int_0^\infty [1 - 2B_1x^2\exp(-mx^2) + 3B_2x^2\exp(-2mx^2)] \times Z(t, x)G(t, x)\frac{dx}{x^2} \\ & + \int_0^\infty (3B_2\exp(-mx^2) - B_1)Z^2(t, x)G(t, x)\,dx \\ & + \int_0^\infty B_2Z^3(t, x)G(t, x)\,dx. \end{aligned} \tag{3.3}$$

Here,

$$Z(t, y) = \frac{T_0\alpha_0(t, y)}{q_0 - \varphi_1^0} - \exp[-m(y)t^2],$$

where $m(y)$ is an unknown parameter, which, in general, can depend on y,

$$t = \sqrt{(1 - 2\xi_0\varphi_1^0)\varphi_1^0}\,\tau_0,$$

$$B_1 = \frac{1 - 3\xi_0^0\varphi_1^0}{1 - 2\xi_0^0\varphi_1^0}\frac{q_0 - \varphi_1^0}{\varphi_1^0},$$

$$B_2 = \frac{\xi_0^0\varphi_1^0}{1 - 2\xi_0^0\varphi_1^0}\left(\frac{q_0 - \varphi_1^0}{\varphi_1^0}\right)^2,$$

$$\xi_0^0 = \xi_0(0, y) = \xi(0, y, 0), \quad \varphi_1^0 = \varphi_1(0, y),$$

$$G(t, x) = xK_1(t)I_1(x), \quad x \leq t;$$

$$G(t, x) = xK_1(x)I_1(t), \quad t \leq x,$$

$K_1(x)$ and $I_1(x)$ are the Bessel functions of the first kind of an imaginary argument. Obviously, the following inequality holds:

$$0 \leq q_0(y) \leq \psi(0, y) = \psi^0.$$

Note one feature of the integral equation (3.3) that has a specific physical meaning. In the absence of step ionization, i.e., at $\xi_0^0 = 0$ ($B_2 = 0$), this equation has the required positive solution with an exponential estimate for $t \longrightarrow \infty$ only for $q_0 < \varphi_1^0$ ($B_1 < 0$) [8]. However, $q_0 > \varphi_1^0$ near the burst curve. In this case, Eq. (3.3) can have a positive solution only at $\xi_0^0 > 0$; the larger the excess of q_0 over φ_1^0, the higher the step ionization frequency. As we show below, a large increase in the electric field produced by the plasma space charges can provide this increase in ξ_0^0. The solution of the integral equation (3.3) itself can be constructed using the method of successive approximations. Note yet another property of this equation that is used below. In view of the axial symmetry,

$$\frac{\partial Y(0, y)}{\partial t} = 0,$$

and

$$\frac{\partial^2 Y(0, y)}{\partial t^2} = \frac{1}{2}(1 + B_1 + B_2),$$

where

$$Y(t, y) = \frac{T_0\alpha_0(t, y)}{q_0 - \varphi_1^0}.$$

The explicit form of all the succeeding terms $T_s\alpha_0(\tau_0, y)$, $s \geq 1$, can be found in a similar way.

3.2. The boundary-value problem defining $S_0\alpha_0(\tau_1, y)$, $\tau_1 = (x - 1)/\mu$, is

$$\frac{\partial^2 S_0\alpha_0}{\partial\tau_1^2} = F(1, y; \alpha_{00}(1 + y) + S_0\alpha_0, 0)$$

$$-f(1, y; \alpha_{00}(1, y), 0) = \omega_1^2 - g_0^{(1)}\omega_1 - \xi_0^{(1)}\omega_1^{(3)}, \quad (3.4)$$

$$-\infty < \tau_1 < 0,$$

$$S_0\alpha_0(0, y) = b_0(y) - \varphi_1^{(1)}, \quad S_0\alpha_0(-\infty, y) = 0.$$

Here,

$$\omega_1 = \omega_1(\tau_1, y) = \alpha_{00}(1, y) + S_0\alpha_0(\tau_1, y),$$

$$\xi_0^{(1)} = \xi(1, y, 0), \quad \varphi_1^{(1)} = \varphi_1(1, y), \quad g_0^{(1)} = g(1, y, 0).$$

We write the solution of problem (3.4) as

$$\frac{2}{\sqrt{(\psi_1 - \varphi_1^{(1)})(\varphi_1^{(1)} - \delta_1)}} \ln \frac{\sqrt{\dfrac{\psi_1 - b_0}{(\psi_1 - \varphi_1^{(1)})\left|\varphi_1^{(1)} - b_0\right|}} + \sqrt{\dfrac{b_0 - \delta_1}{(\varphi_1^{(1)} - \delta_1)\left|\varphi_1^{(1)} - b_0\right|}}}{\sqrt{\dfrac{\psi_1 - \omega_1}{(\psi_1 - \varphi_1^{(1)})\left|\varphi_1^{(1)} - \omega_1\right|}} + \sqrt{\dfrac{\omega_1 - \delta_1}{(\varphi_1^{(1)} - \delta_1)\left|\varphi_1^{(1)} - \omega_1\right|}}} = \sqrt{\frac{\xi_0^{(1)}}{2}}\tau_1, \quad (3.5)$$

$$\psi_1 = \psi(1, y), \quad \delta_1 = \delta(1, y).$$

3.3. Let us now find the terms of the boundary series $Q\alpha_0(\tau, \zeta, \mu)$. First, note that we also seek the burst curve C in the form of an expansion in terms of positive powers of the small parameter μ. The principal term of this expansion (curve C_0) is sought in the form $x = \chi(y)$, $\chi(y)$ is the principal l-periodic function, such that $0 < \chi(y) < 1$.

Following the approach described in [10], we introduce a local (ρ, ζ) coordinate system in the neighborhood of C_0, where $|\rho|$ is the distance from the current point $M(x, y)$ to curve C_0 measured along the normal to C_0, and ζ is the ordinate of the base of this normal on curve C_0. For a fairly small neighborhood of curve C_0, the one-to-one correspondence between the coordinates x, y and ρ, ζ is defined by the formulas

$$x = \chi(\zeta) + \frac{\rho}{\sqrt{1 + (\chi')^2}}, \quad y = \zeta - \frac{\rho\chi'(\zeta)}{\sqrt{1 + (\chi')^2}}.$$

In this case, ρ must be assumed to be positive if point M is to the right of C_0 and negative if it is to the left of C_0. We seek an equation for the burst curve C in the local coordinate system in the form

$$\rho = \lambda(\zeta, \mu) = \mu\lambda_1(\zeta) + \mu^2\lambda_2(\zeta) + \ldots,$$

where the l-periodic functions $\lambda_i(\zeta)$ are to be determined. The following condition is used to find them, i.e., to determine curve C:

$$\left.\frac{\partial A_0}{\partial\rho}\right|_C = 0. \quad (3.7)$$

This condition implies that the solution of problem (2.6) as a function of the variable ρ has an extremum on curve C.

Let us introduce an extended variable

$$\tau = \frac{\rho - \lambda(\zeta, \mu)}{\mu}$$

and the following designation convenient in theoretical constructions:

$$\bar{w}(\rho, \zeta, \mu) = w(x(\rho, \zeta), y(\rho, \zeta), \mu).$$

Here, as usual, the entire set of variable quantities is denoted by $w(x, y)$. Let us now write out the boundary condition that the function $Q\alpha_0(\tau, \zeta, \mu)$ must satisfy at $\tau = 0$. It follows from (3.7) that

$$\left.\frac{\partial \bar{\alpha}_0}{\partial \rho}\right|_{\rho = \lambda(\zeta, \mu)} + \frac{1}{\mu}\left.\frac{\partial Q\alpha_0}{\partial \tau}\right|_{\tau = 0} = 0.$$

Using the expression for $\nabla^2 Q\alpha_0$ derived in [10] and defining $Q_0\alpha_0(\tau, \zeta)$ in the new notation, we obtain the problem

$$\frac{\partial^2 Q_0\alpha_0}{\partial \tau^2} = \bar{F}(0, \zeta; \alpha_{00}(0, \zeta) + Q_0\alpha_0, 0)$$

$$= \omega_C^2 - g_0^C\omega_C - \xi_0^C\omega_C^3, \quad -\infty < \tau < \infty, \qquad (3.8)$$

$$\frac{\partial Q_0\alpha_0(0, \zeta)}{\partial \tau} = 0, \quad Q_0\alpha_0(\pm\infty, \zeta) = 0,$$

where

$$\omega_C = \omega(\tau, \zeta) = \bar{\alpha}_{00}(0, \zeta) + Q_0\alpha_0(\tau, \zeta).$$

It was shown in [10] that problem (3.8) for any l-periodic function $\chi(\zeta)$ has a nontrivial l-periodic (in variable ζ) solution with an exponential estimate. The unknown function $\chi(\zeta)$ is determined at the next step. We write the integral of problem (3.8) as

$$\frac{2}{\sqrt{(\psi_C - \varphi_{1C})(\varphi_{1C} - \delta_C)}}\ln\tan\left(\frac{\Lambda}{2} + \frac{\pi}{4}\right) = \sqrt{\frac{\xi_0^C}{2}}|\tau|,$$

$$\Lambda = \arcsin\left[\frac{(\varphi_{1C} - \delta_C)(\psi_C - \omega_C)}{(\psi_C - \varphi_{1C})(\omega_C - \delta_C)}\right]^{1/2}, \qquad (3.9)$$

$$\psi_C = \psi(\chi(\zeta), \zeta), \quad \delta_C = \delta(\chi(\zeta), \zeta),$$

$$\varphi_{1C} \le \omega_C \le \psi_C.$$

Let us formulate the problem defining $Q_1a_0(\tau, \zeta)$:

$$\frac{\partial^2 Q_1\alpha_0}{\partial \tau^2} = F'_\alpha(\tau, \zeta)Q_1\alpha_0 - \beta\frac{\partial Q_0\alpha_0}{\partial \tau} + F_1(\tau, \zeta),$$

$$-\infty < \tau < \infty,$$

$$\left.\frac{\partial Q_1\alpha_0}{\partial \tau}\right|_{\tau = 0} = -\left.\frac{\partial \bar{\alpha}_{00}}{\partial \rho}\right|_{\rho = 0}, \qquad (3.10)$$

$$Q_1\alpha_0(\pm\infty, \zeta) = 0.$$

Here,

$$F'_\alpha(\tau, \zeta) = 2\omega_C - g_0^C - 3\xi_0^C\omega_C^2, \quad \beta = \frac{1 + (\chi')^2 - \chi\chi''}{\chi[1 + (\chi')^2]^{3/2}},$$

$$F_1(\tau, \zeta) = \left[(F'_\alpha - f'_\alpha)\frac{\partial \bar{\alpha}_{00}(0, \zeta)}{\partial \rho}\right.$$

$$\left. + (F'_\xi - f'_\xi)\frac{\partial \bar{\xi}(0, \zeta)}{\partial \rho} + (F'_g - f'_g)\frac{\partial \bar{g}}{\partial \rho}\right]$$

$$\times(\lambda_1(\zeta) + \tau) + (F'_\alpha - f'_\alpha)\bar{\alpha}_{01}(0, \zeta)$$

$$+ (F'_\xi - f'_\xi)\bar{\xi}_1(0, \zeta) + (F'_g - f'_g)\bar{g}_1(0, \zeta),$$

$$F'_\alpha - f'_\alpha = 2\omega_C - g_0^C - 3\xi_0^C\omega_C^2 - (2\alpha_{00} - g_0^C - 3\xi_0^C\alpha_{00}^2).$$

Since the homogeneous problem corresponding to (3.10) has a nontrivial solution, $\partial Q_0\alpha_0/\partial\tau$, the initial problem (3.10) is solvable only if the inhomogeneity $F_1 - \beta\partial Q_0\alpha_0/\partial\tau$ is orthogonal to the solution of the homogeneous equation. Using the expression for $F_1(\tau, \zeta)$ and the property of an integral of even and odd functions with symmetric limits, we can simplify the solvability condition by reducing it to the form

$$\left.\frac{\partial \bar{\xi}}{\partial \rho}\right|_{\rho = 0}\int_{-\infty}^{\infty}(F'_\xi - f'_\xi)\tau\frac{\partial Q_0\alpha_0}{\partial \tau}d\tau$$

$$= \beta\int_{-\infty}^{\infty}\left(\frac{\partial Q_0\alpha_0}{\partial \tau}\right)^2 d\tau. \qquad (3.11)$$

Equation (3.11) is a nonlinear differential equation of the second order for the unknown function $\chi(\zeta)$. The functions $\partial Q_0\alpha_0/\partial\tau$ and $\alpha_{00}^3(\chi(\zeta), \zeta) - \omega_C^3$ in the integrands of this equation also depend on $\chi(\zeta)$ and are l-periodic in explicit variable ζ. Therefore, the following condition must be added to Eq. (3.11):

$$\chi(\zeta) = \chi(\zeta + l).$$

Note that in deriving Eq. (3.11), for the sake of simplicity, we assumed that $g(x, y, \mu) \equiv \text{const}$. However, we cannot assume that $\xi(x, y, \mu) = \text{const}$, since Eq. (3.11) loses its meaning in this case. As was noted above, the electric field of the plasma space charges can provide the dependence of the step ionization frequency ξ on point $M(x, y)$. Curve C_0 can be determined by solving Eq. (3.11), i.e., by finding the function $0 < \chi(\zeta) < 1$. In this way, all terms of the asymptotics of the zeroth order in parameter μ are determined for the function $A_0(x, y, \mu)$.

4. ASYMPTOTICS OF THE SOLUTION FOR $\mathbf{v}_1(x, y, \mu)$

4.1. The problem defining the vector $\mathbf{P}(x, y, \mu) = \mu\mathbf{v}_1(x, y, \mu)$ follows from problem (2.5):

$$\nabla(A_0\mathbf{P}) = \frac{\gamma-1}{\nu(1+\sigma)}\mu\nabla^2 A_0, \quad \nabla\times\mathbf{P} = 0,$$
$$P_x(0, y, \mu) = 0, \quad P_x(1, y, \mu) = P_x^{(1)}, \tag{4.1}$$
$$\mathbf{P}(x, y, \mu) = \mathbf{P}(x, y + l, \mu).$$

We write the solution of this problem as

$$\mathbf{P}(x, y, \mu) = \frac{\gamma-1}{\nu(1+\sigma)}\mu\nabla\ln A_0(x, y, \mu) + \mu\mathbf{B} = -\mu\nabla\varphi_0, \tag{4.2}$$
$$\varphi_0(x, y, \mu) = -\frac{\gamma-1}{\nu(1+\sigma)}\ln A_0 + B_z y + \text{const}.$$

Here, $\mathbf{B}$ is an arbitrary vector with $B_x = 0$. Relations (4.2) allow us to determine all terms of the asymptotics for $\mathbf{P}$ and φ_0, which also consist of the sum of four series of type (2.8):

$$\mathbf{P}(x, y, \mu) = \mathbf{p}(x, y, \mu) + T\mathbf{p}(\tau_0, y, \mu) + Q\mathbf{p}(\tau, \zeta, \mu) + S\mathbf{p}(\tau_1, y, \mu),$$
$$\varphi_0(x, y, \mu) = \sum_{s=0}^{\infty}\mu^s\varphi_{0s}(x, y) + T\varphi_0(\tau_0, y, \mu) + Q\varphi_0(\tau, \zeta, \mu) + S\varphi_0(\tau_1, y, \mu).$$

In determining the terms of the regular series, we have

$$\mathbf{p}_0(x, y) \equiv 0, \quad \mathbf{p}_1(x, y) = \frac{\gamma-1}{\nu(1+\sigma)}\nabla\ln\alpha_{00} + \mathbf{B},$$
$$\varphi_{00}(x, y) = \frac{1-\gamma}{\nu(1+\sigma)}\ln\alpha_{00}(x, y) - \mathscr{E}_0 y + \text{const}, \tag{4.3}$$

$\mathscr{E}_0 = -B_z$ is the constant electric field component directed along the discharge axis. In expressions (4.3), the constant is chosen in such a way that the potential φ_{00} at a given point has a known value. All succeeding terms of the regular series can be determined in a similar way.

4.2. Let us find the terms of the boundary series. It follows from (4.2) that

$$T_0\mathbf{p}(\tau_0, y) = \frac{\gamma-1}{\nu(1+\sigma)}\mu\nabla\ln\frac{\omega_0(\tau_0, y)}{\alpha_{00}(0, y)},$$
$$T_0\varphi_0(\tau_0, y) = \frac{1-\gamma}{\nu(1+\sigma)}\ln\frac{\omega_0(\tau_0, y)}{\alpha_{00}(0, y)}. \tag{4.4}$$

Similarly,

$$Q_0\mathbf{p}(\tau, \zeta) = \frac{\gamma-1}{\nu(1+\sigma)}\mu\nabla\ln\frac{\omega_c(\tau, \zeta)}{\bar{\alpha}_{00}(0, \zeta)}.$$

Hence,

$$Q_0 p_x(\tau, \zeta) = \frac{\gamma-1}{\nu(1+\sigma)}\frac{1}{\sqrt{1+(\chi')^2}}\frac{\partial}{\partial\tau}\ln\frac{\omega_c(\tau, \zeta)}{\bar{\alpha}_{00}(0, \zeta)},$$
$$Q_0 p_z(\tau, \zeta) = \frac{1-\gamma}{\nu(1+\sigma)}\frac{\chi'(\zeta)}{\sqrt{1+(\chi')^2}}\frac{\partial}{\partial\tau}\ln\frac{\omega_c(\tau, \zeta)}{\bar{\alpha}_{00}(0, \zeta)}. \tag{4.5}$$

We see from (4.5) that the magnitude of the vector $Q_0\mathbf{p}(\tau, \zeta)$ depends on the shape of the burst curve. Thus, for example, if this curve is a rectilinear segment, $0 < x < 1$, $y = \text{const}$, then

$$Q_0 p_x(\tau, \zeta) \equiv 0,$$
$$Q_0 p_z(\tau, \zeta) = \frac{1-\gamma}{\nu(1+\sigma)}\frac{\partial}{\partial\tau}\ln\frac{\omega_c(\tau, \zeta)}{\bar{\alpha}_{00}(0, \zeta)}, \tag{4.6}$$
$$Q_0\varphi_0(\tau, \zeta) = \frac{1-\gamma}{\nu(1+\sigma)}\ln\frac{\omega_c(\tau, \zeta)}{\bar{\alpha}_{00}(0, \zeta)}.$$

Finally,

$$S_0 p(\tau_1, y) = \frac{\gamma-1}{\nu(1+\sigma)}\mu\nabla\ln\frac{\omega_1(\tau_1, y)}{\alpha_{00}(1, y)},$$
$$S_0\varphi_0(\tau_1, y) = -\frac{\gamma-1}{\nu(1+\sigma)}\ln\frac{\omega_1(\tau_1, y)}{\alpha_{00}(1, y)}. \tag{4.7}$$

Consequently,

$$\omega_1(\tau_1, y) = \alpha_{00}(1, y)\exp\left(\nu\frac{1+\sigma}{1-\gamma}S_0\varphi_0\right),$$
$$b_0(y) = \varphi_1^{(1)}\exp\left[\nu\frac{1+\sigma}{1-\gamma}S_0\varphi_0(0, y)\right]. \tag{4.8}$$

Similarly,

$$T_0\alpha_0(\tau_0, y) = \varphi_1^{(0)}\left[\exp\left(\nu\frac{1+\sigma}{1-\gamma}T_0\varphi_0\right) - 1\right],$$
$$q_0(y) = \varphi_1^{(0)}\exp\left[\nu\frac{1+\sigma}{1-\gamma}T_0\varphi_0(0, y)\right]. \tag{4.9}$$

Relation (4.8) establishes the physical meaning of the unknown parameter $b_0(y)$ introduced above. Finally,

$$\psi_C = \varphi_{1C}\exp\left(\nu\frac{1+\sigma}{1-\gamma}Q_0\varphi_0(0, \zeta)\right). \tag{4.10}$$

Thus, we have found the principal terms of the asymptotics in parameter μ for all unknown functions.

4.3. The problems defining the principal terms of the boundary series $\Pi w(x_1, y, h)$ and $Rw(x_2, y, h)$ are formulated so as to remove the residuals that can be introduced in the boundary conditions of problem (1.1) when finding the function $A_0(x, y, \mu)$ at both ends of the segment [0, 1]. Since we managed to remove all the emerging residuals at the left end of this segment using the functions $T_0 w(\tau_0, y)$, we have $\Pi w(x_1, y, h) \equiv 0$.

The boundary functions $Rw(x_2, y, h)$ are determined in [11]. As was noted above, when finding these functions, we establish the relationship between the parameter $b_0(y)$ and the concentrations of the charged plasma components on the inner surface of the tube wall $b_0^{(a)}$, $a = e$ and p, which, according to the condition of problem (1.1), are assumed to be known. Therefore, we assume that the formulated problem has been solved.

5. DISCUSSION OF RESULTS. COMPARISON WITH EXPERIMENTAL DATA. CONCLUSIONS

Below, we discuss the main results of the solution of the formulated problem obtained in previous sections and analyze the plots of the functions that describe the distributions of basic physical parameters of a gas-discharge plasma: the concentrations of its charged components, the electric field strength, the potential, and the space charges. Subsequently, we compare the results of this analysis with the probe measurements for the same quantities in moving strata. We do not analyze the wall boundary layer, i.e., do not use the functions $Rw(x_2, y, h)$) in our calculations. Our analysis of the structure of the strata is based only on the solution of problem (2.5). Let us consider the plots of the following functions:

the ion concentration

$$N_p(x, y, h, \mu) \approx A_0(x, y, \mu) \approx \alpha_{00}(x, y) + T_0\alpha_0(\tau_0, y) + Q_0\alpha_0(\tau, \zeta) + S_0\alpha_0(\tau_1, y); \quad (5.1)$$

the potential and the electric field

$$\Phi(x, y, h, \mu) \approx \varphi_0(x, y, \mu) \approx \varphi_{00}(x, y) + T_0\varphi_0(\tau_0, y) + Q_0\varphi_0(\tau, \zeta) + S_0\varphi_0(\tau_1, y), \quad (5.2)$$

$$\varepsilon_x(x, y, h, \mu) = \frac{1}{h}V_x \approx v_{1x}(x, y, \mu) \approx p_{1x}(x, y) + \frac{1}{\mu}[T_0 p_x(\tau_0, y) + S_0 p_x(\tau_1, y)], \quad (5.3)$$

$$\varepsilon_z(x, y, h, \mu) = \frac{1}{h}V_z \approx v_{1z}(x, y, \mu) \approx p_{1z}(x, y) + \frac{1}{\mu}Q_0 p_z(\tau, \zeta); \quad (5.4)$$

and the plasma space charge

$$\rho_e(x, y, h, \mu) = N_p(x, y, h, \mu) - N_e(x, y, h, \mu) = h\nabla\mathbf{V}(x, y, h, \mu) \approx \frac{h^2}{\mu}\nabla\mathbf{P} \approx h^2\nabla\left[\mathbf{p}_1 + \frac{1}{\mu}(T_0\mathbf{p} + Q_0\mathbf{p} + S_0\mathbf{p})\right]. \quad (5.5)$$

The discharge is assumed to take place in a long tube with $R = 1$ cm filled with neon at pressure $p = 2$ Torr and $T = 300$ K. For the remaining parameters in expressions (5.1)–(5.5), we use the following numerical values: $E_0 = 3$ V cm^{-1}, $\nu = 10^{-2}$, $\sigma = 10^{-3}$, $\mu = 0.3$, $g_0 = 1.0$, $\xi_0^c = 0.114$, $\varphi_{1c} = 1.151$, and $\psi_c = 9.94$. These values correspond to the case where the mean ion concentration is 10^{12} cm^{-3} and the mean temperature is $T_e \approx 2 \times 10^4$ K. Since there is no reliable information about the step ionization frequency, we took a value for ξ_0^C at which the burst height is $\psi_C \approx 10$.

Let the burst curve C be a rectilinear segment, $0 < x < 1$, $y =$ const with the period $l_{st} = 6$ cm. As $c_p^{(0)}$, we choose the ion concentration on the axis of the homogeneous part of a nonstratified positive glow-discharge column with a direct current corresponding to the lower boundary of the appearance of strata. An analysis of the results of probe measurements presented in [12, 13] indicates that for this choice of $c_p^{(0)}$, $q_0(y) < 1$ between the strata.

Figure 1 shows the plots of functions (5.1) and (5.5) for $x = 0$ that define the distributions of the ion concentration,

$$N_p(0, y, h, \mu) \approx A_0(0, y, \mu) = \bar{A}_0(\rho, \zeta, \mu) \approx q_0(y) + Q_0\alpha_0(\tau, \zeta),$$

and a quantity proportional to the plasma space charge density,

$$\nu\left(\frac{\mu}{h}\right)^2\rho_c(0, y, h, \mu) = \nu\left(\frac{\mu}{h}\right)^2\bar{\rho}_e(\rho, \zeta, \mu) \approx \nu\mu\nabla p[T_0\mathbf{p} + Q_0\mathbf{p}],$$

along the discharge axis.

These expressions differ from distributions (5.1) and (5.5) in that they do not contain the boundary functions $S_0 w$, since the latter differ markedly from zero only near the $x = 1$ boundary. The quantity $\nabla\mathbf{p}_1$ was not included in the second expression either, because it is

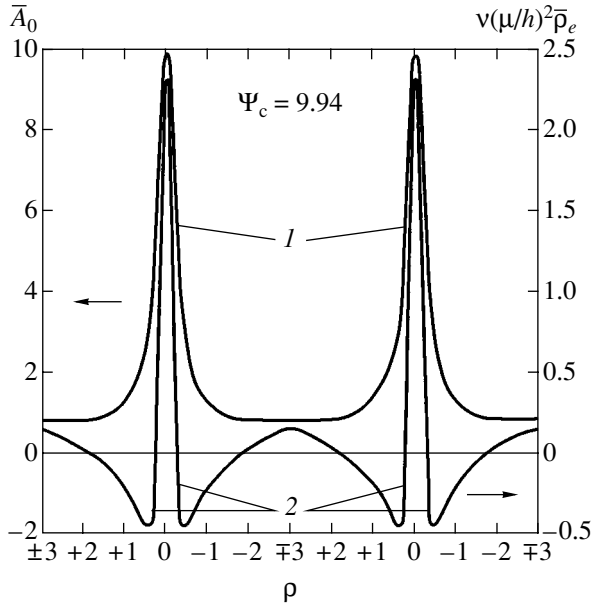

Fig. 1. Plots of the functions $A_0(0, y, \mu) = \bar{A}_0(\rho, \zeta, \mu) \approx q_0(y) + Q_0\alpha_0(\tau, \zeta)$, $y = \zeta - \rho$, $\rho = \mu\tau$ (curve *1*) and $\nu(\mu/h)^2\rho_e(x, y, \mu) = \nu(\mu/h)^2\bar{\rho}_e(\rho, \zeta, \mu) \approx \nu\mu\nabla[T_0\mathbf{p} + Q_0\mathbf{p}]$ (curve *2*) that define the distributions of the concentration of the charged plasma particles and a quantity proportional to the space charge density along the gas-discharge axis ($x = 0$).

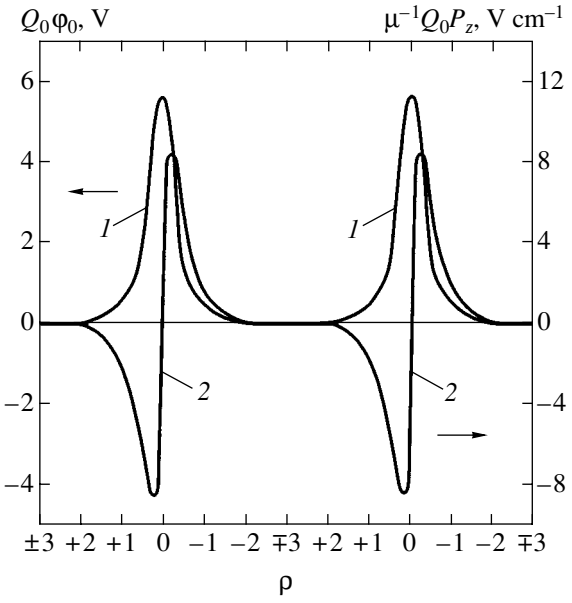

Fig. 2. Plots of the functions $(1/\mu)Q_0p_z(\tau, \zeta)$ and $Q_0\varphi_0(\tau, \zeta)$ that define the distributions of the electric field (curve *2*) produced by the plasma space charge with density ρ_e (see Fig. 1, curve *2*) and the potential of this field (curve *1*) along the discharge axis ($x = 0$).

small. In addition, the expression for N_p includes the boundary condition (3.2).

We see from Fig. 1 that the N_p distribution is typical of the strata. In contrast, the ρ_e distribution cannot be compared with anything, because this quantity is usually not measured by experimenters, but its reliability can be established in an indirect way, namely, from the strength of the field produced by this charge. Here, we note only one feature of this distribution: $\rho_e > 0$ between the strata on the axis.

Figure 2 shows the plots of functions (5.2) and (5.4) without the constant longitudinal electric field E_0, while Figs. 3 and 4 show the plots of the same functions with this field. Our calculations were based on formulas (4.6), (4.3), and (3.9). The potential distribution for the field produced only by the plasma space charge along the axis ($x = 0$) agrees closely with the results of experimental measurements from [12]. This suggests that the ρ_e distribution that follows from the plot in Fig. 1 corresponds to the actual distribution of the plasma space charge along the axis of the stratified discharge.

It follows from the shape of the curves in Figs. 3 and 4 that there is a potential well in front of each stratum on the side of the anode. Its depth decreases with decreasing burst height ψ_c and with increasing E_0. Note that no such potential wells have been detected experimentally in moving strata. This may be because an overestimated ψ_c and an underestimated E_0 were taken in our numerical calculations. For example, it follows from the plots in [12, 13] that the plasma density in a stratum exceeds its mean value by only a few factors, while we used $\psi_c \approx 10$ in our numerical calculations. In addition, it follows from the theory that the depth of the

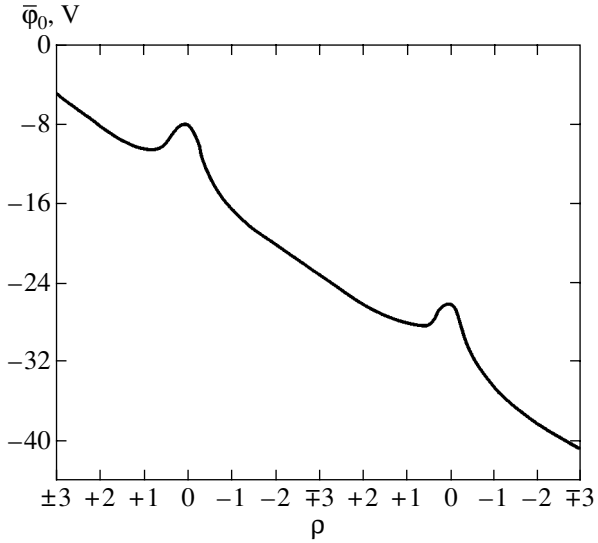

Fig. 3. Plot of the function $\varphi_0(0, y, \mu) = \bar{\varphi}_0(\rho, \zeta, \mu) \approx \varphi_{00}(0, y) + Q_0\varphi_0(\tau, \zeta)$ that defines the distribution of the potential of the total electric field ε_z (see Fig. 4) along the discharge axis.

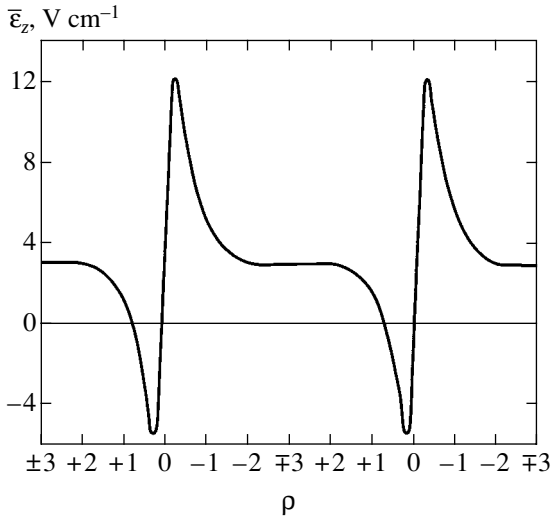

Fig. 4. Plot of the function $\varepsilon_z(0, y, \mu) = \bar{\varepsilon}_z(\rho, \zeta, \mu) \approx p_{1z}(0, y) + (1/\mu)Q_0 p_z(\tau, \zeta)$, $y = \zeta - \rho$, $\rho = \mu\tau$, that defines the distribution of the total electric field whose constant component is $E_0 = 3$ V cm^{-1} along the discharge axis.

potential well also depends on the derivative $\xi'_{0y}(0, y)$. Moreover, if this quantity is positive and fairly large, then potential wells can also appear in front of each stratum on the side of the cathode. Here, we did not consider this case in detail, because there is no comprehensive information on the function $\xi(x, y, \mu)$.

Figures 5 and 6 show the plots that define the distributions of the same physical parameters (5.1)–(5.3) and (5.5), but along the radius of the gas discharge. Our calculations were performed for $\xi_0(x, y_0) \equiv 0$, where y_0 is the ordinate of any point on the discharge axis that is located at equal distances from the two neighboring strata, $q_0(y_0) = 0.85$, and the plasma potential on the inner surface of the tube wall is higher than the potential on the discharge axis. These assumptions are based on the analysis of the experimental data from [12, 13] for moving strata. Let us set a value for the wall potential at which $S_0\varphi_0(0, y_0) = 5$ V (at zero potential on the axis). The parameter $b_0(y_0) = 6.9$ can then be determined from (4.8). We used formula (3.5) to calculate the function $S_0\alpha_0(\tau_1, y)$.

Finally, we used the integral equation (3.3) to determine the function $T_0\alpha_0(\tau_0, y)$. Its solution was constructed by the method of successive approximations. As a first approximation, we took the function

$$Z_1(t, y) = D(y)t^2\exp(-t^2),$$

where, in general, the unknown coefficient D can depend on y. Thus, the solution of Eq. (3.3) depends on two parameters, $m(y)$ and $D(y)$; there are two conditions (see Section 3.1) to calculate them. If we restrict ourselves to the second approximation when constructing the solution of Eq. (3.3), then we will obtain $m(y_0) = 6.2$ and $D(y_0) = 6.4$. This solution was used in our numerical calculations, which underlie the plots in the last two figures. Despite the rough approximation, the plots in Figs. 5 and 6 are consistent with the experimental data.

Thus, our analysis leads us to the following main conclusions:

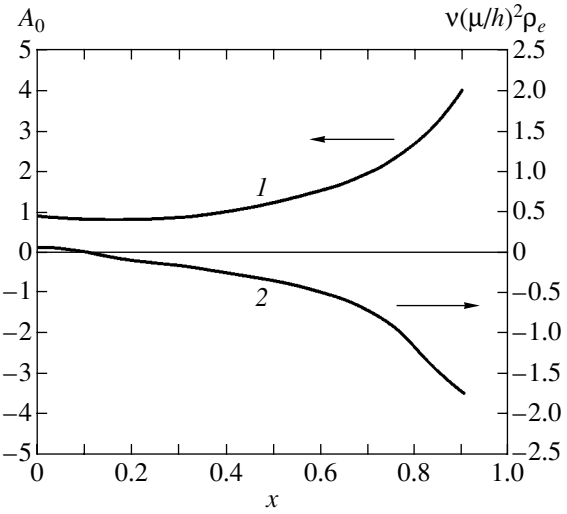

Fig. 5. Plots of the functions $A_0(x, y_0, \mu)$ and $\nu(\mu/h)^2\rho_e(x, y_0, \mu)$ that define the radial distributions of the concentration of the charged plasma components and a quantity proportional to the plasma space charge with density ρ_e (y_0 is the ordinate of any point on the discharge axis that is located at equal distances from the centers of the two neighboring strata).

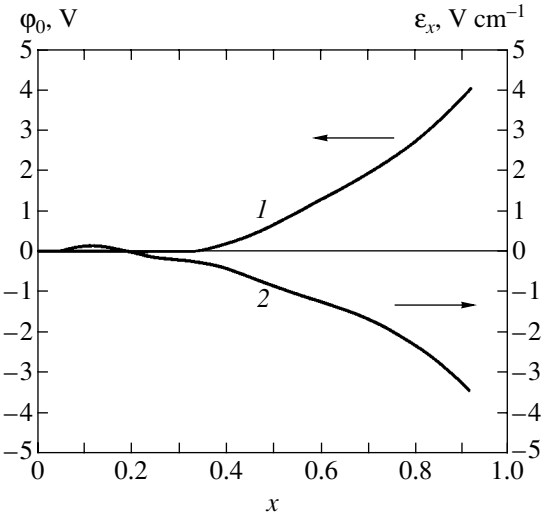

Fig. 6. Plots of the functions $\varepsilon_x(x, y_0, \mu)$ and $\varphi_0(x, y_0, \mu)$ that define the distributions of the electric field and its potential along the tube radius (y_0 has the same meaning as in Fig. 5).

(1) The spatially periodic inner transition layers emerging in the plasma of a positive glow-discharge column are essentially two-dimensional. Not only the longitudinal, but also the radial distributions of all physical plasma parameters depend on the ordinate of the point and differ greatly from the corresponding distributions in a homogeneous column.

(2) The salient features of the structure of the object under study can be established by using only the principal terms of the asymptotics of the solution of the initial mathematical model found above. A comparison of the above dependences with the published probe measurements of the structure of ionization waves reveals a satisfactory (not only qualitative) similarity.

In conclusion, note that the approximate solution used above can be improved, for example, by calculating the succeeding terms of asymptotics (2.8) or through numerical calculations. In the latter case, it is recommended that the approximate solution be used as the initial condition for the problem and then improved numerically.

REFERENCES

1. A. V. Nedospasov, Usp. Fiz. Nauk **94**, 439 (1968) [Sov. Phys. Usp. **11**, 174 (1968)].
2. L. Pekarek, Usp. Fiz. Nauk **94**, 463 (1968) [Sov. Phys. Usp. **11**, 188 (1968)].
3. A. Ya. Voronov, Teplofiz. Vys. Temp. **15**, 28 (1977).
4. A. Ya. Voronov, Available from VINITI, No. 877-79 (1979).
5. L. M. Volkova, A. M. Devyatov, and M. K. Tarakdzhi, Vestn. Mosk. Univ., Fiz., Astron. **35**, 63 (1994).
6. A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Usp. Fiz. Nauk **167**, 1215 (1997) [Phys. Usp. **40**, 1163 (1997)].
7. V. V. Balabanov, L. M. Vasilyak, S. P. Vetchinin, *et al.*, Zh. Éksp. Teor. Fiz. **119**, 99 (2001) [JETP **92**, 86 (2001)].
8. A. Ya. Voronov, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza **3**, 185 (1998).
9. A. B. Vasil'eva and V. F. Butuzov, *Asymptotic Expansions of Solutions of Singular Perturbed Equations* (Nauka, Moscow, 1973) [in Russian].
10. V. F. Butuzov and N. N. Nefedov, Dokl. Akad. Nauk **351**, 731 (1996).
11. A. Ya. Voronov, Izv. Ross. Akad. Nauk, Mekh. Zhidk. Gaza **6**, 105 (1999).
12. K. F. Bessonova, O. N. Oreshak, E. P. Ostapchenko, and V. A. Stepanov, Zh. Tekh. Fiz. **41**, 979 (1971) [Sov. Phys. Tech. Phys. **16**, 770 (1971)].
13. Yu. B. Golubovskiĭ and S. U. Nisimov, Zh. Tekh. Fiz. **65** (1), 46 (1995) [Tech. Phys. **40**, 24 (1995)].

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 156–159.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 179–182.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Mechanism of Convective Instability of a Binary Mixture

L. Kh. Ingel

NPO Tařfun, Obninsk, Kaluga oblast, 249038 Russia
e-mail: lingel@obninsk.com

Received January 31, 2005

Abstract—Even a very slow background vertical motion can strongly affect convection in a stratified fluid. A previously unknown mechanism of convective instability is demonstrated for a binary mixture in the flow field of such motion. © *2005 Pleiades Publishing, Inc.*

It is well known that binary mixtures exhibit some fluid-dynamic and thermodynamic characteristics that are peculiar at first glance (e.g., see [1, 2]). In particular, convective instability can develop, contrary to intuitive expectations, when the background stratification is stable, i.e., when the mixture density decreases with increasing height [1, 3]. Generally, the key role in this phenomenon (called *differential diffusion* or *double-diffusive convection* [1]) is played by the difference between the thermal and species diffusivities. It is believed that double-diffusive convection arises in a stably stratified medium when a slowly diffusing species (e.g., sea salt) has a destabilizing effect on density stratification. Even if this effect is relatively weak, it can be sufficient to destabilize a system characterized by stable density stratification. This paper draws attention to qualitatively different situations. In particular, instability can develop when slow species diffusion stabilizes density stratification. Note that this effect may be substantially stronger than thermal instability. At first glance, this possibility is even more unlikely, because slow species diffusion is generally much more conducive to convective instability than fast heat transfer in a double-diffusive system (transport coefficients are contained in the denominators of the respective Rayleigh numbers).

The analysis that follows is focused on systems with slow background flow parallel to the gravity vector. Convection in vertical background flows that are slow as compared to the convective motion under study is important, in particular, for geophysical applications (e.g., see [4, 5]). Atmospheric and oceanic convection frequently develops against the background of processes of much larger horizontal extent (such as cyclones or anticyclones). These processes involve mean vertical flows whose velocities are several orders of magnitude slower than those associated with convective instability. According to field experiments, even a slow downward flow of the ambient fluid effectively suppresses convection. The nature of this behavior is poorly understood.

Consider the following modified problem of Rayleigh–Bénard instability for a fluid layer between two horizontal plates [1, 6]. Suppose that the background flow is a slow downward motion. For simplicity, the velocity of this motion, $-W < 0$, is assumed to be independent of the vertical coordinate z measured upwards from the lower boundary $z = 0$.[1]

First, consider the case of a one-component medium whose density depends only on temperature T (effects due to species stratification are assumed to be negligible). The temperatures at the lower and upper boundaries, T_d and T_u, are supposed as known. Their difference, $T_d - T_u$, is denoted by ΔT. Heat transfer in the background flow is described by the equation

$$-W\frac{dT}{dz} = \kappa\frac{d^2T}{dz^2}, \qquad (1)$$

where κ is thermal diffusivity. The solution subject to the boundary conditions specified above can be written as

$$\Theta(z) = \frac{\exp(-\xi) - \exp(-w)}{1 - \exp(-w)}, \qquad (2)$$

where $\Theta(z) = [T(z) - T_u]/\Delta T$ is the dimensionless temperature deviation, $\xi = z/h$ is the dimensionless vertical coordinate, and $h = \kappa/W$ is the reference height associated with vertical motion (infinity in quiescent fluid). The key dimensionless parameter is $w = H/h = W(H/\kappa)$, where H is the fluid layer thickness. In the absence of background vertical motion (in the limit of $W \longrightarrow 0$, $h \longrightarrow \infty$, and $w \longrightarrow 0$), the result is the expected linear profile $\Theta = 1 - z/H$, i.e., the solution whose stability is analyzed in the classical Rayleigh–Bénard problem.

[1] The assumption of constant vertical velocity is incompatible with impermeability conditions for the background flow at $z = 0$ and $z = H$. Regimes of this kind can be implemented in practice by pumping the fluid in the vertical direction through porous horizontal boundaries. While this assumption is not necessary for further analysis, it substantially simplifies calculations.

1063-7761/05/10101-0156$26.00

Figure 1 shows the vertical profiles $\Theta(z)$ obtained for $w = 0$ and 10. It is clear that the downward flow "squeezes" the region of vertical temperature difference ΔT into a layer of thickness on the order of $h = \kappa/W$ at the lower boundary. Rigorous stability analysis of the steady state with a nonlinear temperature profile in a background downward flow is a difficult and cumbersome problem. However, a very plausible estimate can be obtained by using simple physical arguments. It is quite obvious that the stability of state *2* in Fig. 1 depends on the convective instability of a bottom "sublayer" of thickness on the order of $h = \kappa/W$ where almost the entire vertical temperature difference ΔT is localized. The effective Rayleigh number corresponding to the sublayer thickness is

$$\mathrm{Ra} \sim \frac{\alpha g \Delta T h^3}{\kappa \nu} \sim \frac{\alpha g \Delta T \kappa^2}{\nu W^3}, \qquad (3)$$

where α is the thermal expansion coefficient of the fluid, ν is the kinematic viscosity, and g is the gravitational acceleration.

According to expression (3), the effective Rayleigh number strongly depends on background vertical motion, which determines the thickness of the sublayer where almost the entire vertical temperature difference is localized. Note also that the effective Rayleigh number defined by (3) rapidly increases with κ, whereas the classical Rayleigh number decreases with increasing κ. Denote by Ra_{cr} the value of the effective Rayleigh number corresponding to loss of stability. Then, the downward velocity required to prevent the onset of convective instability is

$$W_{cr} \sim \left(\frac{\alpha g \Delta T \kappa^2}{\nu \mathrm{Ra}_{cr}}\right)^{1/3}. \qquad (4)$$

For example, setting $\kappa = \nu = 1\ \mathrm{m}^2/\mathrm{s}$ (effective turbulent transport coefficients characteristic of atmospheric boundary layers), $\alpha = 4 \times 10^{-3}\ \mathrm{K}^{-1}$, $\Delta T = 0.1$ K, and $\mathrm{Ra}_{cr} = 10^3$, one obtains $W_{cr} \sim 10^{-2}$ m/s. This velocity is two or three orders of magnitude lower than the characteristic velocity of atmospheric convection. However, results obtained in both field experiments and numerical simulations demonstrate that downward flow with this velocity suppresses convection.

Now, consider a two-component medium. For example, it can be saline water with density stratification determined by the vertical distributions of temperature and salinity, $T(z)$ and $s(z)$. In a commonly used approximation [1, 7, 8], it is assumed that density ρ is a linear function of temperature and salinity:

$$\rho = \rho_0[1 - \alpha(T - T_0) + \beta(s - s_0)]. \qquad (5)$$

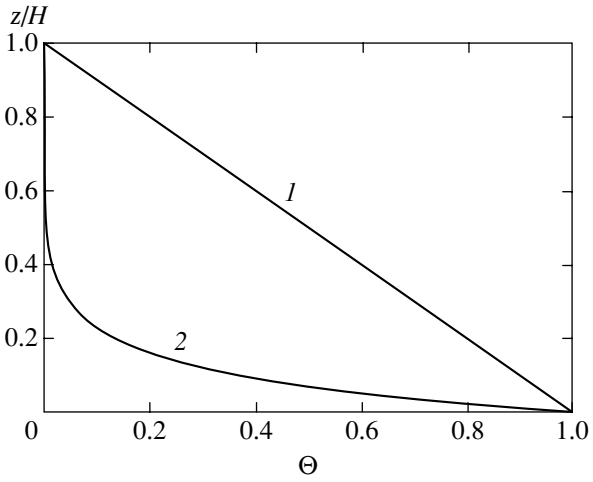

Fig. 1. Distortion of a vertical background temperature distribution by downward motion: (*1*) zero background vertical velocity; (*2*) $w = H/h = 10$.

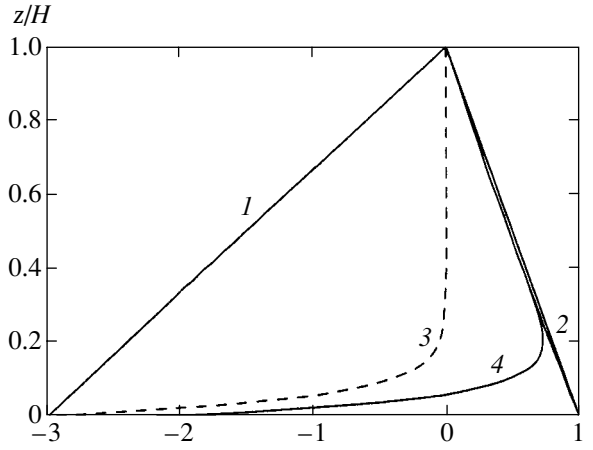

Fig. 2. Instability of density stratification due to downward motion in a two-component medium ($w = 0.2$, $w_s = 20$, $\Delta T > 0$, $\beta\Delta s/\alpha\Delta T = 3$): contributions of salinity, $(-\beta\Delta s\sigma(z)/\alpha\Delta T)$ (*1*), and temperature, $\Theta(z)$ (*2*), to the normalized buoyancy in the absence of background vertical motion; (*3*) profile *1* distorted by downward motion; (*4*) total buoyancy, $b(z)/\alpha\Delta T$, in the presence of downward motion.

The positive coefficient β has an obvious interpretation. (In oceanology, it is called the coefficient of salinity compression.) The subscript 0 denotes constant reference values of the variables (e.g., those at the upper boundary of the fluid layer). The steady-state salinity profile $s(z)$ is described by an equation similar to (1):

$$-W\frac{ds}{dz} = \chi\frac{d^2s}{dz^2}, \qquad (6)$$

where χ is the corresponding transport coefficient. To be specific, assume that the boundary conditions for salinity are analogous to those set above: $s = s_d$ and s_u at $z = 0$ and H, respectively. Then, the solution for s is similar to (2):

$$\sigma(z) = \frac{\exp(-\xi_s) - \exp(-w_s)}{1 - \exp(-w_s)}, \qquad (7)$$

where

$$\sigma = (s - s_u)/\Delta s, \quad \Delta s = s_d - s_u, \quad \xi_s = z/h_s,$$

$$h_s = \chi/W, \quad w_s = H/h_s = W(H/\chi).$$

When χ and κ have different values, the dimensionless functions given by (2) and (7) are quantitatively different. For example, $\chi/\kappa \approx 10^{-2}$ for saline water. The corresponding reference height h_s is smaller than h by one hundred times. Comparing the salinity profile with curve *2* in Fig. 1, one finds that the entire drop in salinity is localized in a layer of thickness much smaller than the thermal sublayer thickness. It can be shown that this disparity can lead to nontrivial convection regimes in the fluid layer.

Consider the dimensionless background buoyancy profile $b(z)$:

$$b(z) = -\frac{\Delta\rho(z)}{\rho_0} = \alpha\Delta T\Theta(z) - \beta\Delta s\sigma(z). \qquad (8)$$

Depending on the signs of ΔT and Δs, four qualitatively different regimes can be distinguished.

1. When $\Delta T < 0$, $\Delta s > 0$ (both fields are stably stratified), it is obvious that downward background flow has only a quantitative effect on the stable density stratification: the gradients of both fields are localized near the lower boundary. The sign of expression (8) remains negative everywhere. Thus, downward flow does not affect convective stability, which is unquestionable in this case.

2. In the absence of vertical motion, density stratification is unstable when at least one scalar field is unstably stratified to a sufficiently high degree. As shown above for a one-component medium, the effective Rayleigh number is reduced in the presence of a downward background flow (because of the reduced thickness of the layer where almost the entire unstable density stratification is localized). Therefore, the onset of convective instability can be prevented. Note that the instability due to unstable stratification of a slowly diffusing species (salt) is suppressed more easily, because the corresponding transport coefficient is contained in the numerators of criteria (3) and (4). Even a slow vertical motion drives the slowly diffusing species into a narrow layer at the boundary, reducing the effective Rayleigh number.

3. When $\beta\Delta s > \alpha\Delta T > 0$ (effects due to unstable temperature stratification are by far counterbalanced by stable species stratification in the absence of vertical motion), even a slow downward motion can change the sign of expression (8) in a considerable part of the layer; i.e., a system with "average" density profile characterized by an ample stability margin can be destabilized. This is clear from Fig. 2, where vertical profiles of buoyancy (8) normalized to $\alpha\Delta T$ and its "thermal" and "saline" components are shown. In the absence of vertical motion, the dominant role is played by the salinity stratification characterized by an ample stability margin (line *1*). Even though the temperature stratification is unstable, its contribution to buoyancy is relatively small (line *2*). Slow downward motion (with $w = 0.2$) only slightly changes the temperature profile (represented by a curve very close to line *2*), whereas the salinity profile is drastically distorted (curve *3*). Since salt is carried downwards by vertical motion, the salinity stratification cannot counterbalance the effects due to unstable temperature stratification, and the resulting buoyancy profile becomes unstable in the greater part of the layer (curve *4*). Note that faster downward background motion can distort the temperature profile as well, stabilizing the system. Thus, the domain of instability corresponds to an intermediate range of downward flow velocity.

4. When $\alpha\Delta T < \beta\Delta s < 0$, the species stratification is unstable, but the system is stabilized by the stable temperature stratification in the absence of vertical motion.[2] It is clear that even a slow downward motion may trigger convective instability. Indeed, it follows from the discussion above that the vertical gradient of species concentration grows steeper in the presence of downward motion; i.e., the instability of the species stratification increases. This effect can be strong when $w_s > 1$. At the same time, the value of w can be much less than unity, in which case the temperature stratification remains almost constant and cannot counterbalance increasingly unstable species stratification.

Thus, whereas vertical background motion can prevent the onset of convective instability in certain cases, it can destabilize a double-diffusive layer. It is important that the phenomena described above can take place when the vertical velocity is very low. For example, in cases 3 and 4 discussed above, an essential condition for instability of the layer is

$$w_s = WH/\chi \geq 1, \quad W \geq \chi/H.$$

For seawater, $\chi \approx 1.5 \times 10^{-9}$ m^2/s, and destabilization can be caused even by a very slow vertical motion with $W \sim 10^{-8}$ m/s. This phenomenon can be interpreted as

[2] The stability condition for this double-diffusive system is more complicated than in the classical Rayleigh–Bénard problem [1, 8, 9]. In particular, it can be much more restrictive. It is assumed here that the condition is satisfied in the absence of vertical motion.

instability of a new type. Note, however, that the times required for a system to evolve into the unstable steady states considered above may be very large.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 04-05-64027 and by ISTC, grant no. G-553.

REFERENCES

1. J. S. Terner, *Buoyancy Effects in Fluids* (Cambridge Univ. Press, Cambridge, 1973; Mir, Moscow, 1977).
2. L. Kh. Ingel', Usp. Fiz. Nauk **172**, 691 (2002) [Phys. Usp. **45**, 637 (2002)].
3. L. Kh. Ingel', Zh. Éksp. Teor. Fiz. **122**, 1019 (2002) [JETP **95**, 878 (2002)].
4. R. C. J. Sommerville and T. G. Chen, J. Atmos. Sci. **36**, 805 (1979).
5. T. B. Glukhovskaya and A. E. Ordanovich, Meteorol. Gidrol., No. 3, 31 (1984).
6. A. V. Getling, *Rayleigh–Bénard Convection: Structure and Dynamics* (Editorial URSS, Moscow, 1999) [in Russian].
7. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics*, 3rd ed. (Nauka, Moscow, 1986; Pergamon, New York, 1987).
8. G. Z. Gershuni and E. M. Zhukhovitskiĭ, *Convective Stability of Incompressible Fluid* (Nauka, Moscow, 1972) [in Russian].
9. G. Walin, Tellus **16**, 389 (1964).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 160–168.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 183–192.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Shock-Wave and Plasma-Pinch Mechanisms of Galactic Cosmic-Ray Production

B. A. Trubnikov

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

e-mail: batrub@nfi.kiae.ru

Received February 2, 2005

Abstract—Based on recent discoveries, we show that it is appropriate to complement the standard shock-wave model for the production of galactic cosmic rays by a plasma-pinch model. The latter describes well the production of high-energy cosmic rays, yields a simple formula for their intensity, and allows the threshold pattern of the knee-type kink in the secondary particle spectrum and a number of unusual phenomena observed above the threshold to be explained. *© 2005 Pleiades Publishing, Inc.*

1. THE SHOCK-WAVE MODEL

The Sun is known to be the source of solar cosmic rays with energies $E < 10^8$ eV (Fig. 1). Cosmic rays with energies $E > 10^8$ eV are commonly called galactic cosmic rays (GCRs), but their acceleration mechanism has not yet been firmly established. In his well-known paper [1], Ginzburg summarized the results of a prolonged study of this problem that he formulated in three conclusions:

—the galactic rather than metagalactic model (theory) is valid;

—the galactic model with a large halo (without a disk) is valid;

—supernova explosions are the main sources of cosmic rays in the Galaxy.

However, new facts that require complementing the shock-wave model have been discovered in recent years.

A supernova explosion is accompanied by a shock wave whose remnant is an expanding (Crab-like) plasma nebula with a pulsar at its center. A rapidly rotating pulsar generates shock waves and turbulent plasma motions in the nebula on which the particle acceleration is possible in accordance with the equation [2]

$$\frac{df(t,\mathbf{p})}{dt} = \mathrm{div}_{\mathbf{p}}(\hat{D}_{\mathbf{p}}\mathrm{grad}_p f). \qquad (1)$$

This equation describes the particle momentum (**p**) dif-

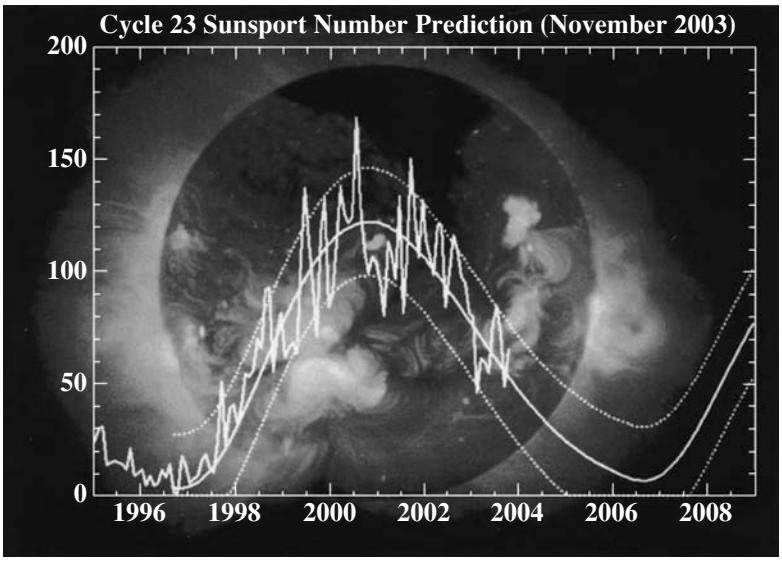

Fig. 1. Giant solar flare of November 28, 2003.

1063-7761/05/10101-0160$26.00

fusion with the diffusion tensor $\hat{D}_p$ proportional to the plasma pulsation intensity; f is the particle distribution function.

Since $\hat{D}_p > 0$, even initially cold particles gradually acquire large momenta and are then caught up by the shock waves. Subsequently, they diffuse through the Galactic disk and halo in accordance with the ordinary diffusion equation

$$\mathrm{div}_{\mathbf{r}}[\hat{D}_{\mathbf{r}}\mathrm{grad}_{\mathbf{r}}n(\mathbf{r})] = Q(\mathbf{r}), \qquad (2)$$

where $n(\mathbf{r})$ is the cosmic-ray density and $Q(\mathbf{r})$ is the distribution of sources over the Galactic disk. Figure 2 elucidates this shock-wave model.

However, it was qualified in [1] that relatively low-intensity cosmic rays with energies above 10^{15}–10^{16} eV were not considered. Therefore, in this paper, we study the range of ultrahigh energies in which many important discoveries have been made since about 2000.

2. RECENT DISCOVERIES

Below, we provide information published on the Internet [3] and in print abstracts [4–7]. Thus, for example, it is reported in [3] that five protons with an energy of 10^{18} eV that arrived from the same source, a cluster of colliding galaxies 450 million light years away (!), have been detected in extensive air showers (EASs). The proton paths are rectilinear and are not distorted by intervening magnetic fields. Since their velocity is $v = c(1 - 10^{-22})$, where c is the speed of light in a vacuum, they produce long rectilinear Cherenkov tracks in air that specify the source's direction with a high accuracy.

The paper [4] reported the discovery of EASs with energies $E > 10^{20}$ eV far above the Greisen–Zatsepin–Kuzmin threshold (against the cosmic microwave background). The specific mechanism of their acceleration remains a mystery. The distribution of arrival directions is consistent with isotropic distribution. EAS doublets and triplets from the same direction are observed. Protons are the dominant particles of the primary spectrum in the range of ultrahigh energies. The discovery of entirely new phenomena is possible.

It is noted in [5] that the following five difficult-to-explain events are observed in EASs even at $E = 10^{16}$ eV:

anticentaurs—electromagnetic cascades of gamma-ray photons and $e^{\pm}$ pairs;

centaurs—cascades composed mostly of hadrons;

alignment of cascade tracks along one straight line (on Earth);

halo events—cascades with a peculiar pattern;

penetrating cascades from unknown particles.

No such events are observed before the kink in the spectrum (see Fig. 5 below). It is also noted in [5] that

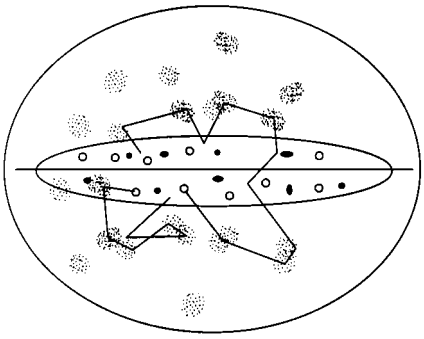

Fig. 2. Scheme of the shock-wave model: the GCR particles produced during supernova explosions in the Galactic disk are scattered by clouds of magnetized plasma in the halo, where they become isotropic, and then diffuse from the Galaxy.

the spectrum $dN/dE \propto E^{-\nu}$ with the index $\nu = 1 + \sqrt{3} = 2.73$ (N is the number of recorded events) suggested previously [8] is universal, while none of the other models yields an unambiguous spectral index.

The numerical calculations for the energy range 10^{14}–10^{19} eV [6] show that the observed GCR spectrum (to be more precise, calculated for the location of the Sun in the Galaxy) steepens significantly compared to their spectrum in the sources only in a relatively narrow (one to two orders of magnitude) range of primary energies. This makes it difficult to reconcile the observed spectrum with predictions of the standard model for the acceleration of cosmic rays by shock waves.

The paper [7] reported that ten successful flights of balloons with emulsion chambers at an altitude of 32 km (where the spectrum is primary) were carried out in 1995–1999. The proton and helium spectra were shown to have similar indices and to exhibit no appreciable steepening at energies as high as 10^{15} eV per particle.

The above list of facts allows us to formulate two main questions: (i) about the spectrum of the primary GCR particles and their maximum energy; and (ii) about the new physics of the cascades of secondary particles in EASs.

In the standard shock-wave model, the produced GCRs in the sources are assumed to have a power-law spectrum, $dN/dE \propto E^{-k}$, with an index near $k = 2$. A steeper spectrum with an index near $k = 2.7$ observed on Earth must form only during diffusive GCR propagation through the Galaxy, from sources in the disk to the solar neighborhood.

However, as is noted in [7], the primary spectrum observed on Earth shows no tendency to steepen up to $E = 10^{15}$ eV. It can be assumed that particles with energies above 10^{15} eV are unlikely to be deflected by any galactic magnetic fields and most likely fly toward Earth directly from their sources, which are distributed isotropically over the sky. However, when the particles

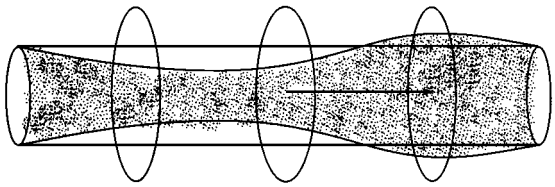

Fig. 3. Scheme of the plasma-pinch model.

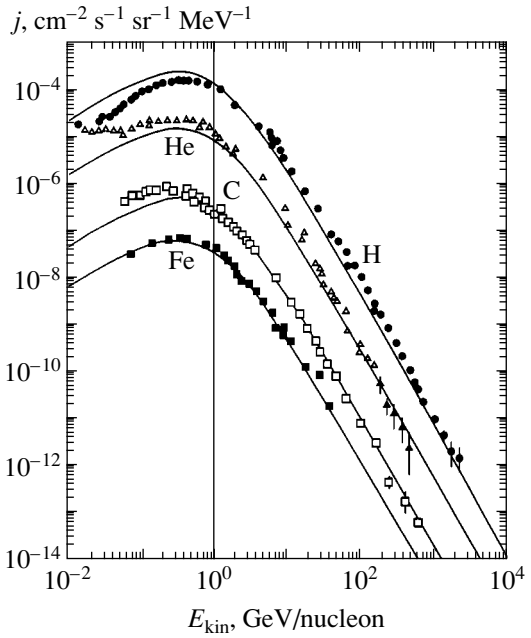

Fig. 4. Comparison of the primary observed GCR spectra (dots) with the curves constructed using formula (3).

move along straight trajectories, their spectrum on Earth must not differ from the source spectra for all energies above 10^{15} eV.

Meanwhile, the calculations [6] yield a primary spectrum with a kink in a narrow energy range, 10^{15}–10^{17} eV. As was noted by the authors of [6], this makes it difficult to compare the shock-wave model with observations.

In our opinion, the shock-wave model requires a refinement at high energies, and we should return to the well-known (since 1949) Fermi acceleration mechanism of a charged particle that is reflected from two approaching clouds of magnetized plasma. As we see from Fig. 2, such clouds in the halo are needed, and they manifest themselves in the shock-wave model as observations of galactic radio halos. It is assumed in the Fermi hypothesis that, if the separation L between the clouds decreases and the adiabatic invariant Lv = inv is conserved, then the velocity v increases.

However, there exists a much more efficient mechanism where a neutral current sheet is formed at the collision boundary between the two magnetized plasma clouds; this sheet soon breaks up into cylindrical current pinches with a longitudinal internal magnetic field. During their prior constriction, this field must straighten the pinches, and they can become very long and almost rectilinear.

Subsequently, necks must grow on them, as was first shown in 1952 [9].

3. THE PLASMA-PINCH MODEL

Figure 3 elucidates this model. The quasi-neutral plasma squeezed out of the necks into the thickenings accelerates along its way, and the particles (electrons and ions) can gain very high velocities. In several papers [10], we showed that the following formula for the GCR intensity is derived (see Appendix) in the relativistic case for an idealized model (composed of a set of isotropically directed pinches without longitudinal fields and with skinned currents):

$$j(\varepsilon) = \alpha_i \beta^2 \gamma^{-\nu}, \qquad (3)$$

where

$$\beta = \frac{v}{c}, \quad \gamma = \frac{1}{\sqrt{1-\beta^2}}, \quad \varepsilon = \gamma - 1, \quad \nu = 1 + \sqrt{3},$$

where $\gamma = 1 + \varepsilon/2$ for hydrogen. This formula

(1) contains no parameters (except the normalization factor α_i, which is different for different GCR nuclei);

(2) describes well the primary GCR spectrum (Fig. 4) even at energies $E < 10^{15}$ eV, at which the shock-wave model is assumed to be applicable, but this is probably a coincidence;

(3) predicts no limiting GCR energy (and could describe the entire observed energy range if there were no kinks in the primary spectrum shown in Fig. 5 for secondary EAS spectra).

4. PRIMARY AND SECONDARY SPECTRA

At $E < 10^{14}$ eV, the primary GCR particles can be recorded by relatively light detectors on satellites beyond Earth's atmosphere. However, no satellite measurements are possible at $E > 10^{14}$ eV, and all further information can be obtained only from the observations of secondary particles in the EASs generated during the collisions of a primary particle with air atoms. However, the interpretation of the EAS data is ambiguous.

Figure 5, in which kinks can be seen, was constructed by assuming that the EAS detectors catch all of the total energy of the secondary particles that is attributed to the primary particle. However, part of the energy in strong EASs can elude detection, for exam-

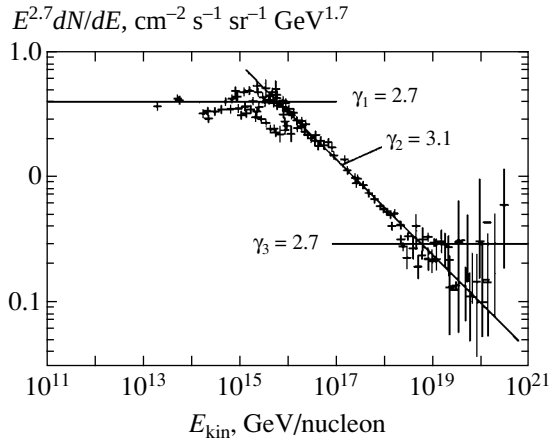

Fig. 5. Kinks in the secondary GCR spectrum [5].

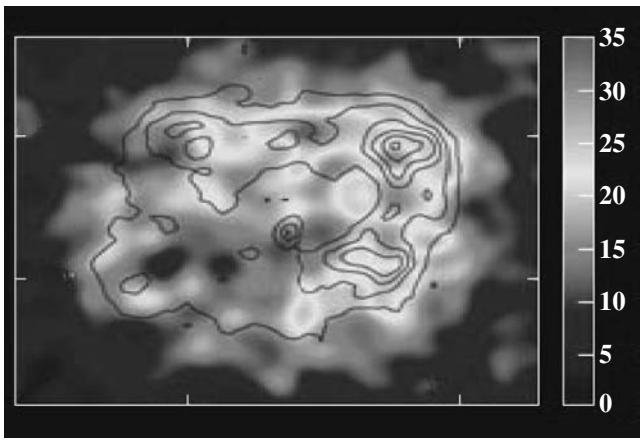

Fig. 6. The first gamma-ray image of the pulsar with an envelope [11]

ple, in the form of difficult-to-record energetic neutrinos and muons.

Some researchers believe that there are no kinks in the primary spectrum at all. In the standard shock-wave model, the first knee-type kink at $E_1 = 5$ PeV (1 PeV = 10^{15} eV) is explained as follows: at $E < E_1$, the GCR particles get entangled in nonuniform Galactic magnetic fields, while at $E > E_1$, their diffusive escape from our Galaxy begins.

Our calculations of the diffusive escape mechanism yield a smooth kink in the spectrum, while the observed kink has a sharp threshold pattern. Circumstantial evidence for the spectral indices of GCRs in supernova remnants could be the gamma-ray photons arriving from them to Earth that are not deflected by intervening magnetic fields. In recent years, several large gamma-ray telescopes have been put into operation, and the cumulative energy spectra of the gamma-ray photons from various sources observed by them have a power-law form:

$$N_\gamma(E > 0.8 \text{ TeV}) \propto E^{-k}.$$

We added unity to the indices of the gamma-ray spectrum from the source (k_γ) and the spectra of the events from the source (k_{on}) and those taken simultaneously with observation of the sources (k_{off}) given in [4–7, 10] for the cumulative spectra to obtain the differential spectra. The latter are presented in the table. The values of k_γ in the first column have a very large spread. The detection of gamma-ray photons with energies $E >$ 30 TeV would be evidence of their hadron rather than electron origin.

According to a recent report [11], the HESS system of four advanced gamma-ray telescopes began operation in Namibia in 2004; this system has imaged a supernova remnant (the pulsar J1713.7–3946; see Figs. 6 and 7) directly in gamma rays for the first time. Figure 7 shows two gamma-ray spectra: the upper spectrum for the entire envelope has a power-law form, $dN_\gamma/dE \propto E^{-k}$, with an index $k = 2.2$, while the lower spectrum (taken previously in [12] only for the upper right hottest part of the envelope) has an index near $k = 3$.

In our opinion, these data and those given in the table have such a large spread that they are consistent with the plasma-pinch model; the latter predicts an index $\nu = 1 + \sqrt{3} = 2.732$, which will possibly be detected during the observations of gamma-ray photons with higher energies, $E > 100$ TeV.

5. THE PROBLEM OF THE KINKS IN THE SPECTRUM

Thus, we can assume that the primary GCR spectrum has no kinks, but they are typical of the secondary EAS spectra. Schematically, an EAS resembles a fir tree with a central hadron stem and side branches of secondary particle cascades emanating from it downward.

Table

	k_γ	k_{on}	k_{off}
Crab Pulsar	2.45	2.61	2.73
Tycho Brahe Pulsar	2.00	2.12	2.73
Geminga Pulsar	1.65	1.88	2.77
Pulsar J1713.7–3946	2.2	–	–
Hole Cygnus 3-X	3.55	2.94	2.74
Quasar Mkn 421	2.87	2.53	2.76
Quasar Mkn 501	2.82	2.85	2.76
Quasar 1739 + 522	3.64	3.29	2.77

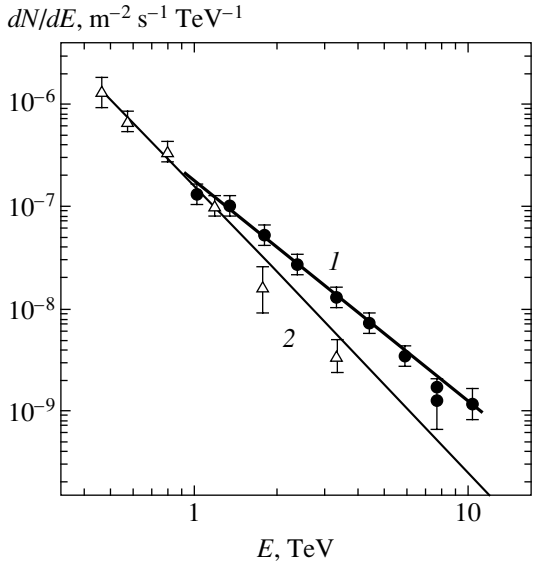

Fig. 7. Gamma-ray spectra: (1) for the entire envelope, $k_\gamma = 2.2$ [11]; (2) for the hottest part of the envelope (upper right in Fig. 6), $k_\gamma = 3$ [12].

The simplest explanation of the first kink in the GCR spectrum at $E_1 = 5$ PeV was offered in [13], where it was assumed that the lower end of the hadron stem in an intense EAS goes to the bottom and part of the total EAS energy simply cannot be recorded by existing instruments. This methodological error may manifest itself as an apparent kink in the secondary spectrum (an artifact) without any kink in the primary spectrum. If the assumption made in [12] is valid, then the maximum GCR energy detected to date, $E_{\max} = 2 \times 10^{20}$ eV, may actually prove to be even higher.

In the opinion of Petrukhin [5], the second cause of the kink could be the formation of new heavy particles or states of nuclear matter. Recall that the theory predicts energies of 10^{23}–10^{25} eV for the Grand Unification of three forces of Nature—strong, weak, and electromagnetic forces. However, there are also models of the early Grand Unification at energies $E = 10^{15}$–10^{17} eV, which are close to the energy of the first knee-type kink at $E_1 = 5 \times 10^{15}$ eV. The late Grand Unification may be responsible for the appearance of the second ankle-type kink observed in EASs at $E_2 = 5 \times 10^{18}$ eV.

6. COLLECTIVE QUARK SELF-MULTIPLICATION

In our opinion, however, a third cause is also possible, which allows one to explain more easily (without introducing new, as yet unknown particles) the threshold pattern of the kink or, more specifically, the switching-on of the processes of collective multiplication of quark pairs described by the equation

$$\frac{dn}{dt} = \frac{1}{2}n^2\langle\sigma v\rangle - \frac{n}{\tau}.$$

Here, n is the density of the quark pairs, σ is their multiplication cross section, and τ is their lifetime in the reaction volume.

The quark multiplication begins at $dn/dt > 0$, i.e., at a critical value of $n_{\rm cr} = 2/\langle\sigma v\rangle\tau$. For example, let us consider the situation where a primary proton with energy $E_0 = m_p c^2\gamma_0$ impinges on a resting nucleus, for instance, of nitrogen ^7N, which contains 14 nucleons, i.e., 42 quarks and 21 quark pairs.

The proton radius and volume are assumed to be $r_p = 1.2 \times 10^{-13}$ cm and $V_p = (4\pi/3)r_p^3$, respectively. Since the proton contains three quarks, the quark density in it is $n_0 = 3/V_p$, and it is the same in all nuclei at rest, including nitrogen. Therefore, the density of the quark pairs at rest is $n_0 = 3/2V_p$; if the medium moves at velocity $v = c\beta$, then this density is

$$n_{\rm eff} = n_0\gamma_{\rm eff}, \quad \gamma_{\rm eff} = 1/\sqrt{1-\beta^2}.$$

The above formula can then be rewritten as a threshold value:

$$\gamma_{\rm eff}^{\rm cr} = \frac{2}{n_0\langle\sigma v\rangle\tau} = \frac{16\pi}{9}\frac{M}{m_p}\frac{r_p^2}{\sigma}\frac{r_p}{v\tau}. \qquad (4)$$

Further, we assume that a primary ultrarelativistic proton with momentum $P = m_p c\gamma_0$ and a nitrogen nucleus at rest are combined into a single complex of mass M with the same momentum $P = m_p c\gamma_0 = Mc\gamma_{\rm eff}$ for a short time. In this case, the factor $\gamma_{\rm eff} = (m_p/M)\gamma_0$ is smaller than the initial factor γ_0. Finally, we have for the threshold

$$\gamma_0^{\rm cr} = \frac{16\pi}{9}\frac{M}{m_p}\frac{r_p^2}{\sigma}\frac{r_p}{v\tau}. \qquad (5)$$

It is from this formula that we wish to obtain the factor $\gamma_0 = 5 \times 10^6$ for the proton, which corresponds to the kink energy $E_1 = 5$ PeV $\approx 5 \times 10^6$ GeV $= m_p c^2\gamma_0$. However, this requires finding the factors on the right-hand side, and additional assumptions are needed here.

7. THE MASS OF THE "BALL" AND THE KINK THRESHOLD

To the best of my knowledge, only individual particle collisions have been investigated previously in quantum chromodynamics and no collective processes have been considered. However, they are phenomeno-

logically taken into account in the theory of quark–gluon plasma (the Fermi–Landau theory), in which it is assumed that the number of produced quarks during head-on collisions between two nuclei is

$$N_q = \sqrt{E_{\rm cis}/E_q},$$

where $E_{\rm cis}$ is the total energy in the center-of-inertia system, and $E_q = m_q c^2$ is the rest energy of the quark (the gluons are assumed to be massless). If a proton of initial energy $E_0 = m_p c^2 \gamma_0$ impinges on a resting nitrogen nucleus of mass $M = 14 m_p$, then

$$E_{\rm cis} = Mc^2\sqrt{1+\left(\frac{m_p}{M}\right)^2 + 2\left(\frac{m_p}{M}\right)\gamma_0} \approx 5.3 m_p c^2 \sqrt{\gamma_0}. \quad (6)$$

For a proton with $\gamma_0 = 5 \times 10^6$, we find $N_q \approx 200$ quarks and, having added 45 initial quarks to them, obtain about 125 quark pairs in the formed overall coalesced complex with the rest mass $M = 245 m_q$. We then have the factor $M/m_p \approx 82$, and the above formula for the kink threshold can be rewritten as

$$\gamma_0^{\rm cr} = 5 \times 10^6 \approx 456 \frac{r_p^2}{\sigma}\frac{r_p}{v\tau}.$$

Taking $v = c$, we have $r_p/c \approx 4 \times 10^{-24}$ s, but we assume that the lifetime of the quark pairs in the overall complex is appreciably longer, say, by a factor of 100. We then obtain the following estimate of the cross section:

$$\sigma \approx 10^{-6} r_p^2 \approx 10^{-32}\ {\rm cm}^2,$$

which seems plausible for the following reasons.

This cross section probably corresponds to such close quark collisions that the gluon forces are weak due to their so-called asymptotic freedom. The Coulomb forces with the Rutherford cross section $\sigma_R \approx (q^2/E_q)^2$, where $q = e/2$ is the mean electric charge of the quarks, play a major role. For our estimation, we assume that $m_q = m_p/3$, which yields the required cross section

$$\sigma_R \approx r_0^2 (3m_e/4m_p)^2 \approx 10^{-32}\ {\rm cm}^2. \quad (7)$$

Here, $r_0 = e^2/m_e c^2$ is the classical electron radius.

8. ON THE ALIGNMENT OF TRACKS AND NEW PHYSICS

Note that the beams of ultrahigh-energy GCR particles in the plasma-pinch model must have the pattern of individual "machine-gun bursts" that are isotropic in arrival directions. This may be the reason why close (in arrival time) doublets and triplets of showers are occasionally observed on the same EAS area from the same direction (i.e., as if from a single machine-gun burst).

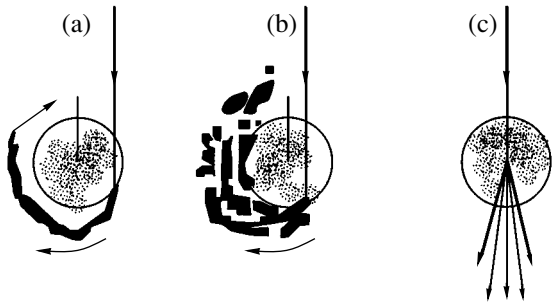

Fig. 8. (a) Peripheral, (b) middle, and (c) central collisions.

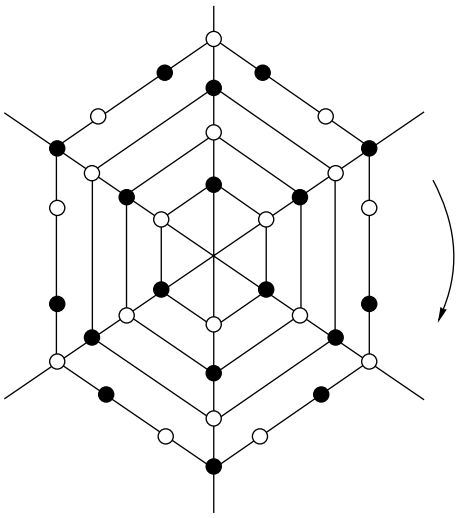

Fig. 9. The stretchable chain of quark–antiquark pairs winds around the core of the nucleus and forms a spider web.

Below, we consider three types of possible collisions (peripheral, middle, and central) between a primary proton and the nucleus of an air atom, which are schematically shown in Fig. 8.

The tracks of the secondary cascades are probably aligned not in a head-on collision, but in an off-center collision, in which a turning ball of secondary quark pairs arises, which is schematically shown in Fig. 9 and resembles a spider web.

Recall that different atoms in a molecule are bound by the "residual" electromagnetic forces, while different nucleons in a nucleus are pulled together by the "residual" gluon forces, despite the electrostatic repulsion of the protons. Therefore, the bunch of quark–antiquark pairs multiplied in the nucleus will continue to be attracted to the core by these residual gluon forces at the exit from the nucleus and to be wound around the

core. According to our hypothesis, this can give rise to a rotating spider web.

Since the proton comes from above, the spider web must rotate in the vertical plane. As the spider web breaks up, the radial jets of secondary particles (along the "spokes" of the spider web) probably yield the cascade tracks aligned along one straight line on Earth. Apart from the alignment of the tracks of the radial jets, we also see from this scheme that only two breakup scenarios are possible for the vertical spider web: into neighboring quark–antiquark doublets, which gives rise to electromagnetic anticentaurs composed of pions decaying into neutrinos, gamma-ray photons, and electron–positron pairs according to the scheme

$$\pi^+ \longrightarrow \pi^0 + e^+ + \nu_e, \quad \pi^0 = \frac{\bar{u}u - \bar{d}d}{\sqrt{2}} \longrightarrow \gamma + \gamma, \quad (8)$$

and into quark triplets, which gives rise to centaurs composed mostly of hadrons with high angular momenta L, i.e., "resonances" with large spins S. The relation $J = a + m_R^2 b$, where m_R is the resonance particle mass and a and b are constants, is known to be obtained for the quantized total angular momentum $J = S + L$. To bind the triplets (either only of quarks or antiquarks), it is probably necessary that the spider web in the second scenario be multilayered and the triplet members be taken from the neighboring layers.

Observations show that in a central head-on collision of a primary GCR proton with the nucleus of an atom in air, the multiplication process does not develop in the way predicted by the Fermi–Landau theory for the formation of an ellipsoid of quark–gluon plasma. More specifically, if the above quark self-multiplication threshold is exceeded, then a third scenario takes place—instead of the ellipsoid, a torus like a "smoke ring" with a vertical axial jet is formed, which produces a halo event—a cascade surrounded by a peculiar pattern, the tracks of the torus.

Finally, the penetrating cascades are probably attributable [5] to a long-range component, whose role can be played by weakly interacting ultrahigh-energy muons.

As we see, the proposed inclusion of the collective quark self-multiplication processes not only can quantitatively explain the threshold value of the knee-type kink in EASs at $E_1 = 5$ PeV, but can also qualitatively explain the cascades of unusual types that, according to our hypothesis, are associated precisely with the pattern of the rotating spider web constructed like parquet from the neighboring quark–antiquark pairs. This "parquet" or even the multilayered "quark crystal" may be amenable to a more accurate theoretical analysis, probably, using group or string theory.

9. DISCUSSION OF THE MODELS

As was noted above, the theory of the standard shock-wave model runs into difficulties, particularly near the kink and at high energies. At the same time, the plasma-pinch model with its simple formula for the intensity, $j = \alpha\beta^2\gamma^{-\nu}$, allows us to explain the following:

(1) the fall-off of the left edge of the spectrum due to the natural factor β^2;

(2) the pinch mechanism of high-energy GCR particle acceleration;

(3) the GCR spectrum without any restriction on the maximum energy;

(4) the threshold of the kink in the secondary EAS spectrum;

(5) the doublets and triplets of close (in time) showers in EAS;

(6) the alignment of the tracks of cascades from the vertical spider web of quarks;

(7) the emergence of centaurs—cascades composed mostly of hadrons;

(8) the formation of anticentaurs—mostly electromagnetic cascades;

(9) the halo events—cascades surrounded by a peculiar pattern.

Points (4)–(9) are related to our hypothesis about the collective self-multiplication of quarks, which is not considered in chromodynamics. This possibility could also be introduced in the shock-wave model, but it has no "machine-gun bursts," and events 6–9 are possible only above the threshold, starting from an energy of 10^{16} eV. In general, in our opinion, the plasma-pinch model complements well the standard shock-wave model, particularly at high GCR energies.

In the plasma-pinch model, both the plasma envelopes of supernova explosions in the Galactic disk and the colliding clouds of magnetized plasma in the Galactic halo are assumed to be the GCR sources. Therefore, the energy supply required for the two models is either the same or even greater in the plasma-pinch model, because GCRs can be produced not only in the disk, but also in the halo of the Galaxy. Undoubtedly, the main difficulty of the plasma-pinch model is that the pinches of the required type have not yet been observed in the Galactic halo. Whether they can be detected by currently available instruments of a particular type, even if they exist in the halo, remains an open question.

Colliding plasma clouds are seen, for example, in the Crab Nebula, where even the optical radiation from neighboring cells has different polarizations, suggesting that the fields magnetizing them have different directions. The plasma pinches of the type required for our model must be formed at the contact boundaries between these cells. Although plasma filaments are actually seen there, no individual pinches and "bolts of lightning" in the Crab can probably be detected in the optical range so far. The pinches and bolts of lightning

are even more difficult to find in the halos of galaxies. However, individual plasma clouds are observed in galactic halos in the radio range, and they are used in the shock-wave model as the presumed cause of the isotropization of the GCRs produced in the Galactic disk. However, if the clouds exist, then the currents magnetizing them also exist!

The acceleration mechanism itself rather than the GCR birthplace and the energy supply should be considered to be the main difference between the two models under consideration. The acceleration mechanism is attributed to shocks in the shock-wave model and to pinches, which periodically break and are again occasionally formed at the boundaries of the neighboring colliding plasma clouds that are magnetized differently in the neighboring cells, in the plasma-pinch model.

Above, we considered only the simplest idealized model composed of a set of isotropically directed pinches with skinned currents and without longitudinal internal magnetic fields, whose role deserves special study. On the one hand, these fields can contribute to the straightening and formation of the required long pinches during their prior constriction and formation. On the other, they complicate the neck breaking process. If the current is not skinned, but distributed over the pinch cross section, then there is a field component B_φ inside the pinch that defocuses the ions moving against the current and focuses the ions moving along the current. Clearly, only the latter will yield the pencil paraxial beams that require a separate numerical calculation.

Note also that, although the shock-wave model is most popular, this does not rule out the possibility of constructing other models. In particular, a surfatron mechanism of particle acceleration by relativistic plasma waves in the halo was suggested in [14]. The author points out that the most remarkable property of the surfing is the absence of synchrotron radiation, the most dangerous channel of energy losses by ultrarelativistic particles. Clearly, the produced GCRs must also be isotropic in directions and can acquire energies up to 10^{16}–10^{20} eV per nucleon in this model as well.

APPENDIX

To derive the formula $j = \alpha\beta^2\gamma^{-\nu}$, we use the equations of relativistic hydrodynamics

$$\partial_i n u^i = 0, \quad \partial_i T^i_k = 0 \tag{9}$$

with the adiabat $p = p_0(n/n_0)^s$, given that the skinned current $I_0 = \text{const}$ produces the pressure $p = p_0(a_0/a)^2$ in a pinch of radius a. In the one-dimensional approximation of a "narrow channel," the two arguments are $\tau = ct$ and z.

Let us introduce dimensionless quantities,

$$c_0^2 = \frac{sp_0}{(s-1)n_0 m_i c^2} = \text{const}, \quad \rho = \frac{na^2}{n_0 a_0^2}, \tag{10}$$

and two functions, $x(\tau, z) = c_0^2/\rho$ and $y(\tau, z) = \operatorname{arccosh}\gamma$, for which we obtain the nonlinear equations

$$\left(\frac{\gamma}{x}\right)'_\tau + \left(\frac{u}{x}\right)'_z = 0, \quad u'_\tau + \gamma'_z + \gamma x'_z + u x'_\tau = 0. \tag{11}$$

If we introduce the inverse functions $\varphi(x, y)$ and $\psi(x, y)$ using the formulas

$$\begin{aligned} \tau &= T(x, y) = (\beta\psi - \varphi)W, \\ z &= Z(x, y) = (\psi - \beta\varphi)W, \quad W = xe^{-x}\gamma, \end{aligned} \tag{12}$$

then we obtain the linear equations

$$\begin{aligned} \psi'_y &= x(\varphi - \varphi'_x), \quad \varphi'_y = \psi'_x + \frac{\psi}{x}, \\ \hat{L}\varphi &= -\varphi''_{yy} = x\varphi''_{xx} + (2 - x)\varphi'_x - 2\varphi, \end{aligned} \tag{13}$$

where $\hat{L}$ is the operator of the Laguerre polynomials $\lambda_k = L_k^{(1)}(x)$ that satisfies the equation $\hat{L}\lambda_k = -(k + 2)\lambda_k$ for $k = 0, 1, 2, 3, \ldots$ Note that the term with $k = 0$ and $\lambda_0 = 1$ does not depend on the argument x.

The general solution is

$$\varphi = \sum_k C_k \lambda_k(x)\mu_k(y).$$

For the functions μ, we obtain the main equation of the problem

$$\mu''_{yy} = (k + 2)\mu, \quad \mu_k = \exp(-|y|\sqrt{k + 2}).$$

In the ultrarelativistic case ($\gamma \gg 1$), we can retain two terms, with $k = 0$ and $k = 1$. However, the zeroth term with $k = 0$ describes periodic (over the entire infinite pinch length) perturbations that cannot arise in cosmic conditions spontaneously. Therefore, the term with $k = 0$ should be discarded, and only the local solution

$$\phi_1 \approx C_1\lambda_1(x)\gamma^{-\sqrt{3}}$$

remains.

The number of particles accelerated in a single pinch is

$$N = \int(\pi a^2 n\gamma - \pi a_0^2 n_0)dz = \int\frac{dN}{du}du = \int\frac{dN}{d\varepsilon}d\varepsilon,$$

where $\varepsilon = E_{kin}/m_Z c^2$, m_Z is the mass of the nucleus with charge Z. Hence, we have the formulas

$$\beta\frac{dN}{d\varepsilon} = \frac{dN}{du} = \pi a_0^2 n_0\left(\rho - \frac{1}{\gamma}\right)dZ, \qquad (14)$$

where

$$\frac{dZ(x,y)}{dy} = Z'_x\frac{dx}{dy} + Z'_y.$$

The total derivative dZ/dx is calculated for a fixed time when

$$dT = T'_x dx + T'_y dy = 0, \quad dZ/dy = Z'_y - Z'_x(T'_y/T'_x).$$

Denoting $\Phi = \varphi - \psi'_y$ and $\Psi = \psi - \varphi'_y$, we find the partial derivatives

$$T'_x = \left(\beta\Psi - \frac{\Phi}{x}\right)W, \quad T'_y = (\Psi - \beta\Phi)W,$$
$$Z'_x = -\left(\frac{\beta}{x}\Phi + \Psi\right)W, \quad Z'_y = (\beta\Psi - \Phi)W. \qquad (15)$$

The spectrum is

$$\beta\frac{dN}{d\varepsilon} = \frac{dN}{du} = -\pi a_0^2 n_0\frac{e^{-x}}{\gamma}\left(c_0^2 - \frac{x}{\gamma}\right)\frac{\Phi^2 + x\Psi^2}{\Phi + x\beta\Psi}. \qquad (16)$$

We see from the relations $x = c_0^2/\rho = c_0^2(a_0/a)^{4/5}$ that the particles are accelerated while moving from the narrow places to the thickenings, where $a \longrightarrow \infty$ and $x \longrightarrow 0$. In the limit $x \longrightarrow 0$, when $\psi = 0$, we obtain a spectrum of the form

$$\beta\frac{dN}{d\varepsilon} = -\frac{A_0}{\gamma}\phi_1(x = 0, y),$$

where $A_0 = \pi a_0^2 n_0 c_0^2 = \text{const}$. For $y \gg 1$, we have $\phi_1 \propto \gamma^{-\sqrt{3}}$, and the expression $dN/d\varepsilon \propto \gamma^{-\nu}$ with the index $\nu = 1 + \sqrt{3}$ describes well the entire ultrarelativistic energy range.

Multiplying the spectrum by the velocity $v = c\beta$ yield the intensity of the particle beam:

$$j = v\frac{dN}{du} = c\beta^2\frac{dN}{d\varepsilon} \approx \alpha_i\beta^2\gamma^{-\nu}, \qquad (17)$$

where the coefficients $\alpha_i = \text{const}$ are different for nuclei with different charges Z_i. However, since the charge of the nucleus is proportional to its mass m_Z, a single argument is suitable for all nuclei: the dimensionless kinetic energy $\varepsilon = \gamma - 1 = E_{kin}/m_Z c^2$. Thus, we see that all bare nuclei are accelerated at the same potential difference V, acquiring energies $E_{kin} = Z_i|e|V$ proportional to the nuclear masses.

ACKNOWLEDGMENTS

I wish to thank A.A. Petrukhin, A.V. Byalko, V.P. Vlasov, G.I. Merzon, and D.V. Shuvaev for the stimulating discussions of the questions touched upon here.

REFERENCES

1. V. L. Ginzburg, Usp. Fiz. Nauk **51**, 343 (1953); **163** (7), 45 (1993) [Phys. Usp. **36**, 587 (1993)].
2. A. A. Vedenov, E. P. Velikhov, and R. Z. Sagdeev, Yad. Sint. **2**, 465 (1962).
3. Nature online, http://www.nature.com/news/2005/050110/full/050110-18.html.
4. L. G. Dedenko, in *Abstracts of 28th All-Russian Conference on Cosmic Rays, Moscow, 2004* (MIFI, Moscow, 2004), p. 44.
5. A. A. Petrukhin, in *Abstracts of 28th All-Russian Conference on Cosmic Rays, Moscow, 2004* (MIFI, Moscow, 2004), p. 85; Nucl. Phys. B (Proc. Suppl.) **136**, 218 (2004).
6. N. N. Kalmykov and A. V. Timokhin, in *Abstracts of 28th All-Russian Conference on Cosmic Rays, Moscow, 2004* (MIFI, Moscow, 2004), p. 37.
7. V. I. Galkin *et al.*, in *Abstracts of 28th All-Russian Conference on Cosmic Rays, Moscow, 2004* (MIFI, Moscow, 2004), p. 21.
8. V. P. Vlasov, S. K. Zhdanov, and B. A. Trubnikov, Pis'ma Zh. Éksp. Teor. Fiz. **49**, 581 (1989) [JETP Lett. **49**, 667 (1989)].
9. B. A. Trubnikov, in *Plasma Physics and Problem of Controlled Thermonuclear Reactions*, Ed. by M. A. Leontovich (Akad. Nauk SSSR, Moscow, 1958), Vol. 1, p. 289 [in Russian].
10. B. A. Trubnikov, *The Plasma Theory* (Atomizdat, Moscow, 1996) [in Russian]; S. K. Zhdanov and B. A. Trubnikov, *Quasi-Gaseous Unstable Media* (Nauka, Moscow, 1991) [in Russian]; S. Zhdanov, S. Zverev, and B. Trubnikov, *Hydrodynamics of Unstable Media* (CRC Press, Boca Raton, 1996); B. A. Trubnikov, in *Book of Abstracts of 10th International Conference and School on Plasma Physics and Controlled Fusion* (Alushta, Crimea, 2004), p. 14.
11. F. A. Aharonian *et al.*, Nature **432**, 75 (2004).
12. R. Enomoto *et al.*, Nature **416**, 823 (2002).
13. Yu. V. Stenkin, Mod. Phys. Lett. A **18**, 1225 (2003).
14. G. N. Kichigin, in *Abstracts of 28th All-Russian Conference on Cosmic Rays, Moscow, 2004* (MIFI, Moscow, 2004), p. 25.

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 169–176.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 193–200.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Surface Gravity Waves in Deep Fluid at Vertical Shear Flows¶

G. Gogoberidze[a], L. Samushia[a], G. D. Chagelishvili[a], J. G. Lominadze[a], and W. Horton[b]

[a]*Center for Plasma Astrophysics, Abastumani Astrophysical Observatory, 0160, Tbilisi, Georgia*
[b]*Institute for Fusion Studies, The University of Texas at Austin, Austin, Texas 78712, USA*
e-mail: gogober@geo.net.ge
Received October 28, 2004

Abstract—Special features of surface gravity waves in a deep fluid flow with a constant vertical shear of velocity is studied. It is found that the mean flow velocity shear leads to a nontrivial modification of the dispersive characteristics of surface gravity wave modes. Moreover, the shear induces generation of surface gravity waves by internal vortex mode perturbations. The performed analytical and numerical study show that surface gravity waves are effectively generated by the internal perturbations at high shear rates. The generation is different for the waves propagating in the different directions. The generation of surface gravity waves propagating along the main flow considerably exceeds the generation of surface gravity waves in the opposite direction for relatively small shear rates, whereas the latter wave is generated more effectively for high shear rates. From the mathematical standpoint, the wave generation is caused by non-self-adjointness of the linear operators that describe the shear flow. *© 2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Generation of surface gravity waves (SGW), which are the best known sea and oceanic waves, is naturally associated with winds. Momentum transfer from wind to undulating movement of the ocean, which is the basic mechanism of the generation of surface waves, is investigated since Kelvin's pioneering work [1]. Independent and intercomplementary theories of Phillips [2] and Miles [3–6] provide the basics of theoretical understanding of surface wave generation by wind. Phillips' resonant mechanism is responsible for excitation and initial rising of wave motion on an unexcited surface of the fluid; Miles' mechanism—energy transfer from wind to fluid as a consequence of the interaction between wind shear flow and surface waves—is responsible for subsequent amplification of the waves. According to Miles' mechanism, the energy source is the wind shear flows situated outside the fluid. Other ways of SGW generation have also been studied, such as the possibility of SGW generation by earthquakes [7, 8] and the theory of SGW generation by intrafluid explosions [9]. In the theories mentioned above, the sources of SGW generation are extrinsic to the fluid.

The question arises as to whether sources intrinsic for the fluid (shear flows and vortex perturbations, for example) can generate SGW.

This question becomes especially interesting in view of the impressive progress made in the understanding of spectrally stable shear flow phenomena by the hydrodynamic community in the past ten years. The early transient period for the perturbations has been shown to reveal rich and complicate behavior in smooth (without inflection point) shear flows. In particular, it has been shown that the linear dynamics of perturbations in the flows are accompanied by intense temporal energy exchange processes between the background flow and perturbations and/or between different modes of perturbations. From the mathematical standpoint, these effects are caused by the non-self-adjointness of the linear operators in shear flows and are adequately described in the framework of the so-called nonmodal approach (see, e.g., [10–12]). The nonmodal approach involves a change of independent variables from the laboratory frame to a moving frame and the study of temporal evolution of spatial Fourier harmonics (SFHs) of perturbations without any spectral expansion in time.

We examine the linear dynamics of surface waves and internal perturbations in deep fluid in the absence of wind and in the presence of the fluid flow with a vertical shear of velocity. Dispersive characteristics of shear-modified SGWs and the linear mechanism of the generation of surface waves in deep fluid by internal perturbations are studied in detail in the framework of the nonmodal approach.

The paper is organized as follows. The mathematical formalism is presented in Section 2. Shear-modified SGWs and their generation are analyzed in Section 3. Applications of the obtained results to the concrete physical problems are discussed in Section 4. Conclusions are given in Section 5.

¶ This article was submitted by the authors in English.

1063-7761/05/10101-0169$26.00 © 2005 Pleiades Publishing, Inc.

2. MATHEMATICAL FORMALISM

We consider deep fluid with the flat outer surface at $z = 0$ and a constant shear flow $\mathbf{U}_0 = (Az, 0, 0)$ for $z < 0$. The shear parameter A is considered positive for simplicity. The gravitational field is considered uniform, $\mathbf{g} = (0, 0, -g)$. Generally, four modes of perturbation (SGW, internal gravity waves, sound waves, and vortex mode) can exist in the system. To reduce the mathematical complications as much as possible but still keep the basic physics of our analysis, we consider fluid to be incompressible (neglecting sound waves) and disregard the stratification effects (assuming that the frequency of internal gravity waves is much less than the frequency of SGWs, i.e., considering internal gravity waves as aperiodic/vortex mode perturbations). We also ignore the effects of viscosity in what follows. After these simplifications, we keep just two modes of perturbation, SGW and the vortex mode, and write the differential equations for the linear dynamics of perturbations of velocity ($\mathbf{u}'$) and normalized pressure ($p' = p/\rho_0$) as

$$\frac{\partial u_x'}{\partial x} + \frac{\partial u_y'}{\partial y} + \frac{\partial u_z'}{\partial z} = 0, \tag{1}$$

$$\frac{\partial u_x'}{\partial t} + Az\frac{\partial u_x'}{\partial x} + Au_z' = -\frac{\partial p'}{\partial x}, \tag{2}$$

$$\frac{\partial u_y'}{\partial t} + Az\frac{\partial u_y'}{\partial x} = -\frac{\partial p'}{\partial y}, \tag{3}$$

$$\frac{\partial u_z'}{\partial t} + Az\frac{\partial u_z'}{\partial x} = -\frac{\partial p'}{\partial z}, \tag{4}$$

with the boundary condition on the surface $z = 0$:

$$\left(\frac{\partial p'}{\partial t} - gu_z'\right)\Big|_{z=0} = 0. \tag{5}$$

We use the standard technique of the nonmodal approach [10]: introduction of comoving variables ($x' = x + Azt$, $y' = y$, $z' = z$, $t' = t$) allows us to transform the spatial inhomogeneity presented in Eqs. (1)–(5) into a temporal one. Then, after the Fourier transformation with respect to x' and y',

$$\mathbf{u}'(\mathbf{r}, t) = \frac{1}{4\pi^2}\int \mathbf{u}(k_x, k_y, z', t) \times \exp[i(k_x x' + k_y y')]dk_x dk_y, \tag{6}$$

the dynamic equations are reduced to

$$ik_x u_x + ik_y u_y + \left(\frac{\partial}{\partial z'} - iAt'k_x\right)u_z = 0, \tag{7}$$

$$\frac{\partial u_x}{\partial t'} + Au_z = -ik_x p, \tag{8}$$

$$\frac{\partial u_y}{\partial t'} = -ik_y p, \tag{9}$$

$$\frac{\partial u_z}{\partial t'} = -\left(\frac{\partial}{\partial z'} - iAt'k_x\right)p, \tag{10}$$

$$\left(\frac{\partial p}{\partial t'} - gu_z\right)\Big|_{z'=0} = 0. \tag{11}$$

Hereafter, the primes of the z' and t' variables are omitted.

From this set, we readily obtain the equation for the perturbation of the vertical component of velocity,

$$\frac{\partial}{\partial t}\left(\left[\tilde{k}^2 - \left(\frac{\partial}{\partial z} - iAtk_x\right)^2\right]u_z\right) = 0, \tag{12}$$

where $\tilde{k} = \sqrt{k_x^2 + k_y^2}$.

All other perturbed quantities (u_x, u_y, and p) can be readily expressed through u_z by combining Eqs. (7)–(10); e.g., for p, we have

$$p = -\frac{1}{\tilde{k}^2}\left(\frac{\partial}{\partial t}\left[\left(\frac{\partial}{\partial z} - iAtk_x\right)u_z\right] - iAk_x u_z\right). \tag{13}$$

Integration of Eq. (12) with respect to time yields

$$\left[\tilde{k}^2 - \left(\frac{\partial}{\partial z} - iAtk_x\right)^2\right]u_z(k_x, k_y, z, t) = F(k_x, k_y, z), \tag{14}$$

where $F(k_x, k_y, z)$ is the constant (in time) of integration and defines the internal vortex mode perturbation in the flow: $F(k_x, k_y, z) = 0$ relates to the case where the internal perturbation is absent.

The Fourier transformation with respect to z,

$$\begin{bmatrix} u_z(k_x, k_y, z, t) \\ F(k_x, k_y, z) \end{bmatrix} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\begin{bmatrix} u_z(k_x, k_y, k_z, t) \\ \tilde{F}(k_x, k_y, k_z) \end{bmatrix} e^{ik_z z}dk_z, \tag{15}$$

reduces Eq. (14) to

$$k^2(t)u_z(k_x, k_y, k_z, t) = \tilde{F}(k_x, k_y, k_z) + 4i\pi\tilde{k}C(k_x, k_y, t), \tag{16}$$

where

$$C \equiv \frac{1}{4i\pi\tilde{k}} \times \left[\left(\frac{d}{dz} - 2iAtk_x - ik_z\right)u_z(k_x, k_y, z, t)\right]\Big|_{z=0}. \quad (17)$$

Defining $u_z(k_x, k_y, k_z, t)$ from Eq. (16), making the inverse Fourier transform with respect to k_z, taking the boundary condition $|u_z| < \infty$ at $z = -\infty$ into account, and recalling that $C(k_z, k_y, t)$ is independent of z, we obtain

$$u_z(k_x, k_y, z, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\tilde{F}(k_x, k_y, k_z)}{k^2(t)}\exp(ik_z z)dk_z + C(k_x, k_y, t)\exp[(\tilde{k} + iAtk_x)z], \quad (18)$$

where $k^2(t) = \tilde{k}^2 + k_z^2(t)$ and $k_z(t) \equiv k_z - Atk_x$.

The first term in Eq. (18) is related to the vortex mode perturbation [11, 13], whereas the second term, which is exponentially decreasing with the depth, is related to the SFHs of shear modified surface waves.

Substituting Eq. (18) in Eq. (13) and using boundary condition (Eq. (11)), we obtain

$$\frac{d^2C}{dt^2} + \frac{iAk_x}{\tilde{k}}\frac{dC}{dt} + \tilde{k}gC = I(k_x, k_y, t), \quad (19)$$

where

$$I(k_x, k_y, t) \equiv \int_{-\infty}^{\infty}\left[8iA^2k_x^2\tilde{k}\frac{k_z(t)}{k^6(t)} - \frac{\tilde{k}g}{k^2(t)}\right]\tilde{F}(\mathbf{k})dk_z. \quad (20)$$

Generally, Eqs. (19) and (20) describe the dynamics of surface wave SFHs in the presence of the internal vortical source: the term $I(k_x, k_y, t)$ is the result of an interplay of the mean flow shear and the internal vortical perturbations and couples the latter perturbation to the surface one. Hence, there is no coupling between these perturbations in the absence of the shear. Indeed, if there are no surface perturbations initially [$u_z(k_x, k_y, z = 0, t = 0) = 0$], then we readily obtain from Eqs. (16) and (20) that $I(k_x, k_y, t) \sim u_z(k_x, k_y, z = 0, t = 0)$ at $A = 0$, i.e., $I(k_x, k_y, t) \equiv 0$. Thus, if there is no the source in a shearless flow initially, it does not appear afterward.

3. SGWs AND THEIR GENERATION IN SHEAR FLOW

We can see from Eqs. (19) and (20) that there are two main effects of the shear: first, the second term in the left-hand side of Eq. (19) indicates that the velocity shear affects the frequencies of SGWs. Second, the source term $I(k_x, k_y, t)$ caused by the internal perturbations couples the internal and surface perturbations and results in the emergence/generation of SGW in the flow. Our further attempts are focused on the study of these effects.

3.1. Shear Modified SGWs

In this subsection, we study shear-induced modifications of the properties of SGWs. For this, we assume that there were no vortex mode perturbations initially, $\tilde{F}(k_x, k_y, k_z) = 0$. Consequently, $I(k_x, k_y, t) = 0$ [see Eq. (20)], and Eq. (19) reduces to a homogeneous one, with the solution

$$C_h(k_x, k_y, t) = C_1(k_x, k_y) \times \exp(-i\Omega_1 t) + C_2(k_x, k_y)\exp(-i\Omega_2 t), \quad (21)$$

where $C_{1,2}(k_x, k_y)$ are determined by the initial conditions and

$$\Omega_{1,2} = \pm\sqrt{\tilde{k}g + \frac{A^2k_x^2}{4\tilde{k}^2}} - \frac{Ak_x}{2\tilde{k}} = \sqrt{\tilde{k}g}\left(\pm\sqrt{1 + S^2\frac{k_x^2}{\tilde{k}^2}} - S\frac{k_x}{\tilde{k}}\right) \quad (22)$$

represents shear-modified frequencies of SFH of SGWs propagating in the opposite directions and $S \equiv A/(4\tilde{k}g)^{1/2}$ is the dimensionless shear rate. This equation shows that, in contrast to acoustic and magnetohydrodynamic wave modes [14–16], the presence of the shear does not lead to the time variability of the frequency. However, velocity shear leads to a nontrivial modification of the frequencies and, consequently, phase velocities of SFH [17, 18]. Indeed, for the value of the phase velocity, Eq. (22) gives

$$V_{\mathrm{ph}}(S, \phi) = \sqrt{\frac{g}{\tilde{k}}}(\sqrt{1 + S^2\cos^2\phi} - S\cos\phi), \quad (23)$$

where $\phi \equiv \arccos(k_x/\tilde{k})$.

The phase velocity is isotropic in the shearless limit ($S = 0$) and depends on ϕ in the shear flow. The anisotropy increases with the shear rate. The value of the phase velocity is minimal at $\phi = 0$, $V_{\mathrm{ph}}^{\min} = \sqrt{g/\tilde{k}}(\sqrt{1 + S^2} - S)$, and is maximal at $\phi = \pi$, $V_{\mathrm{ph}}^{\max} = \sqrt{g/\tilde{k}}(\sqrt{1 + S^2} + S)$. We suppose that a SGW is emitted by a point source situated on the surface at $x = y = 0$.

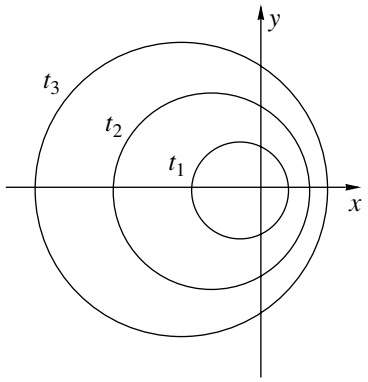

Fig. 1. Shear-induced anisotropy of SGW propagation. The leading wave crest at three different time instants t_1, t_2, t_3, with $t_2 = 2t_1$, $t_3 = 3t_1$, which are circular but not concentric. A point source of the SGW is located at $x = y = 0$.

From Eq. (23), it then follows that the propagation of the leading wave crest is described by

$$r(S, \phi, t) = V_{\rm ph}(S, \phi)t = \sqrt{\frac{g}{\tilde{k}}}(\sqrt{1 + S^2\cos^2\phi} - S\cos\phi)t. \quad (24)$$

Figure 1 shows the leading wave crest of the SGW for three different time instants t_1, t_2, t_3, with $t_2 = 2t_1$, $t_3 = 3t_1$, which are circular but not concentric.

3.2. Generation of SGWs by Internal Vortices

We first analyze the source term $I(k_x, k_y, t)$, which is determined by $\tilde{F}(k_x, k_y, k_z)$. We assume that $\tilde{F}(k_x, k_y, k_z)$ is a localized function in the wavenumber space, with the center of localization at $\mathbf{k}_0 = (k_{x0}, k_{y0}, k_{z0})$. We note that the first factor in the integrand in Eq. (20) reaches its maximum when $k_z - Ak_x t = 0$. Consequently, the maximum of the integral is in the vicinity of the time instant $t = t_* \equiv k_{z0}/(Ak_{x0})$. Equation (20) implies that generally, $I(k_x, k_y, t)$ tends to zero in both limits $t \longrightarrow \pm\infty$. Actually, there exists some time interval $2\Delta t$ around t_*, where the source term differs from zero. The value of Δt depends on the degree of localization of the internal perturbation, i.e., of $\tilde{F}(k_x, k_y, k_z)$, in the wavenumber space. (The source localization is demonstrated below in a specific example.) Thus, in the case of a localized source, the coupling between surface (gravity wave) and internal (vortex mode) perturbations takes place in some time interval $2\Delta t$ around t_*, and these perturbations can be considered separately at $|t - t_*| > \Delta t$.

The general solution of the inhomogeneous equation (Eq. (19)) is the sum of the general solution of the corresponding homogeneous equation and a partial solution of the equation

$$C(k_x, k_y, t) = C_h(k_x, k_y, t) + C_i(k_x, k_y, t). \quad (25)$$

The general solution $C_h(k_x, k_y, t)$ is given by Eq. (21), whereas a partial solution of Eq. (19) is

$$C_i = \frac{1}{2\Omega_0}\exp(-i\Omega_1 t)\int_{t_0}^{t} I(k_x, k_y, t')\exp(i\Omega_1 t')dt' - \frac{1}{2\Omega_0}\exp(-i\Omega_2 t) \times \int_{t_0}^{t} I(k_x, k_y, t')\exp(i\Omega_2 t')dt', \quad (26)$$

where

$$\Omega_0 = \sqrt{\tilde{k}g + \frac{A^2k_x^2}{4\tilde{k}^2}} = \sqrt{\tilde{k}g}\sqrt{1 + S^2\frac{k_x^2}{\tilde{k}^2}}. \quad (27)$$

We assume that the coupling between the surface and internal modes can be neglected at the initial time instant t_0, i.e., $t_0 < t_* - \Delta t$. After passing through the coupling time interval, for any $t > t_f = t_* + \Delta t$, the modes become independent again. However, during the time interval $[t_0, t_f]$, internal vortices generate SGWs with frequencies Ω_1 and Ω_2 [see Eq. (22)]. As follows from Eqs. (21), (25), and (26), if there are no SGWs ($C_{1,2} = 0$) initially, then the generated SFH amplitudes ($Q_{1,2}$) are

$$Q_1(k_x, k_y) = \frac{1}{2\Omega_0} \times \left|\int_{t_0}^{t_f} I(k_x, k_y, t')\exp(i\Omega_1 t')dt'\right|, \quad (28)$$

$$Q_2(k_x, k_y) = \frac{1}{2\Omega_0} \times \left|\int_{t_0}^{t_f} I(k_x, k_y, t')\exp(i\Omega_2 t')dt'\right|. \quad (29)$$

We can replace the integration limits by $\pm\infty$. After integration over time, this yields

$$Q_{1,2} = \frac{\pi k_x}{\tilde{k}^3}\left(\frac{A}{2\Omega_0} \mp \frac{k_x}{\tilde{k}}\right)\exp\left[-\frac{(\Omega_0 \mp A/2)\tilde{k}}{Ak_x}\right]$$

$$\times \int \tilde{F}(k_x, k_y, k_z)\exp\left(-i\frac{(\Omega_0 \mp A/2)k_z}{Ak_x}\right)dk_z$$
$$= \frac{2\pi^2 k_x}{\tilde{k}^3}\left(\frac{A}{2\Omega_0} \mp \frac{k_x}{\tilde{k}}\right)\exp\left[-\frac{(\Omega_0 \mp A/2)\tilde{k}}{Ak_x}\right] \quad (30)$$
$$\times F\left(k_x, k_y, -\frac{\Omega_0 \mp A/2}{Ak_x}\right).$$

We note that the last factors in Eq. (30) are proportional to the vorticity of the initial perturbations at $z_{1,2} = -(\Omega_0 \mp A/2)/(Ak_x)$, respectively. The second factors indicate that, at small shear rates ($S \equiv A/\sqrt{4\tilde{k}g} \ll 1$), the amplitudes of the generated SGWs are exponentially small with respect to the large parameter $1/S$. Equation (30) also indicates that, for a fixed k_x, the generation is most efficient in the two-dimensional case ($k_y = 0$).

We now describe the dynamic picture for a specific example, where a pure internal vortex-mode perturbation (without any admixture of surface waves) is imposed in the flow initially. For simplicity, we consider the two-dimensional problem, where $\partial/\partial y = 0$. The vertical velocity of the imposed perturbation is given by

$$u_z(x, z, t_0) = z^3\eta(-z)$$
$$\times \exp\left(-\frac{[(z + z_0)\cos\phi + x\sin\phi]^2}{L_1^2}\right) \quad (31)$$
$$\times \exp\left(-\frac{[(z + z_0)\sin\phi - x\cos\phi]^2}{L_2^2}\right),$$

where $\eta(z)$ is Heaviside function, $(0, -z_0)$ is the center of the localization, $L_{1,2}$ characterize the vertical and horizontal scales, respectively, and ϕ is the slope of the perturbation.

Numerical solution of the problem was performed as follows. Fourier transformation of Eq. (31) with respect to the x variable allows us to determine $F(k_x, z)$ through Eq. (14). Another Fourier transformation with respect to z yields $\tilde{F}(k_x, k_z)$. Then the source function $I(k_x, t)$ is found by Eq. (20). Thus, the solution of the problem for a fixed k_x reduces to the numerical solution of the inhomogeneous equation (Eq. (19)) with the known $I(k_x, t)$.

The dependence of the source function $I(k_x, t)$ on t at $L_1 = 1$, $L_2 = 7$, $\phi = \pi/18$, $k_x = 1$, and $z_0 = 2$ for two different values of the shear rate $S = 0.08$ (dashed line) and $S = 0.32$ (solid line) is presented in Fig. 2. As was mentioned above, the source term is a localized function and considerably differs from zero only in the interval $t \in (20, 40)$ for $S = 0.08$ and $t \in (5, 10)$ for $S = 0.32$.

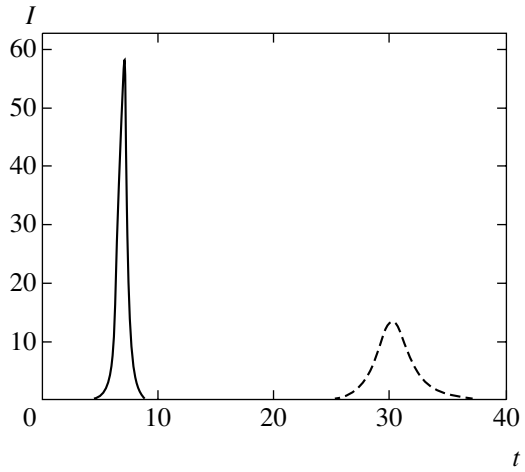

Fig. 2. $I(k_x, k_y, t)$ vs. time at $S = 0.32$ (solid line) and $S = 0.08$ (dashed line), $k_x = 1$, $L_1 = 1$, $L_2 = 7$, $z_0 = 2$, and $\phi = \pi/18$.

To analyze the wave generation efficiency, it is useful to introduce the generation coefficients that characterize the ratio of the generated wave energy density and the maximum energy density of the initial vortex mode perturbations for a fixed value of k_x. Taking into account that the maximum energy density of the vortex mode perturbations is

$$E_v = \frac{1}{2k_x^4}\int_{-\infty}^{\infty}|F(k_x, z)|^2 dz \quad (32)$$

and the energy density of the generated waves is

$$E_{w1,2} = \frac{1}{k_x}Q_{1,2}^2(k_x), \quad (33)$$

we define the dimensionless generation coefficients as

$$G_{1,2} = Q_{1,2}(k_x)\left(\frac{2k_x^3}{\int_{-\infty}^{\infty}|F(k_x, z)|^2 dz}\right)^{1/2}. \quad (34)$$

Figure 3 represents the generation coefficients G_1 (dashed line) and G_2 (solid line) vs. the shear rate S at $L_1 = 1$, $L_2 = 7$, $\phi = \pi/18$, $k_x = 1$, and $z_0 = 2$. As can be seen, at small values of the shear rate, generation of SGW with the frequency Ω_1 (i.e., propagating along the x axis) considerably exceeds the generation of SGW with the frequency Ω_2 (i.e., propagating against the x

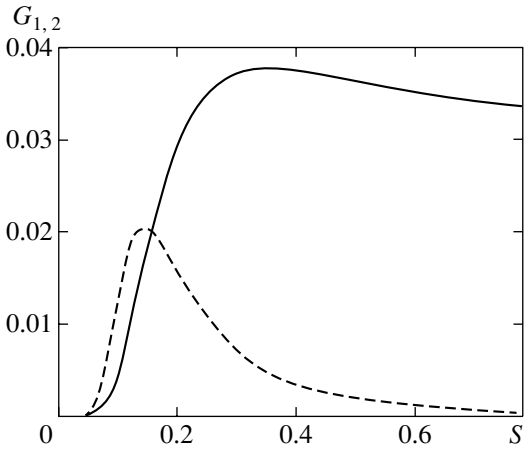

Fig. 3. The generation coefficients G_1 (dashed line) and G_2 (solid line) vs. the shear rate S at $k_x = 1$, $L_1 = 1$, $L_2 = 7$, $z_0 = 2$, and $\phi = \pi/18$.

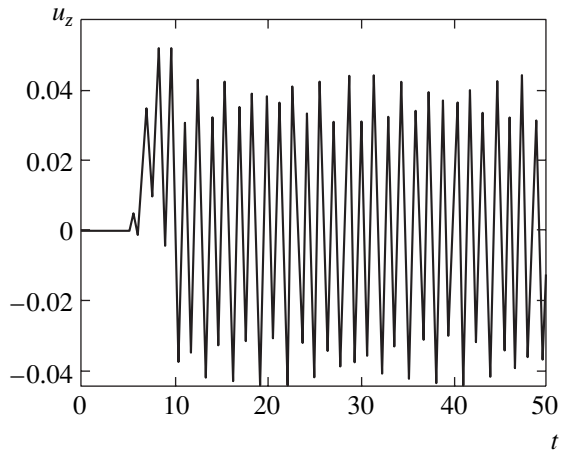

Fig. 4. $u_z(k_x, t)$ vs. time at $S = 0.32$, $k_x = 1$, $L_1 = 1$, $L_2 = 7$, $z_0 = 2$, and $\phi = \pi/18$.

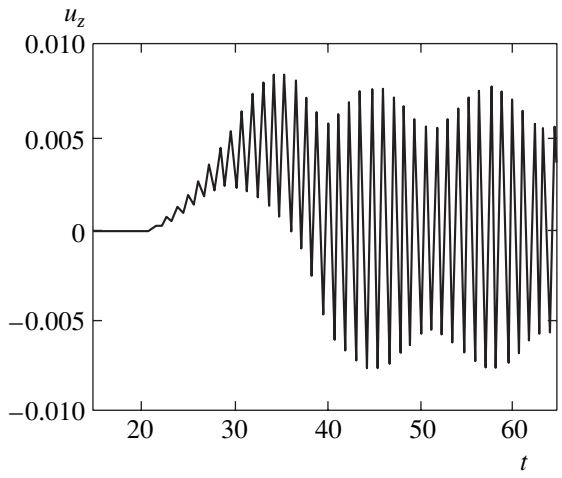

Fig. 5. $u_z(k_x, t)$ vs. time at $S = 0.08$, $k_x = 1$, $L_1 = 1$, $L_2 = 7$, $z_0 = 2$, and $\phi = \pi/18$.

axis), whereas the latter wave is generated more efficiently at $S > 0.15$.

The wave generation is well traced in Figs. 4 and 5, where the temporal evolution of the vertical component of velocity perturbation at the surface obtained by numerical solution of Eqs. (19) and (20) is presented for $S = 0.32$ and $S = 0.08$, respectively. The other parameters are the same as in Fig. 2. A purely internal vortex mode perturbation is imposed in the equations initially. The generation occurs in the time interval where $I(k_z, t)$ noticeably differs from zero. Afterwards, just (two) waves with different frequencies and amplitudes exist. At $S = 0.32$, presented in Fig. 4, the generation occurs in the time interval $t \in (5, 10)$. Besides, the SGW propagating against the x axis is mainly generated. In contrast to this, at $S = 0.08$, presented in Fig. 5, the generation of SGW propagating along the x axis dominates. These numerical results are in agreement with the analytical ones (see Eq. (30) and Fig. 3).

4. DISCUSSION

In the previous sections, a simplified model was considered that allowed us to simplify the mathematical description and study shear-induced effects in a "pure" form. For instance, the influence of the viscosity was ignored and the density ratio ρ_a/ρ_0 of the fluids above and below the surface $z = 0$ was assumed to be zero. The last assumption allows us to ignore all the dynamical processes in the upper fluid. On the other hand, it is well known that, in the case of ocean waves, the wind is the most important and powerful source of the waves. In this section, we discuss possible applications of the studied linear effects to the concrete physical situations.

4.1. Ocean Waves

It is well known [2–6] that the wind is the main source of ocean SGWs. In the context of future discussion, the papers of Chalikov's group [19, 20] should also be noted, where the influence of small-scale turbulence in the air on the wave growth was studied in detail. At present, there exists a well-developed theory of both SGW generation and nonlinear evolution that is mainly confirmed by experiments as well as numerical simulations (see, e.g., [21] for a recent review). After development of a wind-driven instability, nonlinear four-wave resonant interactions transfer the wave energy to smaller scales. The existing theory predicts that, for relatively small frequencies, the Zakharov–Philonenko [22] spectrum $E(\omega) \propto \omega^{-4}$ of SGW fluctuations (sometimes called Toba's spectrum) should be observed (in this context, see also [23]), whereas for relatively high wavenumbers, nonlinearity becomes strong and the Phillips spectrum $E(\omega) \sim \omega^{-5}$ of the wave turbulence should develop. The existing observations confirm these predictions and provide that, in the range

$\omega_p/3 < \omega < 3\omega_p$, where ω_p is the peak frequency, the Zakharov–Philonenko spectrum is usually observed. For $\omega > 3\omega_p$, the spectrum becomes very close to the Phillips one [21]. The properties of the wave spectrum in the very short wavelength region, as well as dynamics of dissipation of SGW turbulent fluctuations, are much less clear [24].

In the case of ocean waves, the presented linear mechanism of SGW generation can make an important contribution to the balance of small-scale SGW fluctuations. Indeed, a characteristic length scale of the turbulence is much smaller at the ocean surface than in the air. Namely, the characteristic length and velocity scales are $u_* \sim 1$ cm/s and $l \sim 1$ cm, respectively [25]. On the other hand, in the presence of the wind, the strong velocity shear $A \sim 10$ s^{-1} is present in the so-called "buffer layer" [26] of the water, with the thickness $l_1 \sim (20\text{–}100)l_0$, where $l_0 \approx \nu/u_*$ is the dissipation length scale and ν is the kinematic viscosity of water. Simple estimates yield $l_1 \sim (0.5\text{–}1)$ cm. The linear mechanism presented implies that vortical perturbations generate SGWs with the same length scale. Therefore, in the case of ocean waves, internal vortex-mode perturbations should effectively generate small-scale SGWs, with the wavelength just above the capillary length scale $\lambda_c = 0.39$ cm [27]. In this context, the study of the influence of capillary effects on the processes discussed above seems to be interesting. Analysis of this problem will be presented elsewhere.

4.2. Interfacial Gravity Waves

In the analysis in Sections 2 and 3, the density ratio ρ_a/ρ_0 of the fluids above and below the surface $z = 0$ was assumed to be zero. The obtained results can be readily generalized to the case of interfacial GWs. If the densities of the upper and lower fluids are ρ_1 and ρ_2 and the shear rates are A_1 and A_2, respectively, then the shear-modified dispersion of interfacial GWs is given by the same expression (22) with g and A replaced by g_* and A_*, where

$$g_* = g\frac{\rho_2 - \rho_1}{\rho_2 + \rho_1}, \quad A_* = \frac{A_2\rho_2 - A_1\rho_1}{\rho_2 + \rho_1}. \tag{35}$$

This equation implies that the influence of shear on both the wave dispersion and the coupling with internal vortex perturbations, which is determined by the dimensionless parameter

$$S_* \equiv \frac{A_*}{\sqrt{4\tilde{k}g_*}} = S_2\frac{1 - \rho_1A_1/\rho_2A_2}{\sqrt{1 - \rho_1^2/\rho_2^2}}, \tag{36}$$

is much more notable when the fluids have comparable densities if ρ_1A_1 is not very close to ρ_2A_2. Therefore, the described shear-induced effects should usually have much stronger effect on the dynamics of interfacial waves than on ocean waves.

5. SUMMARY

We summarize the main features of the linear dynamics of surface gravity waves in a simplified deep fluid (at $z < 0$) flow with vertical shear of the mean velocity $\mathbf{U}_0 = (Az, 0, 0)$. The simplification lies in neglecting the fluid compressibility and stratification, in other words, in the consideration of the system containing just two modes of perturbation: the surface gravity wave mode and the internal vortex mode. Special features of SGW in the system are as follows.

The mean flow velocity shear causes a nontrivial modification of the frequencies and phase velocities of SGWs. The frequencies are defined by Eq. (22). The phase velocity becomes anisotropic (see Eq. (23) and Fig. 1): its value is minimal for SFH propagating along the x axis [$V_{\rm ph}^{\rm min} = \sqrt{g/\tilde{k}}\,(\sqrt{1 + S^2} - S)$] and maximal for SFH propagating against the x axis [$V_{\rm ph}^{\rm max} = \sqrt{g/\tilde{k}}\,(\sqrt{1 + S^2} + S)$].

The mean flow velocity shear leads to the appearance of an intrinsic (to the fluid) source of SGW generation via coupling the wave to the internal vortex-mode perturbations; the coupling results in the emergence/generation of SGWs by internal vortex-mode perturbations at $S \gtrsim 0.05$. The generation is different for the waves propagating in the different directions (see Eq. (30)). The generation of SGW with the frequency Ω_1 considerably exceeds the generation of SGW with the frequency Ω_2 for relatively small shear rates S, whereas the latter wave is generated more effectively for high shear rates ($S > 0.15$).

ACKNOWLEDGMENTS

This research is supported by ISTC (grant no. G 553). The work was supported in part by the Department of Energy (grant no. DE-FG03-96ER-54346).

REFERENCES

1. W. Kelvin, Philos. Mag. **42**, 368 (1871).
2. P. O. Phillips, J. Fluid Mech. **2**, 417 (1957).
3. J. W. Miles, J. Fluid Mech. **3**, 185 (1957).
4. J. W. Miles, J. Fluid Mech. **6**, 585 (1959).
5. J. W. Miles, J. Fluid Mech. **10**, 496 (1961).
6. J. W. Miles, J. Fluid Mech. **13**, 433 (1962).
7. K. Kajiura, J. Oceanogr. Soc. Jpn. **18**, 51 (1962).
8. K. Kajiura, J. Oceanogr. Soc. Jpn. **28**, 32 (1972).
9. H. C. Kranzer and J. B. Keller, J. Appl. Phys. **30**, 398 (1959).
10. P. Goldreich and D. Lynden-Bell, Mon. Not. R. Astron. Soc. **130**, 125 (1965).

11. W. O. Criminale and P. G. Drazin, Stud. Appl. Math. **83**, 123 (1990).
12. G. D. Chagelishvili, A. D. Rogava, and D. G. Tsiklauri, Phys. Rev. E **53**, 6028 (1996).
13. S. J. Chapman, J. Fluid Mech. **451**, 35 (2002).
14. G. D. Chagelishvili, A. G. Tevzadze, G. Bodo, and S. S. Moiseev, Phys. Rev. Lett. **79**, 3178 (1997).
15. G. D. Chagelishvili, R. G. Chanishvili, J. G. Lominadze, and A. G. Tevzadze, Phys. Plasmas **4**, 259 (1997).
16. A. D. Rogava, S. Poedts, and S. M. Mahajan, Astron. Astrophys. **354**, 749 (2000).
17. A. D. D. Craik, J. Fluid Mech. **37**, 531 (1968).
18. V. I. Shrira, J. Fluid Mech. **252**, 565 (1993).
19. D. V. Chalikov, Dokl. Akad. Nauk SSSR **229**, 1083 (1976).
20. D. V. Chalikov and V. K. Makin, Boundary-Layer Meteorol. **56**, 83 (1991).
21. J. P. Janssen, *The Interaction of Ocean Waves and Wind* (Cambridge Univ. Press, Cambridge, 2004).
22. V. E. Zakharov and N. N. Philonenko, Sov. Phys. Dokl. **11**, 881 (1967).
23. S. A. Kitaigorodskii, J. Phys. Oceanogr. **13**, 816 (1983).
24. V. E. Zakharov, Eur. J. Mech. B **18**, 327 (1999).
25. S. A. Kitaigorodskii and J. L. Lumley, J. Phys. Oceanogr. **13**, 1977 (1983).
26. P. A. Chang, U. Piomelli, and W. K. Blake, Phys. Fluids **11**, 3434 (1999).
27. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 6: *Fluid Mechanics*, 4th ed. (Nauka, Moscow, 1988; Pergamon, New York, 1987).

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 177–185.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 201–210.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

Attraction Mechanism of Like-Charged Aerosol Particles in a Moving Conductive Medium

A. V. Gaponov-Grekhov, D. I. Iudin*, and V. Yu. Trakhtengerts**

Institute of Applied Physics, Russian Academy of Sciences, ul. Ul'yanova 46, Nizhnii Novgorod, 603950 Russia
**e-mail: iudin@nirfi.sci-nnov.ru*
***e-mail: vyt@appl.sci-nnov.ru*
Received February 9, 2005

Abstract—We analyze the structure of the space electric charge that appears in the vicinity of a charged dust particle in a moving conductive medium. We show that when the conduction currents play a major role, the screening space charge is concentrated in the form of a thin wake behind the dust particle, while the total Coulomb field forms a dipole structure and serves as an attractive center for other particles with charges of the same sign. We consider the pairing conditions for such particles. Including the polarization contribution from the dust component to the permittivity radically changes the field structure when the dust particle concentration approaches the dissipative instability threshold. In this case, the zone of attraction of like-charged dust particles expands sharply. Estimates suggest that the effects under consideration can govern the formation of regular structures in a moving dusty plasma at fairly high pressures, $P > 0.1$–1 mbar. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

At present, the attraction mechanisms of like-charged aerosol particles in gaseous, partially ionized media are being widely discussed in connection with the problem of the formation of regular structures, in particular, dust plasma crystals in these media [1–6]. In laboratory experiments with dusty plasmas, organized structures appear under widely varied conditions and can be observed in the form of one-dimensional chains and low-dimensional clusters composed of several charged dust particles as well as in the form of fairly extended two-dimensional and three-dimensional lattices with a regular arrangement of dust particles at the lattice sites. This problem is also important for natural media. Dust clouds are typical of cosmic conditions [2]; aerosol streams in partially ionized air are formed in thunderclouds and at mesospheric altitudes in the atmosphere [7–9].

The most popular theory [1, 2, 10] is based on the combined effect of the repulsive Debye–Hückel potential and the dust-particle-confining external electric potential applied to the chamber walls. A crystal under these conditions is formed when the plasma nonideality parameter $\Gamma = q^2/\Delta T$ is larger than a critical value of $\Gamma_{cr} \sim 10^2$, where q, T, and $\Delta \sim N_a^{-1/3}$ are, respectively, the charge, temperature, and characteristic interparticle distance of the dust component, and N_a is its concentration. In the presence of plasma flows, other mechanisms, in particular, the wake [11] and shadow [12] attraction of like-charged particles can also play a significant role.

Fundamentally new phenomena arise in a dusty plasma, given that this system is open. In contrast to an ordinary plasma, the electron–ion component in a complex plasma is absorbed on dust particle surfaces, which requires continuous energy supply to the system [2–5]. This effect can significantly change the formation of a space electric charge around a dust particle. At present, these questions are being actively studied [13].

It should be noted that in the papers cited above, the thermal motion and the Debye screening by the electron and ion charges played a major role in the formation of a space electric charge around a dust particle and, hence, in the formation mechanism of plasma crystals. In general, the conduction currents were disregarded. At the same time, in a fairly dense medium where the ion mean free path is much smaller than the Debye length and is comparable to the dust particle radius, these currents can play a crucial role. A thundercloud is a typical example of such a medium [7, 9]. Similar conditions are realized in laboratory experiments at pressures $P > 1$ mbar (see below).

In this paper, we consider the limiting case of a conductive medium moving relative to the dust particles with a zero Debye length. The physical picture of the interaction in this case is very simple if the hydrodynamic effects are disregarded (the particle is "transparent" for the moving medium): a space charge wake of the opposite sign whose diameter is equal to the particle diameter and whose length is defined by the characteristic scale length

$$l_0 \approx \frac{v_0}{4\pi\sigma}, \tag{1}$$

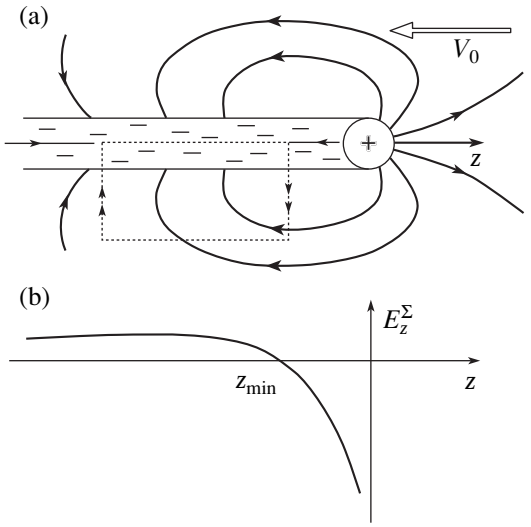

Fig. 1. (a) Image of the particle and the space charge wake; the arrows indicate the path of integration in (7). (b) Total electric field E_z^Σ versus coordinate z.

where v_0 and σ are the velocity and conductivity of the medium, respectively, is formed downstream of the charged dust particle. Basically, this is an analog of the Coulomb screening for the current formation mechanism of a space charge. Together the particle and the wake form a dipole electric field that leads to the "pairing" of like-charged particles under certain conditions. Such a structure of the screening field in a conductive medium was first obtained in [14]. This structure is analyzed in detail in Section 2. In Section 3, we analyze the interaction between "tailed" particles and derive a criterion for their pairing. In Section 4, we consider the collective wake by taking into account the contribution from the dust component to the polarization of the medium. This problem is of interest in connection with the possibility of reaching the dissipative instability threshold [7, 15] as the dust particle concentration increases. The instability is accompanied by the excitation of collective electrostatic modes, with the particle short-range field also changing radically. In Section 5, we summarize our results and discuss the range of their applicability to the conditions of available experiments.

2. THE WAKE IN A CONDUCTIVE MEDIUM

We begin our analysis with the simplest model of the interaction between a spherical charged particle and a conductive medium moving with velocity $\mathbf{v}_0$ whose electric properties are characterized by conductivity σ and permittivity ε. The stability of such a particle in the flow is determined by external forces, including those of a nonelectric origin (the gravity and the drag from the neutral medium), while the constancy of the electric charge is specified by the acting charging mechanism. For a given (unperturbed) motion of the conductive medium, the structures of the electric field **E** and the current **j** are described by the continuity equation for **j** and Ohm's law. In the case of interest, these are

$$\operatorname{div}\mathbf{j} = 0, \quad \mathbf{j} = \sigma\mathbf{E} + \eta\mathbf{v}, \tag{2}$$

where η is the induced space charge in the medium. Given that the flow is one-dimensional (in this case, $\mathbf{v} = -v_0\mathbf{z}_0$, $v_0 > 0$) and using Poisson's equation, we find from (2) that

$$\frac{d\eta}{dz} - k_0\eta = k_0\eta_0(z, \rho), \quad k_0 = \frac{4\pi\sigma}{v_0}, \tag{3}$$

where η_0 is the given charge distribution on the particle, $\mathbf{z}_0$ is a unit vector, and ρ is the transverse coordinate.

We are interested in the case where

$$k_0 a \ll 1, \tag{4}$$

i.e., where the length of the space charge wake k_0^{-1} is much larger than the particle radius a. The pattern of the transverse dependence $\eta_0(z, \rho)$ for the overall structure of the electric field is unimportant, and it may be taken to be uniform. For $\eta_0 \equiv q/s$ at $-a \le z \le 0$ ($s = \pi a^2$ is the cross-sectional area of the wake, and q is the dust particle charge), the resulting space charge distribution is

$$\begin{aligned} \eta &= -\eta_0(e^{k_0 z} - 1), \quad -a \le z \le 0, \\ \eta &= -\eta_0(1 - e^{k_0 a})e^{k_0 z}, \quad z \le -a. \end{aligned} \tag{5}$$

As can be easily verified,

$$s\int_{-\infty}^{0} \eta\, dz = -q.$$

Clearly, the total electric field **E** is axially symmetric (relative to the z axis). The central section of this field containing the z axis is qualitatively shown in Fig. 1. The field strength and the coordinate of the bottom of the potential well, which basically serves as an attractive center for another dust particle with a charge of the same sign, can be determined from the given charge (5) in quadratures. They can be estimated from the following considerations. For a long charged cylindrical rod, the radial field component E_ρ behaves at distances $k_0\rho < 1$ as follows (in CGS units):

$$E_\rho = \frac{2\kappa}{\rho}, \tag{6}$$

where $\kappa = \eta s$ is the linear charge density, and s is the wake cross section. The field component E_z produced

by the space charge on the axis can then be determined by integrating **E** over a closed path (Fig. 1a). We have

$$-2\int_{\delta/k_0}^{a}\frac{\kappa(z)d\rho}{\rho}+E_z\Delta z+2\int_{a}^{\delta/k_0}\frac{\kappa(z+\Delta z)}{\rho}d\rho\approx 0. \quad (7)$$

Given that

$$\kappa(z+\Delta z)\;=\;\kappa(z)+\Delta z s\frac{\partial\eta}{\partial z},$$

and using the corresponding dependence $\eta(z)$ (5), we obtain

$$E_z\;=\;2Ms\eta(z)\;=\;2qMk_0^2\exp(k_0z), \quad (8)$$

where q is the particle charge, $M = |\ln(\delta k_0 a)|$, and δ is the numerical coefficient that characterizes the logarithmic accuracy of our calculations. Running ahead, we note that the rigorous calculation in Section 4 yields $\delta = 0.3$. The total field structure on the wake axis at $z < 0$ is given by the sum

$$E_z^{\Sigma}\approx-\frac{q}{z^2}+2Mk_0^2e^{k_0z}. \quad (9)$$

Dependence (9) is shown in Fig. 1b. The field reaches zero at $k_0z_{\min}\approx -0.6$; the depth of the minimum of the potential well is

$$U_m\approx\frac{q}{z_{\min}}\left(1+\frac{1}{k_0z_{\min}}\right)\approx qk_0.$$

It is easy to verify [14] that the total electric field of the particle and the wake at distances $k_0r \gg 1$ has a dipole structure with the dipole moment $\mathbf{p} = -qk_0^{-1}\mathbf{z}_0$ directed along the velocity vector of the medium $\mathbf{v}_0$.

Thus, the wake of a dust particle in a moving conductive medium differs significantly in scale and shape from the wake of the space charge produced by the Debye screening and has the shape of an elongated (in the direction of motion of the medium) thin rod similar to the valence bond of ions in material. The potential well corresponding to the electric field structure (9) can serve as an attractive center for other dust particles, giving rise to dimeric and, possibly (since a potential well is also present in the common wake, the tail, of the two particles), polymeric structures. To quantitatively answer this question, we must consider the interaction between such particles. This interaction is analyzed in the next section.

3. PAIR INTERACTION OF CHARGED DUST PARTICLES

In the case under consideration, the particle pair interaction is peculiar in that it cannot be reduced to the attraction (or repulsion) between two electric dipoles, since the force acting on the space charge is applied to the medium rather than to the particle. As a result, the problem is reduced to analyzing the relative motion of two charged dust particles with the dipole field of the "tailed" particle acting on each of them. Basically, this allows the electric field acting on a charged particle in plasma to be identified with the mean macroscopic field. In actual conditions, apart from the interaction through the electric field, the dust particles are also acted upon by various nonelectric forces, such as the gravity, the drag from the medium, and the fluctuational forces, which in the long run determine the effective temperature of the dust component. In addition, the particle size dispersion in the field of gravity also leads to dust particle velocity dispersion. All of these forces are rather difficult to take into account, while the principal possibility of particle pairing is entirely determined by the electric force. To ascertain whether this is possible, let us consider the interaction between two identical tailed particles that have a relative velocity $\mathbf{v}^{(0)}$ at the initial time and that are separated by a distance $|\mathbf{r}_0|$, i.e.,

$$t\;=\;0\colon\;(\dot{\mathbf{r}}_2-\dot{\mathbf{r}}_1)_0\;=\;\mathbf{v}^{(0)},\quad(\mathbf{r}_2-\mathbf{r}_1)_0\;=\;\mathbf{r}_0, \quad (10)$$

where $\mathbf{r}_{1(2)}$ is the coordinate of particle 1 (2). At a distance much larger than the wake length where the inequality

$$k_0|\mathbf{r}|\equiv k_0|\mathbf{r}_2-\mathbf{r}_1|\;\gg\;1 \quad (11)$$

holds, the electric field **E** of the tailed particle can be represented as the dipole field

$$\mathbf{E}\;=\;\frac{3\mathbf{n}(\mathbf{p}\cdot\mathbf{n})-\mathbf{p}}{|\mathbf{r}|^3}. \quad (12)$$

Here, $\mathbf{n} = \mathbf{r}/|\mathbf{r}|$, **r** is the radius vector from the location of the dipole to the point of observation, and the dipole moment **p** is

$$\mathbf{p}_{1,2}\;=\;p_0\left(\mathbf{z}_0+\frac{\dot{\mathbf{r}}_{1,2}}{v_0}\right),\quad p_0\approx\frac{q}{k_0}. \quad (13)$$

Formula (13) takes into account the fact that the length of the space charge wake is proportional to the velocity of the medium and that the wake orientation when the particle motion is slowly rearranged ($|\dot{\mathbf{r}}_{1,2}| \ll v_0$) closely follows the total velocity vector relative to the conductive medium.

If we now take into account the fact that $\mathbf{n}_1 \downarrow\uparrow \mathbf{n}_2$, we will then obtain the following equation of motion for the difference vector $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$:

$$\ddot{\mathbf{r}}\;=\;\Delta\frac{3\mathbf{n}(\dot{\mathbf{r}}\cdot\mathbf{n})-\dot{\mathbf{r}}}{r^3}, \quad (14)$$

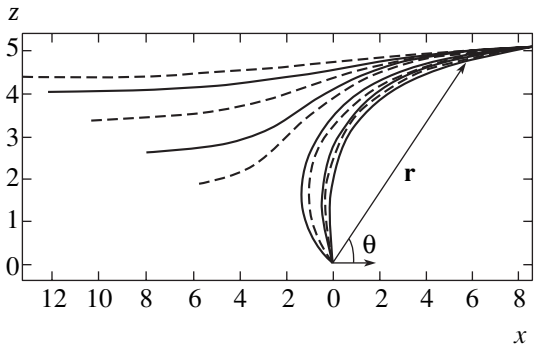

Fig. 2. Examples of the trajectories of relative motion of interacting particles for an equidistant increase in the parameter $\Delta/\gamma d^2$ from 0.2 to 0.8 and at $r_0 = 10$, $\theta_0 = 5\pi/6$, $d = 5$ with (dotted line, $(dk_0)^{-2} = 1$) and without (solid line, $(dk_0)^{-2} = 0$) allowance for the finite depth of the potential well.

where $\mathbf{n} = \mathbf{r}/|\mathbf{r}|$, $|\mathbf{r}| \equiv r$,

$$\Delta = \frac{q^2}{4\pi\sigma m_a}$$

is the interaction parameter, and m_a is the dust particle mass. Equation (14) has a vector integral of motion that can be represented as

$$\dot{\mathbf{r}} + \Delta\frac{\mathbf{r}}{r^3} = \mathbf{v}^{(0)} + \Delta\frac{\mathbf{r}_0}{r_0^3}, \tag{15}$$

where $\mathbf{v}^{(0)}$ and $\mathbf{r}_0$ are the initial relative velocity and radius vector, respectively. Next, let us consider, for simplicity, the two-dimensional motion of particles in a medium with velocity $\mathbf{v}^{(0)}$. In the polar coordinate system on the *xz* plane with the coordinate origin at the point at which one of the particles is located, Eq. (15) transforms to

$$\begin{aligned} r\frac{d\varphi}{dt} &= \beta\cos\varphi - \alpha\sin\varphi \equiv \gamma\sin(\psi - \varphi), \\ \frac{dr}{dt} + \frac{\Delta}{r^2} &= \alpha\cos\varphi + \beta\sin\varphi \equiv \gamma\cos(\psi - \varphi), \end{aligned} \tag{16}$$

where

$$\alpha = v_x^{(0)} + \Delta\frac{x_0}{r_0^3}, \quad \beta = v_z^{(0)} + \Delta\frac{z_0}{r_0^3}, \quad r_0 = \sqrt{x_0^2 + z_0^2},$$

$$\gamma = \sqrt{\alpha^2 + \beta^2}, \quad \psi = \arctan\frac{\beta}{\alpha}.$$

Dividing the second equation in (16) by the first equation yields

$$\frac{dr}{rd\theta} = \frac{\Delta/\gamma r^2 - \cos\theta}{\sin\theta}, \quad \theta = \psi - \varphi. \tag{17}$$

The solution of Eq. (17) that describes the particle trajectory $r(\theta)$ is

$$r\sin\theta = \pm\left[d^2 + \frac{2\Delta}{\gamma}(\cos\theta_0 - \cos\theta)\right]^{1/2}, \tag{18}$$

where $d = r_0\sin\theta_0$ is the impact parameter, which is equal to the minimum distance between the particles as they move in a straight line (when $\Delta = 0$ and $r_0 \longrightarrow \infty$). The fall of the particle to the center, i.e., solution (18) $r = 0$, corresponds to the pairing. Examples of the trajectories corresponding to solution (18) are indicated in Fig. 2 by the solid lines.

In general, the pairing depends on the starting conditions. For fairly large distances $(d/r_0)^2 \ll 1$, we have $\cos\theta_0 \approx -1$ for the colliding particles, and, given the pattern of the trajectories shown in Fig. 2, the pairing condition can be written as

$$\frac{2\Delta}{\gamma d^2} \equiv \frac{q^2}{2\pi\sigma v^{(0)} m_a d^2} > 1, \tag{19}$$

where $v^{(0)}$ is the magnitude of the starting particle velocity; it is assumed that $v^{(0)} \gg \Delta/r_0^2$.

Recall that solution (18) is applicable if inequality (11) holds. If (11) is violated, then the short-range field of the tailed particles must be taken into account, and the depth of the potential well becomes finite. In this case, Eq. (17) takes the form

$$\frac{dr}{rd\theta} = \frac{\dfrac{\Delta}{\gamma(r^2 + k_0^{-2})} - \cos\theta}{\sin\theta}, \tag{20}$$

where the depth of the potential well was taken in accordance with relation (9). The numerical solution of the latter equation is indicated in Fig. 2 by the dotted lines. We see from this figure that the pairing condition becomes slightly more stringent in impact parameter. Other things being equal, the interaction parameter Δ required for pairing increases by about one and a half times.

As we see, the interaction between two like-charged tailed particles differs radically from the interaction between the same particles in the absence of a wake. However, this is true only for distances $k_0 r > 1$. Taking into consideration the electric field structure (9), we conclude that an ordinary Coulomb field of repulsion acts at distances $k_0 r \ll 1$. At the same time, it is not obvious that the field around the particle will not change at close distances either if the contribution from the dust component to the polarization of the medium is taken into account, especially since the medium becomes unstable to electromagnetic disturbances

starting from some critical dust concentration. These questions are analyzed in the next section.

4. THE COLLECTIVE WAKE

As was noted in the Introduction, an increase in the dust particle concentration can fundamentally affect the electric properties of the medium under consideration. In particular, when a certain threshold is exceeded, which, according to [7], is equal to

$$(\Omega_p^2/\nu_a^2)_{\text{thr}} = 1, \tag{21}$$

where

$$\Omega_p = \left(\frac{4\pi q^2 N_a}{m_a}\right)^{1/2}$$

is the plasma frequency of the aerosol particles, and ν_a is the effective collision frequency that determines the force of friction acting on the dust particle from the neutral component—collective electrostatic modes leading to the bunching of dust particles can be excited. It is interesting to trace how the space electric charge is redistributed around the dust particle and how the structure of the electric field found above changes with increasing dust concentration. We restrict ourselves to analyzing this problem in the approximation of a point dust particle. In this approximation, we can use the general method for calculating the electric field produced by a uniformly moving particle in which the potential $\varphi(\mathbf{r}, t)$ is [16]

$$\varphi(\mathbf{r}, t) = \int d^3k \int d\omega \frac{q\delta(\omega - \mathbf{k}\cdot\mathbf{v})}{2\pi^2 k^2 \varepsilon(\omega, \mathbf{k})} \times \exp\{i\mathbf{k}\cdot\mathbf{r} - i\omega t\}, \tag{22}$$

where $\delta(x)$ is the delta function and $\varepsilon(\omega, \mathbf{k})$ is the permittivity.

In the coordinate system with the z axis directed along the particle velocity and given the axial symmetry, formula (22) after integration over ω and the azimuthal angle transforms to

$$\varphi(z, \rho) = \frac{q}{\pi}\int_{-\infty}^{\infty} dk_z \times \int_0^{\infty} \frac{k_\perp dk_\perp J_0(k_\perp \rho)\exp[ik_z(z - vt)]}{(k_\perp^2 + k_z^2)\varepsilon(k_z, k_\perp, \omega = k_z v)}. \tag{23}$$

Here, $J_0(k_\perp\rho)$ is the zero-order Bessel function, and ρ is the distance from the z axis (in the cylindrical coordinate system). Let us first consider a single dust particle. In the ($\xi = z - v_0 t$, ρ) coordinate system, in which the particle is stationary, the flow of conductive gas blows over the particle with velocity v_0 in the $-\xi$ direction. In this case,

$$\varepsilon = 1 - i\frac{k_0}{k_z}, \tag{24}$$

where, as above, $k_0 = 4\pi\sigma/v_0$. The wake is formed at $\xi < 0$, and the potential after integrating (23) over k_z and $k_\perp$ using (24) is ($\xi \le 0$)

$$\varphi(\xi, \rho) = \frac{q}{\sqrt{\rho^2 + \xi^2}} + \pi k_0 q N_0(k_0\rho)\exp(-k_0|\xi|), \tag{25}$$

where $N_0(k_0\rho)$ is the zero-order Neumann function. At $k_0\rho < 1$,

$$N_0(k_0\rho) \approx \frac{2}{\pi} J_0(k_0\rho)\left[\ln\frac{k_0\rho}{2} + C\right],$$

$C = 0.57$ is the Euler constant. As a result, the expression for the electric field $E_z = -\partial\varphi/\partial\xi$ at $k_0\rho = k_0 a \ll 1$ is identical to (9) if we set $z \equiv |\xi| \gg a$ and $\delta = 0.3$ under the logarithm.

In general, given that the dust component is stationary, the expression for $\varepsilon(\mathbf{k}, \omega = k_z v)$ is

$$\varepsilon = \left[1 + \frac{ik_0 v_0}{\omega - k_z v_0} - \frac{\Omega_p^2}{(\omega + i\nu_a)(\omega + ik_z^2 D_a)}\right]_{\omega = k_z v}, \tag{26}$$

where k_0, Ω_p, and ν_a were defined above (see (1) and (21)), $D_a = (\Delta v)^2/\nu_a$, $(\Delta v)^2$ is the velocity spread attributable to the size dispersion of the dust particles falling in the field of gravity $\mathbf{g} \parallel \mathbf{z}$; this spread is assumed to dominate over the thermal scatter; it is also assumed that $\nu_a \gg k_z^2 D_a$. Integral (23) with ε in form (26) describes the electric potential of a test particle moving with velocity v along the z axis. Given dependence (26), expression (23) can be integrated over $k_\perp$. As a result, we obtain

$$\varphi(\tilde\xi, \tilde\rho) = \frac{2q\tilde k_0}{\pi}\int_0^{\infty} \frac{dy K_0(\tilde\rho y)}{\varepsilon_1^2(y) + \varepsilon_2^2(y)} \times \{\varepsilon_1\cos(\tilde\xi y) + \varepsilon_2\sin(\tilde\xi y)\}, \tag{27}$$

where $K_0(\tilde\rho y)$ is the zero-order Macdonald function, $\tilde k_0 = 4\pi\sigma/(v_0 - v)$, $\tilde\rho = \tilde k_0\rho$, $\tilde\xi = \tilde k_0\xi$, and the functions

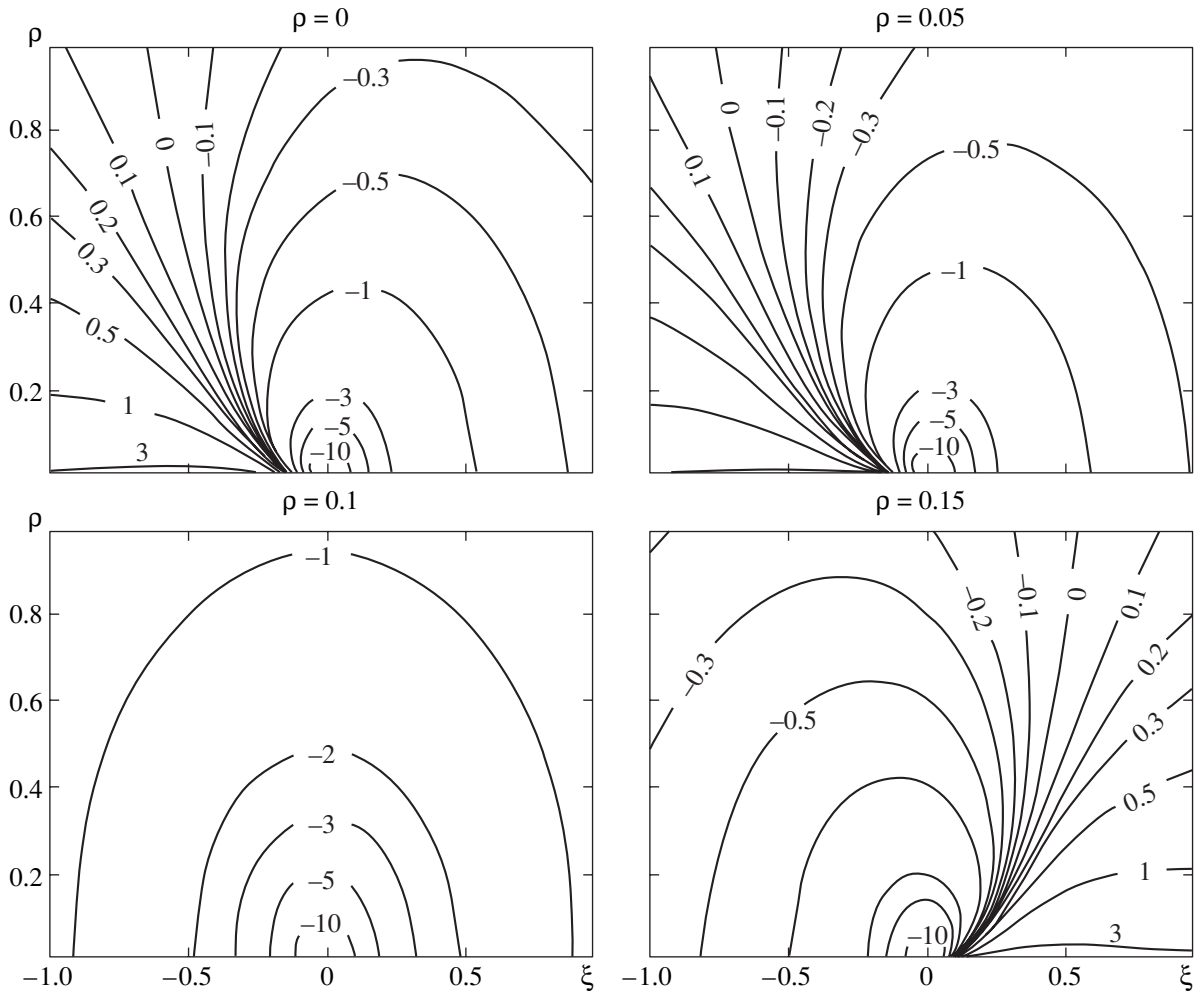

Fig. 3. Change in the structure of the potential as the dust concentration increases at a fixed velocity of the test particle ($c = 0.01$).

$\varepsilon_1 = \mathrm{Re}\varepsilon$ and $\varepsilon_2 = \mathrm{Im}\varepsilon$ are given by

$$\varepsilon_1 = 1 - \frac{p^2(c^2 - c_T^2)}{(1 + c^2y^2)(c^2 + c_T^4y^2)}, \qquad (28)$$

$$\varepsilon_2 = \frac{1}{y}\left\{\frac{p^2c(1 + c_T^2y^2)}{(1 + c^2y^2)(c^2 + c_T^4y^2)} - 1\right\}. \qquad (29)$$

The dimensionless coefficients p, c, and c_T determining the structure of the field φ are

$$p = \frac{\Omega_p}{\nu_a}, \quad c = \frac{\tilde{k}_0 v}{\nu_a}, \quad c_T = \frac{\tilde{k}_0\sqrt{\Delta v^2}}{\nu_a}. \qquad (30)$$

Expression (25) is derived from (27) at $p = 0$. It is easy to see that the dispersion relation $\varepsilon(\omega, \mathbf{k}) = 0$ at $p > 1$ contains solutions increasing with time ($\mathrm{Im}\omega > 0$). Basically, this implies that solution (27), which describes the stationary structure of the field φ, is valid only at $p \lesssim 1$ and is inapplicable at $p > 1$. For $p > 1$, the problem with initial conditions must be solved. Given the aforesaid, let us consider the changes in the structure of the field $\varphi(\tilde{\xi}, \tilde{\rho})$ when p increases from zero to unity while remaining in the domain of stable solutions.

Given the polarization contribution of dust particles to $\varepsilon(\omega = k_z v, \mathbf{k}) = 0$, the three additional parameters (30) appear in the problem. Let us first consider how the structure of the potential $\varphi(\tilde{\xi}, \tilde{\rho})$ changes in the absence of a velocity spread, $c_T = 0$. Here, the velocity of the test particle c plays a significant role. Figure 3 shows the change in the spatial structure of $\varphi(\tilde{\xi}, \tilde{\rho})$ with increasing p.

An interesting feature arises at $p^2 \approx c$, when the dipole structure disappears ($p^2 = c = 0.01$) and then reappears ($p > 0.1$), but with the opposite orientation of the dipole, the screening space charge in front of the dust particle. The equality $p^2 = c$ here is not coincidental, but corresponds to a certain resonance at which ε_2

changes its sign when passing through zero. This resonance is discussed in more detail below. The physical interpretation of this potential rearrangement is fairly simple: at a low relative velocity, the dust particles are deflected from the z axis by the electric field of the test particle without reaching it; the space charge tails of these particles turn inward, producing a charge excess in front of the particle. Since the dust component at $c < 0$ strikes the test particle from the tail, the dipole does not change its orientation, and the depth of the potential well increases due to the tails of the overtaking dust particles. In general, with the exception of the above effect, the dipole structure of the tailed particle qualitatively changes only slightly with p, as illustrated by Fig. 4.

The depth of the potential well slightly increases with p, because $|\varepsilon|$ decreases; at fixed c, the deepest minimum of the potential well is reached at a certain value of $p^2 = c$. The coordinate $\xi_{\min}$ of the potential minimum slightly approaches the coordinate of the test particle with increasing p. The dust particle velocity dispersion, $c_T \neq 0$, decreases the depth of the potential well and can cause it to disappear at fairly large values of c_T and p (Fig. 5).

The largest changes in the structure of the Coulomb potential arise as the threshold of dissipative instability is approached; at $c_T = 0$, the latter is described by the dispersion relation[1]

$$\epsilon(\omega, \mathbf{k}) \equiv 1 - \frac{\Omega_p^2}{\omega(\omega + i\nu_a)} + \frac{4\pi i\sigma}{\omega - \mathbf{k}\cdot\mathbf{v}_0} = 0. \qquad (31)$$

According to (31), the instability threshold $\mathrm{Im}\omega = 0$ is reached for $\omega \longrightarrow 0$ and is equal to $p_{\mathrm{thr}} = 1$ (21); near the threshold, the dispersion relation $\omega(\mathbf{k})$ is

$$\omega = k_z v_0 \frac{1}{1 + 4\pi\sigma\nu_a/\Omega_p^2}. \qquad (32)$$

The forced solution of interest described by (27) at $\omega = k_z v$ is applicable in the stable domain $p < 1$. Clearly, the largest change in potential φ as the instability threshold is approached might be expected for $|\varepsilon| \longrightarrow 0$, i.e., for the particle that is in phase with wave (32) when

$$p \longrightarrow 1, \quad c \approx p^2, \quad y \longrightarrow 0. \qquad (33)$$

The calculations of φ (27) confirm the above reasoning. Figure 6 shows the structure of the potential $\varphi(\tilde{\xi}, \tilde{\rho})$ near the threshold $p \lesssim 1$ for the resonant ($p^2 = c$) and nonresonant ($p = c$) cases. We see a radical difference in the behavior of the potential both near and far from

[1] It should be noted that, strictly speaking, the dispersion relation (31) is valid on scales much larger than the interparticle distance. On the other hand, if many dimeric (polymeric) structures appeared in the medium, then the instability can be modified significantly.

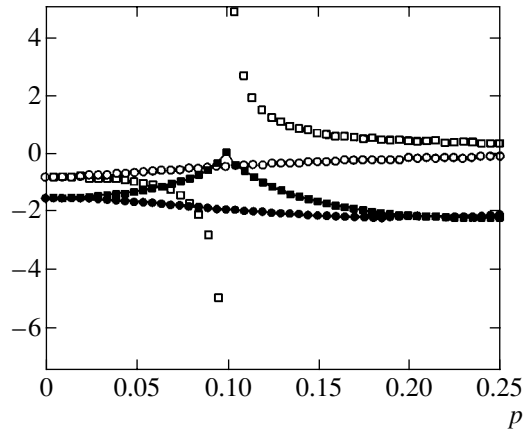

Fig. 4. Evolution of the position (open circles) and value (solid line and filled circles) of the potential minimum as the dust concentration p increases at a fixed test particle velocity: $c = 0.01$ (squares) and -0.01 (circles).

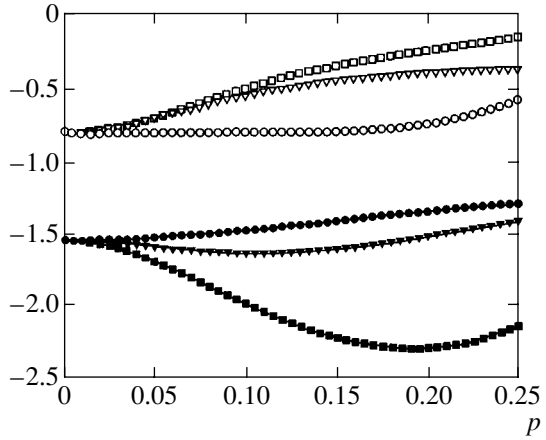

Fig. 5. Evolution of the position (open circles) and value (solid line and filled circles) of the potential minimum as the dust concentration p increases at $c = -0.01$ and for various dust particle velocity dispersions: $c_T = 0$ (squares), 0.05 (triangles), 0.5 (circles).

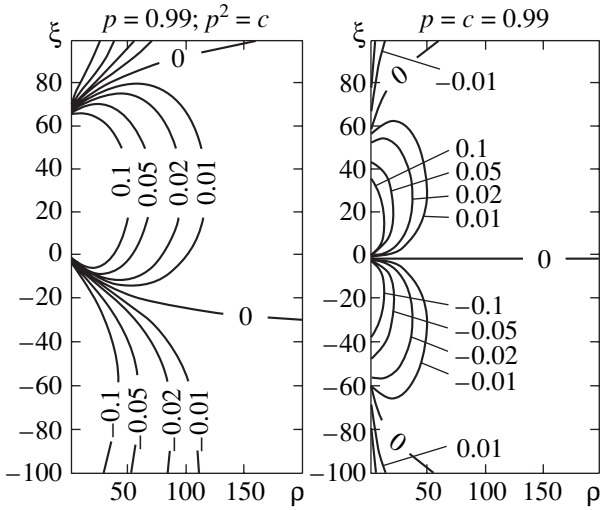

Fig. 6. Change in the structure of the potential near the dissipative instability threshold.

the particle. For $p^2 = c \longrightarrow 1$, the following analytical expression for φ and the electric field **E** can be derived from (27):

$$\varphi(\xi, \rho) = \frac{q}{r} \mp qk_0 \ln\frac{r + |\xi|}{\rho}, \quad (34)$$

where $r = (\rho^2 + \xi^2)^{1/2}$; the electric field has the components

$$E_\xi = \frac{q}{r}\left(k_0 + \frac{\xi}{r^2}\right), \quad E_\rho = \frac{q}{r}\left(\frac{\rho}{r^2} - \frac{k_0\xi}{\rho}\right). \quad (35)$$

The structure of potential (34) is determined not only by the presence of a charged particle at the coordinate origin and a wake of the opposite sign, but also by the charge concentrated on the axis ($\rho = 0$); the latter extends from $-\infty$ to ∞ as the $p = 1$ threshold is approached and under the resonance conditions $p^2 = c = 1$, being positive at $\xi < 0$ and negative at $\xi > 0$. In contrast to the case of a single particle, the electric field amplitude increases sharply, becoming infinite on the axis, which basically corresponds to the resonance $\varepsilon \longrightarrow 0$.

Qualitative changes in the potential structure might be expected during the passage through the instability threshold, when $p > 1$. In this case, ε_1 in (27) changes its sign (at small y) and the Coulomb field of repulsion of like-charged particles can transform into the field of attraction even at small distances between them (similar to the case of $\varepsilon < 0$ in a medium without dispersion). However, it should be borne in mind that the above analysis is generally invalid at $p > 1$, and the excitation of electric eigenmodes of the medium must be taken into account along with the forced solution (27). This problem is beyond the scope of this paper.

5. DISCUSSION AND CONCLUSIONS

The effects considered above relate to the limiting case of a partially ionized plasma, where the conduction currents play a major role in forming the space charge, while the effects of thermal motion can be neglected. If the formation of the space charge is determined by ions, then, as follows from the general expression for ε with the inclusion of thermal motion, this neglect is possible if the ion mean free path

$$\frac{v_{Ti}}{\nu_{in}} = l_{in} \ll r_{Di} = \frac{v_{Ti}}{\omega_{pi}}$$

and, in addition, if the length k_0^{-1} of the space charge wake is larger than the diffusion length, i.e.,

$$\nu_{in} \gg \omega_{pi}, \quad \frac{k_0 v_{Ti}^2}{\nu_{in} v_0} \ll 1, \quad (36)$$

where ν_{in} is the ion–neutral collision frequency, ω_{pi} is the plasma ion frequency, v_{Ti} is the thermal ion velocity, $k_0 = 4\pi\sigma_i/v_0$, v_0 is the directed relative velocity of the ions and dust particles, and the conductivity

$$\sigma = \frac{\omega_{pi}^2}{4\pi\nu_{in}}.$$

If the electrons make a contribution comparable to that of the ions, then the speed of sound $v_s \approx T_e/m_i$ must be substituted for v_{Ti} in relation (34).

Inequalities (36) in thundercloud conditions hold with a large margin. Different situations are encountered in laboratory experiments. At a typical degree of ionization ~10^{-6}–10^{-5}, the first inequality in (36) holds at fairly high pressures, $P \gtrsim 0.1$–1 mbar. The second inequality in (36) holds at a fairly high drift velocity v_0. If this velocity is attributable to the electric field, then, given the relation $j = eN_i v_0 = \sigma E$, the second inequality in (36) transforms to

$$\frac{4\pi N_i T_i}{E^2} \ll 1, \quad T_i = m_i v_{T_i}^2. \quad (37)$$

At typical values of $N_i \sim 10^9$ cm^{-3} and $T_i \sim 300$ K, inequality (37) holds at $E > 7$ V cm^{-1}. Such electric fields definitely take place in rarefied-plasma experiments where dust crystal structures are observed [5, 6]. Let us consider how the current screening mechanism and the dust particle pairing conditions are realized in these experiments. Choosing $P \approx 1$ mbar, $N_i \sim 10^9$ cm^{-3}, and $E \sim 10$ V cm^{-1} as typical values, we find for an argon plasma: $v_0/v_{Ti} \sim 1.5 v_{Ti} \sim 3 \times 10^4$ cm s^{-1}, $\nu_{in} \sim 3 \times 10^7$ s^{-1}, and $\omega_{pi} \approx 10^7$ s^{-1}. Inequalities (36) hold at these parameters. The characteristic scale length of the wake is $k_0^{-1} \sim 10^2\mu$. The pairing criterion (19) at the validity boundary $k_0 d \sim 1$ can be written as

$$\frac{8\pi q^2 \sigma}{v_0^2 m_a \nu^{(0)}} > 1.$$

Assuming that $a \approx 10\mu$, $v_0 \approx 3 \times 10^4$ cm s^{-1}, $q \approx 3 \times 10^4 e$, and $4\pi\sigma \approx 3 \times 10^6$ s^{-1}, we find that the capture criterion is satisfied for a relative dust particle velocity $v_{\tau 0}^{(0)} \lesssim 10^{-3}$ cm s^{-1}. Note that the thermal velocity of dust particles with $a \sim 10\mu$ ($m_a \sim 10^{-9}$ g) at $T \approx 300$ K is $v_{T_a} \approx 6 \times 10^{-3}$ cm s^{-1}. A rise in the dust particle concentration causes an increase in the depth of the potential well in the particle wake and, thus, an increase in the pairing probability. Qualitative changes in the structure of the Coulomb potential might be expected as the dissipative instability threshold dependent on the dust particle concentration (see (27)) is approached, when the capture zone can expand sharply. However, certain

conclusions can be reached here by considering the nonstationary and, in general, nonlinear problem of the formation of an electric field in a complex plasma.[2]

Summarizing our results, we can suggest the following scenario for the formation of structures in a moving conductive medium in the presence of dust charged particles. When inequalities (36) hold, the space charge screening the dust particle is localized in a thin cylindrical wake with the diameter equal to the particle diameter and the length

$$l \sim k_0^{-1} = \frac{v_0}{4\pi\sigma}$$

determined by the velocity v_0 of the flow and the conductivity σ of the medium. The dipole structure of the potential of such a tailed particle is an attractive center for other dust particles. This pairing process has no threshold, but its probability depends on the impact parameter of the interacting dust particles and on their electric charge and increases with dust particle concentration. The electric field structure admits of the principal possibility of the formation of polymeric chains along with dimers, but this question is beyond the scope of this paper. The formation of three-dimensional crystal structures, particularly in the absence of a plasma trap in the transverse (relative to the drift) direction, is still an open question. Here, the reaching of the dissipative instability threshold, which gives rise to a periodic electric potential in the transverse direction [15], may play an important role. The spread in dust particle velocities may prove to be critical for the formation of polymeric structures. In this case, the structural transition in the medium under consideration is possible only when the dissipative instability threshold is reached.

[2] Note that a radical rearrangement of the Coulomb potential accompanied by the attraction effects of like-charged particles when the instability threshold is reached might also be expected in the case of a collisionless plasma. This may prove to be fundamentally important at the nonlinear instability growth stage, leading to the clustering of phase space and to the sticking of interacting particles into bunches.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 04-02-17405).

REFERENCES

1. A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Usp. Fiz. Nauk **167**, 1215 (1997) [Phys. Usp. **40**, 1163 (1997)].
2. V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys. Usp. **40**, 53 (1997)].
3. V. N. Tsytovich, G. E. Morfill, and H. Thomas, Fiz. Plazmy **28**, 675 (2002) [Plasma Phys. Rep. **28**, 623 (2002)].
4. G. E. Morfill, V. N. Tsytovich, and H. Thomas, Fiz. Plazmy **29**, 3 (2003) [Plasma Phys. Rep. **29**, 1 (2003)].
5. H. Thomas, G. E. Morfill, and V. N. Tsytovich, Fiz. Plazmy **29**, 963 (2003) [Plasma Phys. Rep. **29**, 895 (2003)].
6. V. E. Fortov, A. G. Khrapak, and I. T. Yakubov, *Physics of Nonideal Plasma* (Fizmatlit, Moscow, 2004) [in Russian].
7. V. Yu. Trakhtengerts, Dokl. Akad. Nauk SSSR **308**, 584 (1989).
8. V. Y. Trakhtengerts, J. Atmos. Terr. Phys. **56**, 373 (1994).
9. E. A. Mareev, A. E. Sorokin, and V. Yu. Trakhtengerts, Fiz. Plazmy **25**, 289 (1999) [Plasma Phys. Rep. **25**, 261 (1999)].
10. H. Ikezi, Phys. Fluids **29**, 1764 (1986).
11. S. V. Vladimirov and M. Nambu, Phys. Rev. E **52**, 2172 (1995).
12. A. M. Ignatov, Fiz. Plazmy **22**, 648 (1996) [Plasma Phys. Rep. **22**, 585 (1996)].
13. V. N. Tsytovich, Pis'ma Zh. Éksp. Teor. Fiz. **78**, 1283 (2003) [JETP Lett. **78**, 763 (2003)].
14. A. V. Gaponov-Grekhov, I. S. Dolina, B. E. Nemtsov, and L. A. Ostrovskiĭ, Zh. Éksp. Teor. Fiz. **102**, 243 (1992) [Sov. Phys. JETP **75**, 128 (1992)].
15. A. V. Gaponov-Grekhov and V. Yu. Trakhtengerts, Pis'ma Zh. Éksp. Teor. Fiz. **80**, 814 (2004) [JETP Lett. **80**, 687 (2004)].
16. V. D. Shafranov, in *Reviews of Plasma Physics*, Ed. by M. A. Leontovich (Gosatomizdat, Moscow, 1963; Consultants Bureau, New York, 1967), Vol. 3, p. 3.

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 186–195.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 211–223.

STATISTICAL, NONLINEAR, AND SOFT MATTER PHYSICS

The Tube Character of Electron Drift in Condensed Inert Gases¶

E. B. Gordon[a] and B. M. Smirnov[b]

[a]*Institute for Problems in Chemical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

[b]*Institute for High Temperatures, Russian Academy of Sciences, Moscow, 127412 Russia*

e-mail: gordon@binep.ac.ru, bsmirnov@orc.ru

Received February 11, 2005

Abstract—The behavior of an excess electron in condensed inert gases in an external electric field is considered at densities and temperatures at which the mobility of a slow electron is relatively high. On the basis of experimental data and a model of a pair electron interaction with atoms, an effective potential energy surface is constructed for an excess electron inside a dense inert gas. The region available for a slow electron consists of many intersecting channels that form a Delaunay network located between atoms. A drifting electron, as a quantum object, propagates along these channels (tubes), and electron transition between intersecting potential energy tubes of different directions provides an effective electron scattering. This mechanism of electron drift and scattering differs from that in gases and crystals. Peculiarities of electron drift inside dense inert gases are analyzed within the framework of this mechanism of electron scattering, leading to a moderate change of the electron mobility upon melting. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The reduced electron mobility of excess electrons in heavy condensed inert gases (Ar, Kr, Xe) as a function of the atom number density has a sharp maximum at moderate number densities, as it follows from experiments [1–14]. In particular, for xenon, the maximum zero-field reduced mobility exceeds that for a gaseous state by more than three orders of magnitude [6, 10]. Moreover, the maximum reduced zero-field mobility of excess electrons in inert gases exceeds that for coin metals by one order of magnitude [15].

The simplest theoretical models [16–20] consider the drift of an excess electron as a result of pair electron–atom scattering and explain high electron mobility by the Ramsauer effect in electron scattering on individual atoms and negative electron–atom scattering lengths for Ar, Kr, and Xe. But the approach of independent atoms is correct only for gases; at atomic densities corresponding to the maximum of the electron mobility, the distance between nearest-neighbors is comparable with the electron–atom scattering length, and, hence, this approach is not correct and may be considered rather as a model. The effect of high electron mobility at moderate atomic densities results from the collective character of the interaction of an excess electron with atoms.

Recently [21], we showed that the reason of the electron mobility maximum is related to the transition from attraction to repulsion for an excess electron inside an inert gas as the number density of atoms increases starting from the gas density. Indeed, in gases, where an excess electron interacts with each atom independently, the average electron potential energy is negative because of a negative electron–atom scattering length, which leads to an attractive Fermi exchange interaction of the electron with each atom. At high atomic densities, when the distance between neighboring atoms is comparable with the atom size, the average interaction potential for the electron corresponds to repulsion because of the Pauli exclusion principle. Hence, at moderate atom densities, the average interaction potential of an excess electron with inert gas atoms becomes zero.

Therefore, there is a range of inert gas densities with an attractive interaction potential for an excess electron, and these densities correspond to a high electron zero-field mobility, as is observed experimentally. Because penetration of an excess electron into each core leads to repulsion due to the Pauli exclusion principle, points of the maximum attractive potential for an excess electron form a Delaunay network [22–24], whose lines are located between atoms and may be found on the basis of the Voronoi–Delaunay method [25]. This method consists in construction of planes located at identical distances from neighboring atoms. Intersections of these planes form the Delaunay network, and the electric potential has minima on this network. The equipotential surfaces that are close in energy form

¶ This article was submitted by the authors in English.

tubes, which are almost straight. The electron, being a quantum object, can propagate inside an inert gas along these almost straight channels. Electron transitions between channels of different directions in regions of their intersections lead to an effective electron scattering.

We thus conclude that a specific mechanism of the electron drift in this case differs from those in both gases and crystals. Indeed, propagating in gases, a test electron is scattered on individual atoms, whereas scattering of an electron wave in crystals is determined by distortion of the crystal lattice as a result of a shift of atom positions from the equilibrium ones during motion of the electron wave. Then, melting of a solid should seemingly lead to a strong change of the electron mobility. However, in the case under consideration, a slow electron propagates along an individual tube near the Delaunay network, the electron scattering is weak, and, therefore, the electron mean free path inside an inert gas with optimal parameters is large compared to the distance between nearest-neighbors. Correspondingly, change of the phase state does not lead to a significant change of the mobility for a slow electron in an inert gas.

Below, we consider the tube mechanism of electron drift in heavy inert gases and analyze various aspects of the electron drift under conditions of the tube-shape potential of a self-consistent field for an excess electron.

2. ELECTRIC POTENTIAL FOR AN EXCESS ELECTRON IN DENSE INERT GASES

The negative spatial charge created by excess electrons in dense inert gases can result in strong electric fields even at low electron number densities. Therefore, we consider the regime of electron drift in an inert gas neglecting the interaction between individual electrons; that is, an individual electron is considered drifting in an inert gas. We consider peculiarities of the potential energy surface (PES) for an excess electron in an inert gas and, correspondingly, the character of the electron drift in condensed inert gases under the action of an external electric field. Using the analogy with clusters consisting of many atoms with a pair interaction between them [26–28], we represent the PES as a sum of local minima and saddles. At atomic densities, when the electron mobility is high, an excess electron passes over barriers of the PES during its drift in inert gases.

Another peculiarity of the PES at optimal atomic number densities is a large volume inside condensed inert gases where the electron location is prohibited by the Pauli exclusion principle. Indeed, a slow electron cannot penetrate inside an atom where valence atomic electrons are located, and, hence, the excluded region for an excess electron is concentrated near atomic cores. For simplicity, we take the prohibited volume near each atom in the form

$$V_* = \frac{4\pi}{3} r^3, \tag{1}$$

where r is the effective atom radius, which depends on the electron energy ε. We take it from the relation

$$U(2r) = \varepsilon, \tag{2}$$

where $U(R)$ is the interaction potential of two atoms at a distance R between them. In this way, we changed the repulsion of a free electron from the atom core by that of a bound electron. Table 1 lists the values of the atom radii for an exchange electron–atom interaction calculated on the basis of the above formula. This volume is compared in Table 1 with the volumes per atom for the solid V_{sol} and liquid V_{liq} phase states at the triple point, and also with the volume per atom V_{max} at the atomic number density that corresponds to the maximum of the electron mobility. These ratios are given for the electron energy $\varepsilon = 0.1$ eV and for the electron energy $\varepsilon = 1$ eV in parentheses. We can see that the prohibited volume for a free electron at low electron energies may occupy a significant part of the total volume.

Table 1. Parameters of the repulsive interaction potential for an excess electron with individual atoms of inert gases

	Ar	Kr	Xe
$\bar{r}$, a_0	1.663	1.952	2.338
$\overline{r^2}$, a_0^2	3.311	4.455	6.277
r, Å, $\varepsilon = 0.1$ eV	1.63	1.72	1.92
r, Å, $\varepsilon = 1$ eV	1.23	1.28	1.30
V_{sol}, cm^3/mol	24.6	29.6	37.1
V_{liq}, cm^3/mol	28.2	34.3	42.7
V_{max}, cm^3/mol	50.2	43.0	50.2
V_*/V_{sol}	0.44(0.19)	0.44(0.18)	0.48(0.15)
V_*/V_{liq}	0.39(0.17)	0.38(0.15)	0.42(0.13)
V_*/V_{max}	0.22(0.10)	0.30(0.12)	0.35(0.11)

We note that this character of the exchange interaction between an excess electron and valence electrons of atoms of condensed inert gases is preserved up to high atomic densities until electron shells of neighboring atoms overlap significantly. In any case, it is valid at densities related to the solid and liquid aggregate states of inert gases, and the average exchange interac-

Table 2. Parameters of the potential energy for an excess electron inside inert gases

	Ar	Kr	Xe
L, a_0	–1.5	–3.1	–5.7
U_{min}, eV [8, 11, 12, 14]	–0.33	–0.53	–0.77
N_{min}, 10^{22} cm^{-3}	1.1	1.2	1.1
a_{min}, Å	4.8	4.7	4.8
r_{min}, Å	2.8	2.7	2.8
$2\pi\hbar^2 LN_{min}/m_e$, eV	0.41	0.94	1.58
C, eV	0.44	0.71	1.04
α	4	4	4
A, eV	6	10	14
R_{min}, Å	3.6	3.5	3.6
C', eV	0.15	0.25	0.36
A', eV	2.2	3.5	6.5

tion potential at a given atomic density due to this interaction can be approximated by the formula

$$U_{ex} = A\exp\left(-\alpha\frac{V_*}{V}\right), \quad (3)$$

where V is the volume per atom and A and α are parameters.

We now construct the difference between the potential for an excess electron located inside a condensed inert gas and in a vacuum. Taking the electron potential in a vacuum to be zero, we vary the atomic density from low values, when this system of atoms is a gas, up to moderate ones, at which the mobility of an excess electron is of interest. At low atomic densities, an excess electron interacts with individual atoms independently. In regions between atoms and far from them, the interaction potential is zero, and nonzero interaction takes place only near the atoms. On the basis of the Fermi formula [29, 30], the interaction potential between an electron and atoms can be represented as

$$U(\mathbf{r}) = \sum_i \frac{2\pi\hbar^2}{m_e} L\delta(\mathbf{r}-\mathbf{R}_i), \quad (4)$$

where $\hbar$ is the Planck constant, m_e is the electron mass, $\mathbf{r}$ is the electron coordinate, $\mathbf{R}_i$ is the coordinate of the ith atom, and L is the electron–atom scattering length. Because the scattering length L is negative for Ar, Kr, and Xe (see Table 2), this interaction potential corresponds to attraction in the regions close to atoms. Therefore, the potential energy surface consists of regions inside atoms with a sharp electron repulsion, regions near each atom with electron attraction, and regions between atoms with zero interaction potential. The attraction corresponds only to an average interaction of an electron of zero energy with an individual atom in a gas, and, according to formula (4), the average interaction potential of an electron with inert gas atoms is

$$\bar{U}_{at} = \frac{2\pi\hbar^2}{m_e} LN, \quad (5)$$

where N is the atom number density. This interaction leads to a redshift of spectral lines emitted by excited atoms located in inert gases [31]. Because this shift of spectral lines is determined mostly by the exchange electron–atom interaction, and a long-range interaction, including the polarization ion–atom interaction, gives a small contribution to this shift, we account below for the exchange part of the interaction only.

The exchange interaction of a test electron with electrons of an internal atom region corresponds to repulsion of this electron, and we describe it by formula (3). Adding the attractive exchange interaction potential (5) to it, we represent the total electron potential in the form

$$\bar{U}(N) = -C\frac{N}{N_{min}} + A\exp\left(-\alpha\frac{N_{min}}{N}\right), \quad (6)$$

where N is the current number density of atoms and N_{min} is the number density at which the interaction potential has a minimum. The values of N_{min} together with a_{min}, the distances between nearest-neighboring atoms at this density, are given in Table 2.

If formula (6) is valid for a gas, where the second term is zero, the parameter C is equal to

$$C = -\frac{2\pi L\hbar^2}{m_e} N_{min}. \quad (7)$$

In reality, we are based on the experimental dependence $\bar{U}(N)$ that gives another value of C. Indeed, on the basis of experimental data, which are approximated by formula (6), we find the parameters in Eq. (6) in accordance with the formula

$$C = -\frac{d\bar{U}}{dx}(x=0), \quad \alpha = \frac{C}{C-|U_{min}|}, \quad A = (C-|U_{min}|)e^{\alpha}, \quad x = N/N_{min}. \quad (8)$$

Here, U_{min} is the minimum of the electron potential inside an inert gas (the electron potential in a vacuum is zero). Experimental parameters for $U(x)$ together with the parameters in formula (6) are given in Table 2. Figure 1 represents experimental data for the average electron potential energy in xenon.

Based on the experimental data for the electric potential of a condensed inert gas with respect to an excess electron, we construct the potential energy surface for an excess electron inside an inert gas. We rewrite formula (6) for the minimal electron energy of an excess electron as

$$\bar{U}(r_W) = -C\left(\frac{r_{\min}}{r_W}\right)^3 + A\exp\left(-\alpha\frac{r_W^3}{r_{\min}^3}\right), \qquad (9)$$

where

$$r_W = \left(\frac{4\pi N}{3}\right)^{-1/3}$$

is the Wigner–Seits radius and $r_{\min}$ is the radius at the atom number density corresponding to the maximum attraction of an electron inside the inert gas. Formula (9) can be rewritten in terms of the distance a between the nearest-neighbors,

$$\bar{U}(a) = -C\left(\frac{a_{\min}}{a}\right)^3 + A\exp\left(-\alpha\frac{a^3}{a_{\min}^3}\right), \qquad (10)$$

where $a_{\min}$ is the distance between nearest-neighbors at which the electron potential inside an inert gas has the minimum. We note that the atom number densities corresponding to the minimum of the electron potential according to formulas (9) and (10) are equal to $N_{\min}$, the minimal electron potentials in formulas (9) and (10) coincide with $U_{\min}$, and these parameters follow from formula (6). In addition, we assume a classical character of the electron interaction inside an inert gas in this consideration, although the interaction has a quantum character in reality.

3. DELAUNAY NETWORK FOR THE INTERACTION OF AN EXCESS ELECTRON INSIDE AN INERT GAS

Our goal is to construct the potential energy surface for an excess electron inside an inert gas in the range of the atom number densities and temperatures providing an attractive electric potential there. We concentrate on the simplest case where atoms form a crystal lattice and find electron positions with the minimum potential energy. Evidently, because of the repulsive interaction for an excess electron with atom interiors, the points of the minimum electron potential are located equidistantly from the nearest nuclei. For two nearest planes of the crystal lattice, we then draw the Voronoi surfaces between each pair of nearest-neighbors, such that these surfaces separate the action of individual atoms on an electron. Each Voronoi plane is located at identical distances from two nearest atoms, and intersections of the Voronoi surfaces with the two considered planes of atoms are shown in Fig. 2, where they form a net of regular hexagons whose centers are the nuclei of the lattice. Evidently, from the symmetry considerations, the optimal positions of an excess electron with minimal values of the electron potential energy are located in the plane in the middle between the nearest planes of atoms considered. Intersections of the Voronoi surface with this plane form straight lines of three directions (the solid lines in Fig. 2).

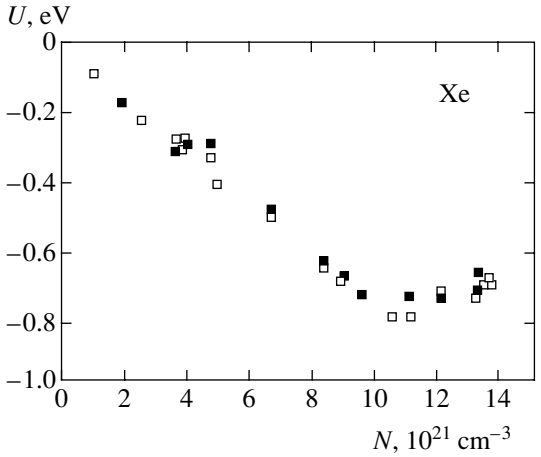

Fig. 1. The potential energy of an excess electron moving in xenon in an external electric field with respect to the vacuum vs. on the number density of atoms according to experiment [8] (symbols).

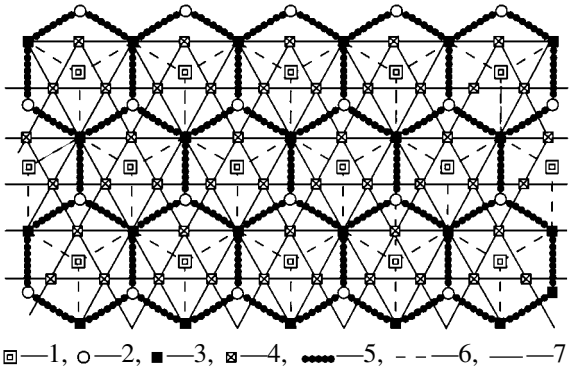

Fig. 2. The character of the behavior of an excess electron between two planes of the crystal lattice of inert gases: (*1*) positions of atoms of the first layer, (*2*) positions of atoms of the second layer, (*3*) vertices of the pentagons that are intersections of the Voronoi surface with the corresponding layer, (*4*) positions of the Voronoi surface for an excess electron in the middle plane between these layers with the strongest interaction between the electron and atoms, (*5*, *6*) hexagons that are intersections of the Voronoi surface with the corresponding layers, (*7*) directions of the electron current if it is located in the middle plane.

Evidently, the electron potential energy is minimal on these lines forming the Delaunay network [22–24]. We note that the Delaunay network is an important mathematical concept (see, e.g., [32–34]). We here use

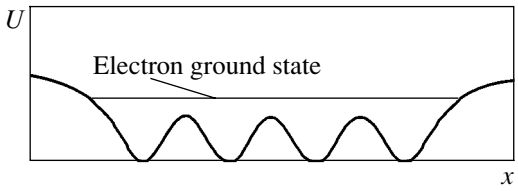

Fig. 3. The form of the potential energy for an excess slow electron in a condensed inert gas along lines of the Delaunay network.

only the applied aspect of this problem related to the construction of lines of the minimum or maximum potential (see, e.g., [25]). Electron drift inside an inert gas proceeds near these lines. We assume that intersection points of these lines, i.e., sites of the Delaunay network, are characterized by minima of the electron potential energy, and their values are identical for all the intersection points (values *4* in Fig. 2) because of the symmetry. Passing to three-dimensional space, we obtain intersections of six straight lines at points whose distance from two nearest-neighbors is $a/2$, where a is the distance between nearest-neighbors of the lattice.

Thus, assuming the optimal distance of an excess electron from nearest nuclei at the optimal number densities of atoms to be maximum for the minimum electron potential energy, we obtain the optimal electron positions for the close-packed crystal lattice to be located on the Delaunay network that consists of intersecting straight lines. We have two types of these lines, which are alternated, and the period of translation symmetry is a for the first-type lines and $a/2$ for the second-type lines. In Table 3, we give the distances from six nearest-neighbors for points that correspond to the minima of the electron potential energy or are located in the middle between nearest such points. The number of nuclei with an indicated distance from a given point of the Delaunay network is given in parentheses.

Table 3. Distances between an excess electron located in minima and maxima of the Delaunay network and six nearest nuclei in the case where atoms form a close-packed crystal lattice. Numbers in parentheses are the corresponding numbers of nearest neighbors

	Points 4 in Fig. 2	In the middle between points 4 in Fig. 2
Lines of the first type	$\frac{a}{2}$ (2), $a\frac{\sqrt{3}}{2}$ (4)	$\frac{a}{\sqrt{2}}$ (6)
Lines of the second type	$\frac{a}{2}$ (2), $a\frac{\sqrt{3}}{2}$ (4)	$\frac{a\sqrt{3}}{4}$ (1), $a\frac{\sqrt{7}}{4}$ (2), $a\frac{\sqrt{11}}{4}$ (1), $a\frac{\sqrt{15}}{4}$ (2)

In the liquid aggregate state, the Voronoi surfaces and Delaunay network may be constructed in the same manner, but the Delaunay network lines become curved. Nevertheless, because of a short order in liquids, the curvature of these lines is not large, and we can take the crystal case as a basis for a qualitative consideration. In any case, the number of lines and the character of their intersection is identical in both cases. Supposing that positions on the Delaunay network correspond to the minimal electron potential inside an inert gas, we obtain that slow electron drifts inside the condensed inert gas near the lines form the Delaunay network. If we move along a given line of the Delaunay network, the electron potential energy oscillates, as is shown in Fig. 3. The behavior of the electron PES on the Delaunay network lines and near them resembles that for bound atoms in clusters [26–28], with the potential energy surface including many potential wells separated by barriers or saddles. But, based on the experimental data for electron mobility, we take the difference between neighboring minima and maxima of the potential energy to be relatively small if the atom number density is near that corresponding to the maximum electron mobility [21].

4. POTENTIAL ENERGY SURFACE FOR AN EXCESS ELECTRON INSIDE INERT GASES

We have found the character of distribution of the electron potential inside condensed inert gases in the density range where the electron potential energy is negative and close to the minimal one. The lines of a significant electron attraction inside an inert gas form a Delaunay network, and this result is not based on the assumption of a pairwise character of the electron–atom interaction. We use this assumption at the next stage of evaluation of the electron PES near the lines of maximum attraction, representing the interaction potential of an electron with surrounding atoms in the form of pair interaction potentials $u(r)$ of this electron with nearest atoms,

$$U = \sum_i u(r_i), \tag{11}$$

where r_i is the distance of the electron from the ith nucleus and the pair interaction potential is taken such that formulas (6) and (9) give the minimal electron energy inside an inert gas. Because of a short-range character of the electron interaction, we account for only six nearest-neighbors. We take the dependence $u(r)$ to be identical to that given by formula (10),

$$u(r) = -C'\left(\frac{R_{\min}}{r}\right)^3 + A'\exp\left(-\alpha\frac{r^3}{R_{\min}^3}\right). \tag{12}$$

This interaction potential has the minimum at the distance $R_{\min}$.

Within the framework of this model, we represent the observed electron potential inside an inert gas as the average for points *3* and *4* in Fig. 2. Then, on the basis of the data in Table 3, we have that the observed electron potential energy $\bar{U}(N)$ at a given number density N of atoms is

$$\bar{U}(a) = u\left(\frac{a}{2}\right) + 2u\left(a\frac{\sqrt{3}}{2}\right) + 3u\left(\frac{a}{\sqrt{2}}\right). \quad (13)$$

Taking this relation at the minimum of the electron potential, i.e., at $a = a_{\min}$, and expanding the interaction potential $u(r)$ near its minimum,

$$u(r) = u(R_{\min}) + \frac{u''(R_{\min})}{2}(r - R_{\min})^2,$$

we obtain the minimal electron potential

$$\bar{U}(a_{\min}) = u\left(\frac{a_{\min}}{2}\right) + 2u\left(a_{\min}\frac{\sqrt{3}}{2}\right) + 3u\left(\frac{a_{\min}}{\sqrt{2}}\right) = 6u(R_{\min} \pm \Delta R), \quad (14)$$

where

$$R_{\min} = \frac{1 + 2\sqrt{3} + 3/\sqrt{2}}{12}a_{\min} = 0.726a,$$
$$\Delta R = \pm 0.123a_{\min}, \quad \frac{\Delta R}{R_{\min}} = \pm 0.17. \quad (15)$$

The above estimates, together with relation (13), allow us to determine the parameters of the interaction potential $u(r)$ taken in form (12); they are listed in Table 2.

This approach allows us to construct the potential energy surface of an excess electron inside inert gases based on experimental data. Although the model uses a pair interaction between an excess electron and inert gas atoms, this is not of importance at the final stage of the analysis, because parameters of this model are taken from experimental results. In other words, the general character of the electron interaction is based on the Delaunay network and does not include the pairwise character of the electron interaction inside an inert gas, whereas the values of the electron potential includes this assumption. Therefore, the above behavior of the electron PES is valid strictly, while the accuracy of the values of the electron potentials at a given electron position are valid qualitatively.

5. ELECTRON DRIFT AT OPTIMAL DENSITIES OF INERT GASES AND LOW ELECTRIC FIELDS

The above analysis allows us to schematically draw equipotential surfaces for an excess electron inside an inert gas at a given number density of atoms, when the mobility of a slow excess electron is high. The lines of the minimum potential energy then form a Delaunay network, and for the crystal state of an inert gas, these lines are straight and pass between nearest atoms. A general shape of lines of the minimum potential energy are also correct for liquids in principle. Indeed, first, a change of the number density of atoms resulting from the solid–liquid phase transition for inert gases is approximately 15%, and a change of the average distance between atoms is correspondingly three times less. Second, the distortion of lines of the minimal potential for an excess electron is also inessential, because a slow electron is a quantum object, and the difference of the electron potentials inside and outside an inert gas allows us to find the energy of the electron level inside the inert gas, but not the minimum potential for an excess electron inside it. Correspondingly, the de Broglie wavelength is not small for a slow excess electron and a weak distortion of straight lines of the minimum electric potential for an excess electron, in passing from a solid to a liquid, is not of importance. Hence, our consideration relates simultaneously to the solid and liquid states of condensed inert gases.

Thus, we consider motion of a slow electron inside an inert gas whose density corresponds to electron attraction inside it. Therefore, an electron is bound inside the inert gas and moves along tubes centered at lines of the minimum electron potential, which are represented in Fig. 2 for a solid inert gas. These tubes of identical potentials are widened slightly near their intersections, and the distances between neighboring points of tube intersections is $a\sqrt{3}/2$, as follows from Fig. 2. When a slow electron propagates along a tube (see Fig. 3), its scattering proceeds in nodes of tube intersections, and as a result of this scattering, it transfers to another tube. We take the probability γ for the electron scattering in an intersection node to be small, and then the mean free path λ of an electron during its propagation along a potential tube is relatively large, $\lambda \sim a/\gamma$ ($\gamma \ll 1$).

Electron scattering in the intersection regions of potential energy tubes is similar to electron scattering on atoms in a gas because the time of strong interaction for an excess electron that causes scattering is a small part of the total time in both cases. In addition, in both cases, an electron is scattered mainly elastically, and only a small part of the electron energy ($\sim m/M$) is transferred to nuclear heating (m is the electron mass and M is the atom mass). Below, we therefore use formulas for the electron drift velocity w and its average velocity v assuming that the electron is scattered in a gas (see,

Table 4. Parameters of the drift of an excess electron in liquid inert gases under optimal conditions and low electric field strengths

	Ar	Kr	Xe
T_{max}, K	155	170	223
v_t, 10^6 cm/s	7.7	8.1	9.3
N_{max}, 10^{22} cm^{-3}	1.2	1.4	1.2
a_{max}, Å	4.9	4.7	4.9
K_{max}, cm^2/ V s	1800	4600	6000
γ_{min}	0.0062	0.0022	0.0015
E_*, V/cm	16	4.5	3.2

e.g., [35–37]). For an electron moving in an external electric field of a strength E, we then have

$$w \approx \frac{eE\lambda}{mv}, \quad v \approx \sqrt{\frac{M}{m}}w, \tag{16}$$

which gives

$$w \approx \left[\frac{eEa}{\gamma(v)\sqrt{mM}}\right]^{1/2}, \quad v \approx \left[\frac{eEaM^{1/2}}{\gamma(v)m^{3/2}}\right]^{1/2}. \tag{17}$$

In these formulas, we take the average electron velocity to be large compared to the electron thermal velocity in the absence of an external electric field. If this electric field is weak and does not change the Maxwell velocity distribution for excess electrons, the zero-field electron mobility K is

$$K \approx \frac{ea}{v\gamma}. \tag{18}$$

In evaluating the parameters of the electron drift, if it proceeds according to the above scheme, we are based on experimental data. Table 4 contains the number densities of atoms N_{max} and temperatures T_{max} of liquid inert gases [13] that provide the maximum zero-field mobility of electrons, the corresponding distances a_{max} between nearest-neighbor atoms, and the thermal electron velocity

$$v_t = \sqrt{8T_{max}/\pi m}$$

under these conditions. Then, the above formulas give the minimal probability γ_{min} of electron scattering, which is a typical probability for the transition to another current tube at a point of tube intersection, and a typical electric field strength E_* at which a change in the average electron velocity due to the electron drift in an electric field is comparable to the initial thermal velocity. Starting from these electric field strengths, the electron drift parameters depend on the electric field strength.

We note that this mechanism for the electron drift, with the electrons propagating along the tubes whose centers form a Delaunay network, is valid only for some range of inert gas parameters at which the electron is locked inside the inert gas in regions near the Delaunay network. This mechanism of the electron drift provides high mobility for slow electrons, which can be used for determination of the range of the inert gas parameters and electric field strength, where this mechanism of the electron drift applies. For xenon at least, these conditions are fulfilled in a wide range of the indicated parameters. An increase of the electric field strength leads to an increase of the electron energy and causes broadening of the region between atoms where an excess electron can be located. Finally, at high electron energies, the electron scattering changes from the tube character to scattering on atomic cores. Then, the electron mobility decreases sharply with an increase of the electric field strength. In reality, for xenon, the experimental data analysis shows that this tendency exists, but the transition is not reached.

Electron scattering is also intensified if the gas parameters differ from the optimal ones. If the atom number density deviates from the optimal one, the attractive electron potential energy on the Delaunay network decreases, which leads to a stronger electron scattering in regions of tube intersection. At a given atom number density, the lower gas temperature, the higher is the electron drift velocity. This is explained by distortions in the atom distributions that increase as the temperature increases.

Thus, we represent the character of the drift of a slow electron in condensed inert gases under the optimal number density and temperature. The electron scattering under these conditions differs in principle from that in gases, where electrons collide with individual atoms separately. In this case, an electron is moving along a certain tube and transfers to another tube at points of their intersections. Axes of these tubes form the Delaunay network. This character of electron scattering also differs from the wave character of scattering in a crystal lattice, where scattering is determined by deviation of atom positions from the crystal lattice sites, such that scattering parameters vary significantly during the melting. In the case of the tube character of electron scattering, melting does not significantly change the electron drift parameters. We add that the tube character of the electron drift is realized in a restricted range of the inert gas parameters and is valid at not too high electric field strengths.

On the basis of this analysis, we can single out the range of parameters that corresponds to the maximum electron mobility in condensed inert gases. As the number density of atoms increases, the effective interaction for an excess electron with atoms of a condensed inert gas varies from attraction due to the exchange interac-

tion with an individual atom because of a negative electron–atom scattering length to repulsion owing to the Pauli exclusion principle when the electron penetrates inside an atom. Evidently, the maximum electron mobility corresponds to moderate atomic number densities corresponding to the transition from the first form of interaction to the second one. Then, the PES part of location of an excess electron consists of narrow tubes with intersections, and the electron can propagate along these tubes. As the number density of atoms increases, these tubes are destroyed in regions near atoms where the electron is locked. If the atomic number density decreases, tubes widen, and the electron may transfer more effectively to tubes of another direction. In both cases, the electron mobility decreases.

We note that a temperature increase leads to an increase of fluctuations in positions of individual atoms, which causes the destruction of a PES tube. But a pressure increase leads to a decrease of these fluctuations and hence stabilizes the PES tube when the tube corresponds to optimal conditions. In analyzing the optimal conditions for the electron mobility, we are mostly based on experimental data. But experimental data of this problem study are fragmentary. Additional experimental studies are required in order to construct the optimal range in coordinates of the atomic number density, temperature, and electric field strength for each heavy inert gas (Ar, Kr, and Xe). We also expect from the subsequent experimental study that, at high pressure, the electron mobility will decrease with an increasing pressure.

We also note the peculiarities of inelastic electron scattering in condensed inert gases. If the electron energy is not small, and the electron can be considered as a classical object, its inelastic scattering inside an inert gas is related to excitation of phonons, and each act of elastic scattering is accompanied by a loss of approximately the m/M portion of the electron energy; in other words, the process of inelastic scattering of a classical electron in a condensed inert gas is similar to that in rare gases. This is used in formulas (16) and (17). But a slow electron is a quantum object, and its inelastic scattering proceeds in another manner. Indeed, the electron states are characterized by discrete levels, and inelastic electron scattering requires its transition to an excited electron level. Therefore, at low electric field strengths, the inelastic electron scattering is weak and becomes the same as in a gas when the electron is excited sufficiently strongly, such that its levels are located sufficiently close.

In considering inelastic electron scattering, we restrict ourselves to just this limiting case. At high electric field strengths, the electron energy acquired from the field suffices for excitation of inert gas atoms. The excitation processes are in principle the same as in a gas, which are analyzed in detail in [37]. The efficiency of this process, that is, the electron energy part consumed to atom excitation, increases with an increase of the average electron energy $\bar{\varepsilon}$ [15, 21] and is of the order of 10% when the ratio $\bar{\varepsilon}/\Delta\varepsilon \gtrsim 0.1$ ($\Delta\varepsilon$ is the atom excitation energy). We are also guided by the experimental efficiency value of 18% in solid xenon [38, 39].

6. PECULIARITIES OF SELF-SUSTAINING DISCHARGE IN CONDENSED INERT GASES

An applied aspect of the phenomenon of electron drift in condensed inert gases is realized in electric discharge, with the electric energy being converted into the energy of emitted photons in the vacuum ultraviolet spectrum range. Excess electrons drifting in condensed inert gases excite inert gas atoms, which leads to transformation of an electrical energy into the energy of emitted photons. Because the electron energy is high, the efficiency of energy transformation is relatively high. During these processes, an excess electron cannot ionize the medium, because its energy is below the ionization potential due to an effective atom excitation. The electrons are therefore injected into a sample from outside and only play the role of carriers of a negative charge, in contrast to standard gaseous discharges with ionization inside a sample. Due to this character of discharge maintenance, excess electrons create a noncompensated negative charge in condensed inert gases. This charge restricts the number density of excess electrons and correspondingly the power of the discharge and the intensity of yield radiation [42]. We find the maximum value $N_e^{\max}$ for the electron number density from the Poisson equation that has the form

$$\frac{dE}{dx} = -4\pi e N_e. \tag{19}$$

Here, E is the electric field strength, e is the electron charge, N_e is the electron number density, which is constant inside the inert gas layer, and the coordinate x is perpendicular to the inert gas layer whose thickness is l. From the Poisson equation, requiring $E = 0$ in the layer middle because of the problem symmetry, we obtain the electric voltage U between the layer boundaries due to excess electrons inside the layer as

$$U = \pi e N_e l^2. \tag{20}$$

Formula (20) implies that the electron number density is the greater, the higher is the electric field voltage and the smaller is the layer thickness. In particular, under typical parameters $U = 1$ keV and $l = 1$ mm realized in experiments [38, 39], this formula gives $N_e^{\max} = 2 \times 10^{11}$ cm^{-3}. This electron number density locks the electric current in discharge. We note that the number density $N_e = 1 \times 10^{11}$ cm^{-3} leads to the electric current

density $j \approx 0.01$ A/cm^2 and the discharge power $P = Uj \approx 10$ W/cm^2.

7. CONCLUSIONS

High electron mobility is observed in heavy condensed inert gases (Ar, Kr, Xe) in a narrow range of atomic densities. A widespread explanation of this effect [16–20] by the Ramsauer effect in electron scattering on an individual atom is not correct because of a large distance of the electron–atom scattering in comparison with the distance between neighboring atoms at these atomic densities. In reality, the nature of high electron mobility is related to the transition from an attractive interaction potential between an excess electron and the atom ensemble to a repulsive one [21]. In this paper, we have proposed a new mechanism of electron drift in some range of atomic densities and temperatures near the optimal ones that provide the maximum electron mobility. This mechanism is additional to the character of electron drift in gases due to electron scattering on individual atoms and to electron drift in crystals due to scattering of the electron wave on nonuniformities of the crystal lattice.

This character of electron drift consists in propagation of an electron along almost straight channels; electron scattering occurs as a result of the electron transition to a propagation channel of another direction. This new mechanism of electron drift follows from the structure of the potential energy surface near its minimum; it consists of almost straight intersecting tubes, and the minimum of potential energy surface forms a Delaunay network. The tube character of the electron drift leads to high electron mobility. The understanding of this phenomenon allows us to choose optimal conditions for a self-sustaining electric discharge in condensed inert gases as a generator of ultraviolet radiation [38–40] and stimulates new experimental investigations.

ACKNOWLEDGMENTS

This work is supported in part by the Russian Foundation for Basic Research (grant nos. 04-03-32684 and LSS-1953.2003.2).

REFERENCES

1. H. Schnyders, S. A. Rice, and L. Meyer, Phys. Rev. **150**, 127 (1966).
2. L. S. Mi, S. Howe, and W. Spear, Phys. Rev. **166**, 871 (1968).
3. J. A. Jahnke, L. Meyer, and S. A. Rice, Phys. Rev. A **3**, 734 (1971).
4. T. Kimura and G. R. Freeman, Can. J. Phys. **52**, 2220 (1974).
5. K. Yoshino, U. Sowada, and W. F. Schmidt, Phys. Rev. A **14**, 438 (1976).
6. S. S.-S. Huang and G. R. Freeman, J. Chem. Phys. **68**, 1355 (1978).
7. S. S.-S. Huang and G. R. Freeman, Phys. Rev. A **24**, 714 (1981).
8. R. Reininger, U. Asaf, and I. T. Steinberger, Chem. Phys. Lett. **90**, 287 (1982).
9. E. M. Gushchin, A. A. Kruglov, and I. M. Obodovskiĭ, Zh. Éksp. Teor. Fiz. **82**, 1114 (1982) [Sov. Phys. JETP **55**, 650 (1982)].
10. V. V. Dmitrienko, A. S. Romanuk, S. I. Suchkov, and Z. M. Uteshev, Zh. Tekh. Fiz. **53**, 2343 (1983) [Sov. Phys. Tech. Phys. **28**, 1440 (1983)].
11. R. Reininger, U. Asaf, and I. T. Steinberger, Phys. Rev. B **28**, 3193 (1983).
12. R. Reininger, U. Asaf, I. T. Steinberger, and S. Basak, Phys. Rev. B **28**, 4426 (1983).
13. F. M. Jacobsen, N. Gee, and G. B. Freeman, Phys. Rev. A **34**, 2329 (1986).
14. A. K. Al-Omari, K. N. Altmann, and R. Reininger, J. Chem. Phys. **105**, 1305 (1996).
15. B. M. Smirnov, Usp. Fiz. Nauk **172**, 1411 (2002) [Phys. Usp. **45**, 1251 (2002)].
16. J. Lekner, Phys. Rev. **158**, 130 (1967).
17. M. H. Cohen and J. Lekner, Phys. Rev. **158**, 305 (1967).
18. B. E. Springett, J. Jortner, and M. H. Cohen, J. Chem. Phys. **48**, 2720 (1968).
19. V. M. Atrazhev and I. T. Iakubov, J. Phys. C **14**, 5139 (1981).
20. V. M. Atrazhev and E. G. Dmitriev, J. Phys. C **18**, 1205 (1985).
21. E. B. Gordon and B. M. Smirnov, Zh. Éksp. Teor. Fiz. **125**, 1058 (2004) [JETP **98**, 924 (2004)].
22. B. N. Delaunay, in *Proceedings of International Mathematical Congress, Toronto, 1924* (Univ. of Toronto Press, Toronto, 1928), p. 695.
23. B. N. Delaunay, Z. Kristallogr. **84**, 109 (1933).
24. B. N. Delaunay, Usp. Math. Nauk, No. 3, 16 (1937).
25. N. N. Medvedev, *Method of Voronoi—Delaunay in Study of Structure of Noncrystal Systems* (Sib. Otd. Ross. Akad. Nauk, Novosibirsk, 2000) [in Russian].
26. R. S. Berry, in *Theory of Atomic and Molecular Clusters*, Ed. by J. Jellinek (Springer, Berlin, 1999).
27. D. J. Wales, Adv. Chem. Phys. **115**, 1 (2000).
28. D. J. Wales, *Energy Landscapes* (Cambridge Univ. Press, Cambridge, 2003).
29. E. Fermi, Nuovo Cimento **11**, 157 (1934).
30. M. Ya. Ovchinnikova, Zh. Éksp. Teor. Fiz. **49**, 275 (1965) [Sov. Phys. JETP **22**, 194 (1966)].
31. H. S. W. Massey, E. H. S. Burhop, and H. B. Gilbody, *Electronic and Ionic Impact Phenomena*, 2nd ed. (Clarendon, Oxford, 1969).
32. R. V. Galiulin, Kristallografiya **25**, 901 (1980) [Sov. Phys. Crystallogr. **25**, 517 (1980)].
33. R. V. Galiulin, Kristallografiya **43**, 366 (1998) [Crystallogr. Rep. **43**, 332 (1998)].

34. R. V. Galiulin, Zh. Vychisl. Mat. Mat. Fiz. **43**, 791 (2003).
35. L. G. H. Huxley and R. W. Crompton, *The Diffusion and Drift of Electrons in Gases* (Wiley, New York, 1974).
36. B. M. Smirnov, *Physics of Weakly Ionized Gases* (Nauka, Moscow, 1978; Mir, Moscow, 1981).
37. B. M. Smirnov, *Physics of Ionized Gases* (Wiley, New York, 2001).
38. E. B. Gordon, G. Frossati, and A. Usenko, Zh. Éksp. Teor. Fiz. **123**, 962 (2003) [JETP **96**, 846 (2003)].
39. A. Usenko, G. Frossati, and E. B. Gordon, Phys. Rev. Lett. **90**, 153201 (2003).
40. A. S. Schussler, J. Burghoorn, P. Wyder, *et al.*, Appl. Phys. Lett. **77**, 2786 (2000).
41. E. B. Gordon and A. F. Shestakov, Low Temp. Phys. **27**, 883 (2001).
42. E. B. Gordon, O. S. Rzhevskii, and V. V. Khmelenko, Quantum Electron. **21**, 209 (1994).
43. E. B. Gordon, V. V. Khmelenko, and O. S. Rzhevskii, Chem. Phys. Lett. **217**, 605 (1994).
44. B. M. Smirnov, Usp. Fiz. Nauk **171**, 1291 (2001) [Phys. Usp. **44**, 1229 (2001)].
45. V. M. Atrazhev, I. V. Chernysheva, and T. D. Doke, Jpn. J. Appl. Phys. **41**, 1572 (2002).

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 22–32.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 30–41.

ATOMS, MOLECULES, OPTICS

Stochastic Modeling of Coherent Phenomena in Strongly Inhomogeneous Media

V. L. Kuz'min[a], I. V. Meglinski[b,c], and D. Yu. Churmakov[c]

[a]*St. Petersburg Institute of Trade and Economics, ul. Novorossiĭskaya 50, St. Petersburg, 194021 Russia*
e-mail: vladimir.kuzmin@paloma.spbu.ru
[b]*Saratov State University, ul. Universitetskaya 42, Saratov, 410026 Russia*
[c]*Cranfield University, School of Engineering, Cranfield, MK43 0AL, UK*

Received November 24, 2004

Abstract—A procedure of numerical simulation for coherent phenomena in multiply scattering media is developed on the basis of the juxtaposition of a Monte Carlo stochastic method with an iterative approach to the solution of the Bethe–Salpeter equation. The time correlation function and the interference component of coherent backscattering are calculated for scalar and electromagnetic fields. The results of simulation are in good agreement with experimental results, as well as with theoretical results obtained by generalizing the Milne solution.

1. INTRODUCTION

Stochastic modeling within a Monte Carlo method [1–10] has recently found wide application in the study of coherent phenomena associated with the propagation of light in randomly inhomogeneous media. These phenomena (coherent backscattering, spatial and time correlations of intensity) are attributed to the wave nature of light and, in spite of the multiple scattering regime, are observed in a wide variety of dielectric systems [11], such as colloidal suspensions, liquid crystals, and biological tissues.

The standard modeling technique for radiation in a random medium is based on the concept of intensity transfer. The phase relations between the fields that contribute to the intensity remain outside the framework of this modeling; one needs a special approach to take into account these relations.

The theory of multiple scattering in random media, including the theory of coherent and interference phenomena, is based on the Bethe–Salpeter equation (see [12]). In the present paper, we juxtapose a stochastic Monte Carlo method [13] with a theoretic method that is based on the representation of the solution to the Bethe–Salpeter equation as a series in scattering orders to demonstrate how the standard Monte Carlo method is generalized to a unified approach to the numerical simulation of the time correlation function of intensity, coherent backscattering, and other coherent phenomena that require consideration of phase shifts.

Due to the multiple scattering, polarized light is completely depolarized; this fact allows one to restrict the consideration to a scalar field. However, in the case of backscattering, the scattered light remains partially polarized due to the contributions of lower order scattering terms. Experiments point to a significant role of polarization in backscattering [14–16].

In [17–22], the problem of multiple scattering by point, Rayleigh, particles was solved by generalizing the Milne solution to the case of electromagnetic waves. In [17. 18], a vector transport equation was solved strictly for backward scattering with regard to the interference component, and, in [19, 20], the angular dependence of the backscattering intensity was calculated with regard to polarization. In [21], a solution for the time correlation function was obtained. In [22], this solution was generalized to the case of finite-size scatterers.

In [2], a Monte Carlo procedure was applied to calculate the intensity of coherent backscattering, including a coherent component; for Rayleigh scattering, the peak of coherent backscattering proved to be much less than the enhancement predicted by the exact solution [19, 22]. In [23], the authors also simulated backscattering of electromagnetic waves; they calculated the rate of depolarization of linearly polarized light as a function of the number of scattering events. For Rayleigh scattering, the function obtained coincides with that predicted in [24], which was obtained within a diffusive approximation; however, as the anisotropy of the single-scattering cross section increases, the result of numerical simulation appreciably differs from theoretical results.

In the present paper, we develop a method for stochastic modeling of coherent phenomena with regard to the polarization of an electromagnetic field. For Rayleigh scattering, the numerical results are in agreement with the theoretical results. This fact suggests that the data obtained in the general case of anisotropic scattering cross section are reliable.

We consider the case most frequently used in theoretical investigations, of the scattering of radiation in a medium that occupies a half-space with a plane boundary. To facilitate consideration of the method, we examine the normal incidence and scattering close to the backward scattering.

In Section 2, we present general expressions for the time correlation function and the interference component of backscattering. In Section 3, we compare the methods of summation of ladder diagrams with a Monte Carlo method and present the results of simulation for the time correlation function and coherent backscattering for a scalar field. In Section 4, we consider the results of simulation for linearly polarized light. In the Conclusions, we discuss the results.

2. TRANSFER OF FIELD CORRELATIONS

Suppose that a medium occupies the half-space $z > 0$, where z is the Cartesian coordinate normal to the boundary of the medium and $\mathbf{k}_i$ and $\mathbf{k}_s$ are the wavevectors of the incident and scattered plane waves.

The transfer of field correlations in an inhomogeneous dispersion medium with random space–time fluctuations of dielectric constant is described by the Bethe–Salpeter integral equation

$$\begin{aligned}\hat{\Gamma}(\mathbf{R}_2, \mathbf{R}_1, t|\mathbf{k}_s, \mathbf{k}_i) &= k_0^4\tilde{G}(\mathbf{k}_s - \mathbf{k}_i, t)\delta(\mathbf{R}_2 - \mathbf{R}_1)\hat{I} \\ &+ k_0^4\int d\mathbf{R}_3\tilde{G}(\mathbf{k}_s - \mathbf{k}_{23}, t)\hat{\Lambda}(\mathbf{R}_2 - \mathbf{R}_3) \\ &\times\hat{\Gamma}(\mathbf{R}_3, \mathbf{R}_1, t|\mathbf{k}_{23}, \mathbf{k}_i).\end{aligned} \tag{2.1}$$

Here, $\Gamma_{\beta_1\beta_2\alpha_1\alpha_2}(\mathbf{R}_2, \mathbf{R}_1, t|\mathbf{k}_s, \mathbf{k}_i)$ is the propagator, or the Green's function of the Bethe–Salpeter equation, which represents a fourth-rank tensor. It describes the transfer of two complex-conjugate fields that arrive at the point $\mathbf{R}_1$ with a time shift t, the wavevectors $\mathbf{k}_i$, and the polarizations described by the Cartesian indices α_1 and α_2, and go out from the point $\mathbf{R}_2$ with the wavevectors $\mathbf{k}_s$ and the polarizations described by the Cartesian indices β_1 and β_2. Here, $k_0 = 2\pi/\lambda$ is the wavenumber; λ is the wavelength; $k_s = k_i = k = nk_0$; n is the refractive index of the random medium: $n = n_1 + in_2$, where n_1 and n_2 are the real and imaginary parts of n, respectively; the imaginary part n_2 defines the photon mean free path $(2n_2k_0)^{-1} = l$. The quantity

$$\mathbf{k}_{ij} = k(\mathbf{R}_i - \mathbf{R}_j)|\mathbf{R}_i - \mathbf{R}_j|^{-1}$$

defines the wavevector between the ith and jth scattering events. The fourth-rank tensor $\hat{\Lambda}(\mathbf{R})$,

$$\begin{aligned}\Lambda_{\alpha\beta\mu\nu}(\mathbf{R}) &= \left(\hat{I} - \frac{\mathbf{R}\otimes\mathbf{R}}{R^2}\right)_{\alpha\mu} \\ &\times\left(\hat{I} - \frac{\mathbf{R}\otimes\mathbf{R}}{R^2}\right)_{\beta\nu}\frac{\exp(-R/l)}{R^2},\end{aligned} \tag{2.2}$$

represents a direct product of the complex-conjugate pair of Green's functions of the Maxwell wave equation in the far-field region and describes the transformation of a pair of fields with polarizations μ and ν into a pair of fields with polarizations α and β in a single scattering event.

In the weak scattering approximation ($\lambda \ll l$), which is usually valid in the dielectric systems under consideration, $\tilde{G}(\mathbf{q}, t)$ represents the Fourier image of the correlation function of space–time fluctuations of the dielectric constant:

$$\begin{aligned}\tilde{G}(\mathbf{q}, t) &= \frac{1}{(4\pi)^2}\int d\mathbf{r}\langle\delta\varepsilon(0, 0)\delta\varepsilon(\mathbf{r}, t)\rangle \\ &\times\exp(-i\mathbf{q}\cdot\mathbf{r}).\end{aligned} \tag{2.3}$$

A key role in the problems of multiple scattering is played by the optical theorem, which relates the single-scattering cross section to the scattering length l_s. For an electromagnetic field, the optical theorem in the approximation of weak scattering, or the Born approximation, is expressed as

$$l_s^{-1} = \Gamma_R^{-1}k_0^4\int d\Omega_s\tilde{G}_0(\mathbf{k}_s - \mathbf{k}_i). \tag{2.4}$$

Here, $\tilde{G}_0(\mathbf{q}) = \tilde{G}(\mathbf{q}, t)$ is the Fourier image of the static correlation function of fluctuations of the dielectric constant, $\Gamma_R = 2(1 + \overline{\cos^2\theta})^{-1}$ is the Rayleigh factor, and

$$\overline{\cos^2\theta} = \frac{\int d\Omega_s\tilde{G}_0(\mathbf{k}_s - \mathbf{k}_i)\cos^2\theta_s}{\int d\Omega_s\tilde{G}_0(\mathbf{k}_s - \mathbf{k}_i)}$$

is the squared cosine, averaged over the single-scattering cross section, of the scattering angle between the wavevectors $\mathbf{k}_i$ and $\mathbf{k}_s$.

The photon mean free path l and the scattering length l_s are related by the formula

$$\frac{1}{l} = \frac{1}{l_s} + \frac{1}{l_a}, \tag{2.5}$$

where l_a is the characteristic length of the absorption due to inelastic scattering. For the media considered here, $l_a \gg l$ and the ratio l/l_s is close to unity.

Let us define a normalized correlation function of fluctuations of the dielectric constant:

$$p(\mathbf{k}_i - \mathbf{k}_s, t) = \frac{\tilde{G}(\mathbf{k}_i - \mathbf{k}_s, t)}{\int\tilde{G}(\mathbf{k}_i - \mathbf{k}_s, 0)d\Omega_s}. \tag{2.6}$$

For $t = 0$, this function coincides with the phase function $p_0(\mathbf{k}_i - \mathbf{k}_s) = p(\mathbf{k}_i - \mathbf{k}_s, 0)$, which describes the cross section of single scattering.

Iterating the Bethe–Salpeter equation and applying the optical theorem, we obtain the series

$$\begin{aligned}\hat{\Gamma}(\mathbf{R}_2, \mathbf{R}_1, t|\mathbf{k}_s, \mathbf{k}_i) &= \Gamma_R l_s^{-1} p(\mathbf{k}_i - \mathbf{k}_s, t)\delta(\mathbf{R}_2 - \mathbf{R}_1)\\ &+ \Gamma_R^2 l_s^{-2} p(\mathbf{k}_s - \mathbf{k}_{21}, t)\hat{\Lambda}(\mathbf{R}_{21}) p(\mathbf{k}_{21} - \mathbf{k}_i, t)\\ &+ \Gamma_R^3 l_s^{-3}\int d\mathbf{R}_3 p(\mathbf{k}_s - \mathbf{k}_{23}, t)\hat{\Lambda}(\mathbf{R}_{23}) p(\mathbf{k}_{23} - \mathbf{k}_{31}, t)\\ &\times \hat{\Lambda}(\mathbf{R}_{31}) p(\mathbf{k}_{31} - \mathbf{k}_i, t) + \ldots,\end{aligned} \tag{2.7}$$

which is usually represented as a series of ladder diagrams.

Let us define a binary correlation function of a field that is observed at a large distance r from the scattering medium as

$$\begin{aligned}&\hat{C}^{(E)}(t|\mathbf{k}_s, \mathbf{k}_i)\\ &= \hat{C}^{(L)}(t|\mathbf{k}_s, \mathbf{k}_i) + \hat{C}^{(V)}(t|\mathbf{k}_s, \mathbf{k}_i),\end{aligned} \tag{2.8}$$

where $\hat{C}^{(L)}(t|\mathbf{k}_s, \mathbf{k}_i)$ is a contribution of ladder diagrams that describes a noncoherent component, while $\hat{C}^{(V)}(t|\mathbf{k}_s, \mathbf{k}_i)$ is the interference component that is observed in the backscattered wave.

In the case of normal incidence and scattering close to the backward scattering, the ladder and interference components of the time correlation function of a field are given by [24, 25]

$$\begin{aligned}&C^{(L)}_{\beta_1\beta_2\alpha_1\alpha_2}(t|\mathbf{k}_s, \mathbf{k}_i)\\ &= \int d\mathbf{R}_1 d\mathbf{R}_2 \Gamma_{\beta_1\beta_2\alpha_1\alpha_2}(\mathbf{R}_2, \mathbf{R}_1, t|\mathbf{k}_s, \mathbf{k}_i)\\ &\times \exp\left(-\frac{z_1 + z_2}{l}\right)\end{aligned} \tag{2.9}$$

and

$$\begin{aligned}C^{(V)}_{\beta_1\beta_2\alpha_1\alpha_2}(t|\mathbf{k}_s, \mathbf{k}_i) &= \int d\mathbf{R}_1 d\mathbf{R}_2\\ &\times \left[\Gamma_{\beta_1\alpha_2\alpha_1\beta_2}\left(\mathbf{R}_2, \mathbf{R}_1, t\left|\frac{\mathbf{k}_s - \mathbf{k}_i}{2}, \frac{\mathbf{k}_i - \mathbf{k}_s}{2}\right.\right)\right.\\ &\left. - k_0^4 \tilde{G}(\mathbf{k}_s - \mathbf{k}_i, t)\delta(\mathbf{R}_2 - \mathbf{R}_1)\delta_{\alpha_1\beta_1}\delta_{\alpha_2\beta_2}\right]\\ &\times \exp\left(-\frac{z_1 + z_2}{l} + i(\mathbf{k}_s + \mathbf{k}_i)_\perp \cdot (\mathbf{R}_2 - \mathbf{R}_1)_\perp\right),\end{aligned} \tag{2.10}$$

respectively, where the subscript "⊥" denotes the component perpendicular to the boundary of the medium. It is easily seen that, for backward scattering, when $\mathbf{k}_s = -\mathbf{k}_i$, the polarized component of the interference contribution $\hat{C}^{(V)}(t|\mathbf{k}_s, \mathbf{k}_i)$ coincides with the polarization component of the main, noncoherent, contribution $\hat{C}^{(L)}(t|\mathbf{k}_s, \mathbf{k}_i)$ before subtracting the contribution of single scattering, while the depolarized components do not coincide.

The noncoherent component defines the time correlation function of a field

$$g_1(t) = \frac{C^{(L)}(t|-\mathbf{k}_i, \mathbf{k}_i)}{C^{(L)}(0|-\mathbf{k}_i, \mathbf{k}_i)}. \tag{2.11}$$

Disregarding non-Gaussian long-range components, we can represent the intensity correlation function as a square of the field correlation function: $g_2(t) = 1 + g_1^2(t)$.

For $t = 0$, the interference component (2.10) describes the peak of coherent backscattering

$$I^{CBS}(\theta_s) = \frac{C^{(V)}(0|\mathbf{k}_s, \mathbf{k}_i)}{(0|-\mathbf{k}_i, \mathbf{k}_i)} \tag{2.12}$$

and its angular dependence.

3. SIMULATION OF A SERIES IN SCATTERING ORDERS: A SCALAR FIELD

Let us juxtapose the procedure of analytic summation of a series of ladder diagrams with a Monte Carlo method that combines the calculation scheme of stochastic trajectories with the application of statistical weights [26, 27]. First, consider a scalar field. When passing to the case of a scalar field, the tensor $\hat{\Lambda}(\mathbf{R})$ is replaced by the scalar function $\Lambda_0(R) = R^{-2}\exp(-R/l)$, and the Rayleigh factor Γ_R is replaced by unity. The first term of the iterative series describes single scattering, the second describes double scattering, etc.

A Monte Carlo method describes the radiation propagation as a random process that consists of one, two, …, n scattering events. The addition of one link of the ladder $\Lambda(R_{jj-1})p_0(\mathbf{k}_j - \mathbf{k}_{j-1})$ in the theoretical description is realized in the numerical experiment by a path R traversed by a photon (a packet of photons) up to the next scattering event. The random photon mean free path R between two successive scattering events is defined by the Poisson distribution [13]:

$$f(R) = l^{-1}\exp(-R/l). \tag{3.1}$$

This distribution implies that $R = -l\ln\xi$, where ξ is the probability of the fact that the mean free path is greater than R; the value of ξ is chosen by a random number generator from the interval [0, 1]. The propagation

direction of a photon packet in each event of elastic scattering is defined randomly with regard to the phase function (2.6). As a result, one obtains a stochastic trajectory of a photon that is emitted at the point $\mathbf{R}_0 = \mathbf{R}_S$, experienced n collisions at the points $\mathbf{R}_1, \ldots, \mathbf{R}_n$, and detected on the surface at the point $\mathbf{R}_{n+1} = \mathbf{R}_D$; this trajectory randomly simulates the nth-order term in the above-mentioned iterative series (Fig. 1).

Let W_i be the statistical weight of the ith photon that reaches a detector. Then, the sum of these weights defines, in arbitrary units, the intensity of scattered radiation at the detection point,

$$I = \Sigma_i W_i.$$

In the case of isotropic scattering, $W_i = 1/N_{\text{ph}}$, where N_{ph} is the sample size; in the case of anisotropic scattering, this quantity is multiplied by the product of phase functions.

The complexity of analytic calculations is associated with the fact that the integrals over $\mathbf{R}_i$ cannot be uncoupled because the phase functions depend on the mutual disposition of scattering particles. The numerical simulation uncouples this chain and, at each step, randomly defines the direction and the magnitude of the mean free path of a photon packet.

By virtue of the normalization condition for the phase function,

$$\int p_0(\mathbf{k}_\text{i} - \mathbf{k}_\text{s})\, d\Omega = 1, \tag{3.2}$$

the statistical weight of a photon packet is not changed after each scattering event. In theoretical description, the conservation of the weight of a packet is fulfilled according to the optical theorem. Indeed, since

$$\int \Lambda_0(R)\, d\mathbf{R} = 4\pi l,$$

the expansion parameter of iterative series (2.7) is given by

$$l_\text{s}^{-1} \int d\Omega_n \int d\mathbf{R}_{i+1} \Lambda_0(\mathbf{R}_{i+1} - \mathbf{R}_i)\, p_0(\mathbf{k}_{i+1} - \mathbf{k}_i) = l_\text{s}^{-1} l. \tag{3.3}$$

Let us show that it is the form of the propagator $\Lambda_0(R)$ that leads to distribution (3.1). In the absence of absorption, the quantity $l_\text{s}^{-1} l$ is exactly equal to unity, which points to the conservation of the weight of the photon packet. In analytic calculations, it is the condition $l_\text{s}^{-1} l = 1$ that makes the method of successive approximations inapplicable for solving the Bethe–Salpeter equation.

A photon contributes to the detected signal provided that it intersects the boundary of the medium at a given angle at the detection point. Then, the trajectory of the next photon is modeled. The sample size of incident photons varied from 10^5 to 10^7. The modeling of a photon trajectory terminated when the number of scattering events became greater than 10^4. When the statistical weight of a photon became less than 10^{-3}, we also ceased to follow upon the trajectory of the photon. According to our estimates, the detection probability of such a photon on the surface is no greater than 10^{-2}; therefore, our approach leads to an error of at most 10^{-5}. We controlled the accuracy of calculated parameters by the stability of their numerical values under increasing sample size. For a sample size of 10^5, the intensity is stable up to at least four digits.

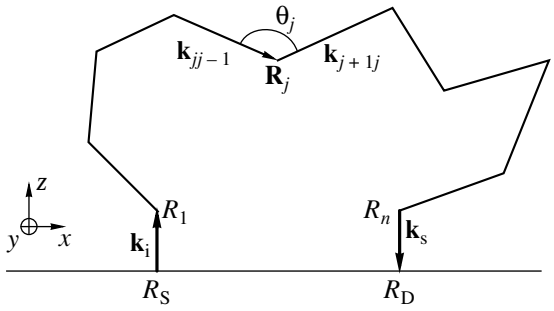

Fig. 1. Trajectory of random walks of a photon from the input point $\mathbf{R}_\text{S}$ to the output point $\mathbf{R}_\text{D}$; $\mathbf{R}_1$ and $\mathbf{R}_n$ are the points of the first and the last, nth, scattering events; $\mathbf{k}_{jj-1}$ and $\mathbf{k}_{j+1j}$ are the wavevectors before and after the jth scattering event; and θ_j is the angle between the above wavevectors.

As the phase function, we used the Henyey–Greenstein function (see [28]). In [2, 29], the authors used the Rayleigh–Gans function.

When a fraction of scattered photons possessing the required properties (for example, phonons scattered into a narrow solid angle when modeling coherent backscattering) is small, the time it takes to accumulate reliable statistics may prove to be extremely large in the above-described standard modeling technique.

To improve the statistics, we applied a semianalytic method of modeling [30], which is also known as the method of local estimation [31], in which each photon contributes to the scattered radiation. Suppose that a trajectory contains N scattering events. The weight of the ith photon after n ($n < N$) scattering events is $W_n^{(i)}$. The contribution of all trajectories with the number of scattering events n is simulated by the term of iterative series (2.7) that describes n-tuple scattering in the ladder component (2.9):

$$C^{(L)} \sim \Sigma_i W_n^{(i)} \exp(-z_n^{(i)}/l),$$

where $z_n^{(i)}$ is the distance from the point of the last, nth scattering event to the boundary of the medium. Thus, in this approach, each photon contributes to the detected radiation. The accuracy of modeling is easily

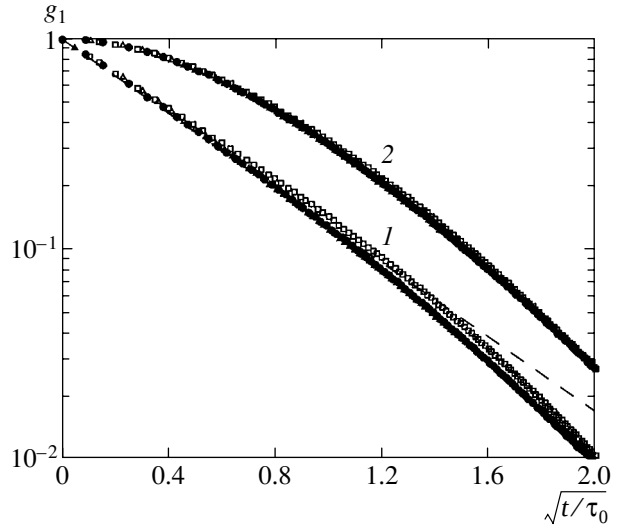

Fig. 2. The time correlation function of a field scattered backward by (*1*) a semi-infinite medium and (*2*) a layer of thickness $L = l^*$ as a function of the time argument $\sqrt{t/\tau}$ for three values of the anisotropy parameter $\overline{\cos\theta} = 0$ (□), 0.5 (●), and 0.9 (△); the dashed curve represents approximation by function (3.6) with $\gamma = 2$.

monitored by comparing with analytic results, which can be obtained for lower orders of scattering. For example, in the case of isotropic scattering, the contributions of single and double scattering to the intensity are given by

$$I_{\text{single}} = l^{-1}\int dz_1 d\mathbf{R}_2 \delta(\mathbf{R}_2 - \mathbf{R}_1)\exp\left(-\frac{z_1 + z_2}{l}\right) = \frac{1}{2},$$

$$I_{\text{double}} = (4\pi)^{-1} l^{-2}\int dz_1 \int d\mathbf{R}_2 \Lambda_0(\mathbf{R}_2 - \mathbf{R}_1)$$
$$\times \exp\left(-\frac{z_1 + z_2}{l}\right) = \ln\sqrt{2} = 0.346\ldots.$$

Using the method described, by simulation we reproduced these theoretical results with a high degree of accuracy.

The analysis involving the juxtaposition of the theoretical approach based on the Bethe–Salpeter equation with the Monte Carlo method allows us to generalize the latter method for modeling coherent multiple scattering phenomena.

The difference between the calculation of the time correlation function and the calculation of intensity lies in the fact that the direction of a scattered photon packet is determined by a generalized phase function $p(\mathbf{k}_j - \mathbf{k}_{j-1}, t)$, which depends on the time shift t, rather than by the phase function itself. In the majority of known applications [15, 32], the authors investigate a diffusion mechanism of the time evolution of irregularities, when the time correlation function of the intensity fluctuations can be represented as a product of a static correlation function and an exponential function:

$$p(q, t) \approx p_0(q)\exp(-D_s q^2 t), \tag{3.4}$$

where D_s is the coefficient of self-diffusion. Thus, the time correlation function is calculated in the Monte Carlo method as

$$g_1(t) = \sum_{i=1}^{N_{\text{ph}}} W_i \exp\left(-2\frac{t}{\tau} n_i\left(1 - \frac{1}{n_i}\sum_j^{n_i}\cos\theta_j\right)\right), \tag{3.5}$$

where $\tau = (D_s k^2)^{-1}$ is the characteristic time of Brownian diffusion of a scatterer to a distance λ and θ_j is the scattering angle in the jth scattering event. The result remains virtually unchanged if we replace the sample average by the average over the phase function

$$\frac{1}{n_i}\sum_j^{n_i}\cos\theta_j \longrightarrow \overline{\cos\theta} = \frac{\int d\Omega_s \tilde{G}_0(\mathbf{k}_s - \mathbf{k}_i)\cos\theta_s}{\int d\Omega_s \tilde{G}_0(\mathbf{k}_s - \mathbf{k}_i)},$$

in agreement with the diffusive approximation.

In the isotropic case, there is an exact Milne solution that allows one to control the results of simulation. The exact solution gives the following value for the ratio of the intensity of the total backscattered radiation to the intensity of single scattering: $I/I_{\text{single}} = 8.455\ldots$ (see [12]). The method of simulation described reproduces this value with an accuracy of at least four digits for a sample size on the order of 10^5. To reduce the simulation time, we calculated, within the diffusive approximation, the contribution of photons for which the distance between the input and output points is several dozens of times greater than the mean free path l, whereas the contribution of photons that are emitted at a distance less than l is calculated by the simulation scheme described.

Figure 2 represents the results of simulation for the time correlation function of field for three scattering media with different values of the anisotropy factor $\overline{\cos\theta} = 0$, 0.5, and 0.9. We chose the value $l = 33$ μm, which corresponds to the values of the transport length $l^* = l(1 - \overline{\cos\theta})^{-1}$, ranging from 33μm for the isotropic case to 333 μm for the case of strong anisotropy ($\overline{\cos\theta} = 0.9$). In terms of $\sqrt{t/\tau}$, the correlation function is practically universal and does not depend on the anisotropy of single scattering. Note also that these results are in good agreement with experimental data [15, 16]. The obtained time correlation function of field is well described by the formula

$$g_1(t) \propto \exp(-\gamma\sqrt{6t/\tau}), \tag{3.6}$$

which was proposed in [15].

The slope ratio γ of the time correlation function of field in a medium with isotropic scattering cross section: Theoretical predictions and results of simulation

Incident radiation	Scattered radiation	Diffusive approximation	Milne solution	Monte Carlo method
Plane wave	Plane wave	$2\frac{(1+z^*)^2}{1+2z^*}$	2 [12, 22]	≈2
Plane wave	Point detector	$1+z^*$	≈1.53 [22]	≈1.53
Point source	Total scattered radiation	z^*		≈1.2

Note: The Milne parameter is $z^* \approx 0.71$.

The deviation from a linear decay for very small times is attributed to insufficient statistics at large distances. As time increases, the contribution of long optical paths decreases, thus weakening the requirements on the statistical sample size.

Strictly speaking, the linear dependence on $\sqrt{t/\tau}$ is realized only for the scattering from a semi-infinite medium. In the case of scattering from a layer of finite thickness, the time correlation function decreases linearly with time t. However, even for layers with a thickness on the order of the transport mean free path, the time correlation function in terms of $\sqrt{t/\tau}$ weakly depends on the anisotropy parameter $\overline{\cos\theta}$.

The decay rate of the correlation function, which weakly depends on the specific character of a medium in dimensionless terms t/τ, depends rather strongly on the geometry of the experiment. The table presents the values of the slope ratio γ for isotropic scattering that are calculated in the diffusive approximation with the use of the Milne solution and obtained by simulation.

One can see that, in the case of a point source or a point detector, the decay rate of correlations is less than that in the case of plane waves. As pointed out in [15], this is associated with the fact that, in the case of plane waves, the relative contribution of longer optical paths increases and leads to a faster decay of the correlation function.

Note that, compared with the expression for the intensity of a noncoherent component, the expression for the intensity of the interference component of backscattering (see (2.10)) contains an additional factor $\exp(i\mathbf{q}_\perp \cdot (\mathbf{R}_1 - \mathbf{R}_2)_\perp)$. Taking into account the translational invariance with respect to the coordinates $\mathbf{R}_\perp$, we can replace this factor by $\cos(\mathbf{q}_\perp \cdot (\mathbf{R}_1 - \mathbf{R}_2)_\perp)$. When calculating the intensity of the coherent component of backscattering, we should multiply the weight of the ith photon arriving at the interface with vector $\mathbf{k}_s$ at distance $|(\mathbf{R}_S - \mathbf{R}_D^{(i)})_\perp|$ from the input point $\mathbf{R}_S$ by the factor $\cos(\mathbf{q}_\perp(\mathbf{R}_S - \mathbf{R}_D^{(i)})_\perp)$ and sum over all photons. As a result, we obtain the following expression for the peak of coherent backscattering:

$$I^{\mathrm{CBS}}(\theta_s) = \sum_i W_i \cos(\mathbf{q}_\perp \cdot (\mathbf{R}_S - \mathbf{R}_D^i)_\perp) - I_{\mathrm{single}}.$$

The height, or enhancement, of the peak of backscattering is given by

$$h^{\mathrm{CBS}} = \frac{2I - I_{\mathrm{single}}}{I},$$

where I is the intensity of the noncoherent component. We obtained $h^{\mathrm{CBS}} = 1.87$ for isotropic scattering, which is in good agreement with the value $h^{\mathrm{CBS}}_{\mathrm{theor}} = 1.88$ [2] obtained on the basis of the Milne generalized solution. For $\overline{\cos\theta} = 0.9$, we obtained $h^{\mathrm{CBS}} = 1.99$, which also agrees with the expected theoretical value $h^{\mathrm{CBS}}_{\mathrm{theor}} = 2$ for $\overline{\cos\theta} \longrightarrow 1$.

Just as in the case of time correlations, the angular dependence of the peak intensity of coherent backscattering calculated in the dimensionless variables $\tilde{q} = kl^*\sin\theta_s$ is close to the universal one (Fig. 3) and is sufficiently well described by the formula

$$I^{\mathrm{CBS}}(\theta_s) \propto \exp(-\gamma kl^* \sin\theta_s)$$

with $\gamma = 2$. Note that this dependence differs significantly from the dependence predicted by the diffusive approximation [24]:

$$I^{\mathrm{CBS}}_{\mathrm{diff}} \propto 1 - 2\frac{(l + z^*l^*)^2}{l + 2z^*l^*}k\sin\theta_s$$

for $kl^*\sin\theta_s \ll 1$, where $z^* = 0.71\ldots$ is the Milne extrapolation parameter. The values of the slope ratio given by this formula for the isotropic scattering and the strongly isotropic scattering differ by a factor of three: $\gamma^{(\mathrm{diff})} = 2.3$ for $\overline{\cos\theta} = 0$, and $\gamma^{(\mathrm{diff})} = 0.71$ for $\overline{\cos\theta} \longrightarrow 1$ (see [18]). Note that the results of simulation also show

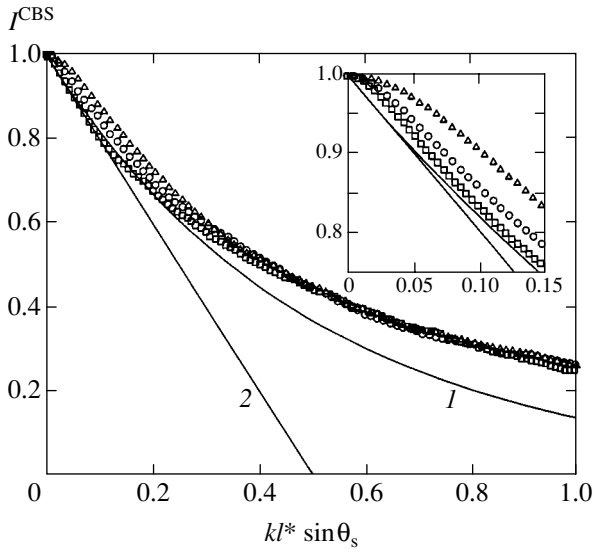

Fig. 3. Peak intensity of coherent backscattering as a function of $kl^*\sin\theta_s$; $\lambda = 0.6$ μm; $l = 33$ μm; and $\overline{\cos\theta} = 0$ (□), 0.5 (○), and 0.9 (△); (*1*) approximation by $\exp(-2kl^*\sin\theta_s)$ and (*2*) function of the form $1 - 2kl^*\sin\theta_s$. The inset shows the initial regions that demonstrate the deviation from universality.

that the decay rate of the peak of coherent backscattering decreases as the anisotropy parameter $\overline{\cos\theta}$ increases in the region of very small scattering angles, $kl^*\sin\theta_s \le 0.1$.

4. CORRELATION EFFECTS FOR LINEARLY POLARIZED LIGHT

In the case of an electromagnetic wave, one should additionally follow up the variation in the direction of the field, characterized by a polarization vector, along a random trajectory of a photon. According to (2.2), to this end, one should calculate the result of the action of the chain of operators [24]

$$\prod_{j=1}^{n}(\hat{I} - (\mathbf{R}_{j+1} - \mathbf{R}_j) \otimes (\mathbf{R}_{j+1} - \mathbf{R}_j)|\mathbf{R}_{j+1} - \mathbf{R}_j|^{-2}) \quad (4.1)$$

on the incident field.

Suppose that, just as in the case of a scalar field, the weight of each incident photon is $1/N_{ph}$. In the electromagnetic field, one defines, in addition to the weight, the initial polarization of the photon field; in the general case, it is defined by three Cartesian coordinates. Suppose that the polarization of each incident photon is defined by a set of three numbers: $\mathbf{P}^{(in)} = (1; 0; 0)$. This vector of initial polarization indicates that the incident field is polarized along the *x* axis.

The polarization of a field is changed under scattering. In addition to the standard procedure of stochastically determining the direction of a photon after a collision and finding the weight function by the phase function, one should calculate a new vector of polarization $\mathbf{P}_{j+1}$ by the preceding vector $\mathbf{P}_j$ for each scattering event:

$$\mathbf{P}_{j+1} = (\hat{I} - (\mathbf{R}_{j+1} - \mathbf{R}_j) \otimes (\mathbf{R}_{j+1} - \mathbf{R}_j)|\mathbf{R}_{j+1} - \mathbf{R}_j|^{-2})\mathbf{P}_j.$$

Suppose that a photon experiences *n* scattering events. Then, after the last, *n*th scattering event, a photon arrives at the observation point $\mathbf{R}_D$ with the polarization vector

$$\mathbf{P}^{(out)} = \prod_{j=1}^{n}(\hat{I} - (\mathbf{R}_{j+1} - \mathbf{R}_j) \otimes (\mathbf{R}_{j+1} - \mathbf{R}_j) \times |\mathbf{R}_{j+1} - \mathbf{R}_j|^{-2})\mathbf{P}^{(in)}. \quad (4.2)$$

Let W_i be the statistical weight of a "scalar" *i*th photon that arrives at the point $\mathbf{R}_D$. Then, summing over all N_{ph} detected photons, we obtain the following expressions for the polarized and depolarized components (for short, we omit the index "out"):

$$I_{pol} = I_{XX} = \sum_{i=1}^{N_{ph}} W_i P_{ix}^2 \Gamma_R^{n_i},$$
$$I_{depol} = I_{YX} = \sum_{i=1}^{N_{ph}} W_i P_{iy}^2 \Gamma_R^{n_i}. \quad (4.3)$$

In the case of backward scattering, there is no *z* component. Note that these formulas describe the noncoherent contribution of ladder diagrams, $I_{\beta\alpha} = C^{(L)}_{\beta\beta\alpha\alpha}(0|\mathbf{k}_s, \mathbf{k}_i)$.

In the case of an electromagnetic field, the polarization vector strongly fluctuates even for very large (on the order of 10^5) statistical sample sizes. In [24], the authors analyzed the rate of depolarization as a function of the number *n* of scattering events in the diffusive approximation. According to [24], in the case of an isotropic single-scattering cross section, the depolarization after *n* scattering events is given by

$$P(n) = \frac{I_{pol}(n) - I_{depol}(n)}{I_{pol}(n) + I_{depol}(n)} = \frac{3(0.7)^{n-1}}{2 + (0.7)^{n-1}}.$$

Figure 4 represents the calculated function $P(n)$. Note that the number of scattering events is actually proportional to the length of the optical path; since the length of the path is proportional to the flight time, the function $P(n)$ illustrates the spreading of a light impulse in a strongly inhomogeneous medium. One can see that the depolarization indeed decays exponentially as the optical path increases; however, the decay rate differs

from that predicted by the diffusive approximation. As the anisotropy increases, the characteristic length of depolarization increases, because, for large values of $\overline{\cos\theta}$, a photon must experience $(1 - \overline{\cos\theta})^{-1}$ times greater number of collisions compared with the isotropic case to appreciably change its direction and, hence, the polarization. For $n > 30$, the results are quite unstable. This is associated with the fact that the products of an odd number of components of the polarization vector directed along the boundary of the medium are equal to zero theoretically, whereas the stochastic result strongly fluctuates.

The calculations are performed as follows. For trajectories with a number of scattering events less than a certain number n_0, which varied from 20 to 50, the calculation was performed by the above-described scheme. We assumed that, for $n > n_0$, the polarized and depolarized components are equal and, hence, can be calculated as half the intensity calculated for a scalar field:

$$I_{\text{pol}}(n) = I_{\text{depol}}(n) = \frac{1}{2} I_{\text{scalar}}(n).$$

To control the process, we compared the results of simulation in the case of point scatterers with the known exact theoretical results.

In the case of Rayleigh scattering for the normal incidence and scattering at an angle of 180°, the exact solution [21, 22] obtained within a generalization of the Milne solution for an electromagnetic field yields a value of $I_{\text{pol}}/I_{\text{depol}} \approx 1.92$ for the ratio of polarized to depolarized components of the noncoherent component. The numerical simulation yields $I_{\text{pol}}/I_{\text{depol}} \approx 1.94$. The known value of the ratio of the polarized component of scattered radiation to the depolarized component allows one to determine the residual polarization of the noncoherent component of backscattering. The generalized Milne solution yields [19, 21, 22]

$$\frac{I_{\text{pol}} - I_{\text{depol}}}{I_{\text{pol}} + I_{\text{depol}}} \approx 0.31,$$

while numerical simulation yields 0.326. A close value of 0.33 was obtained in [23].

Let us define the height of the peak of the polarized component of coherent backscattering as

$$h_{\text{pol}}^{\text{CBS}} = \frac{2I_{\text{pol}} - I_{\text{single}}}{I_{\text{pol}}}.$$

The theoretical value [19, 21, 22] is $h_{\text{pol}}^{\text{CBS}} \approx 1.75$. The analysis of simulation data obtained in [2] yields a value of $h_{\text{pol}}^{\text{CBS}} \approx 1.4$, which is far from the theoretically predicted result. In [29], the authors obtained a value of $h_{\text{pol}}^{\text{CBS}} \approx 1.69$. The value $h_{\text{pol}}^{\text{CBS}} \approx 1.746$ calculated by us agrees much better with theory.

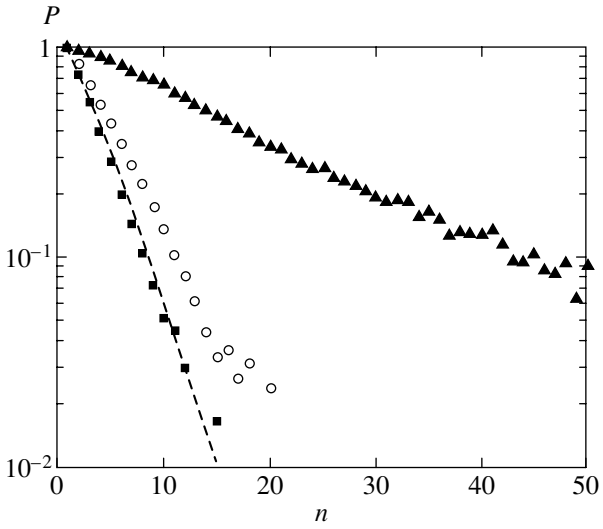

Fig. 4. Depolarization degree P as a function of the number of scattering events n for $\overline{\cos\theta} = 0$ (■), 0.5 (○), and 0.9 (▲). A semi-infinite layer. The straight line represents the diffusive approximation.

We calculated the time correlation function for the polarized and depolarized components of backscattered light by numerical simulation.

To calculate the time correlation functions of an electromagnetic field, we used the following formulas:

$$g_{\text{pol}}^{(1)}(t) = \sum_{i=1}^{N_{\text{ph}}} W_i P_{ix}^2 \Gamma_{\text{R}}^{n_i} \times \exp\left(-2\frac{t}{\tau} n_i \left(1 - \frac{1}{n_i}\sum_{j}^{n_i} \cos\theta_j\right)\right),$$
$$g_{\text{depol}}^{(1)}(t) = \sum_{i=1}^{N_{\text{ph}}} W_i P_{iy}^2 \Gamma_{\text{R}}^{n_i} \times \exp\left(-2\frac{t}{\tau} n_i \left(1 - \frac{1}{n_i}\sum_{j}^{n_i} \cos\theta_j\right)\right), \quad (4.4)$$

where $P_{i\alpha}$ is the polarization vector of the ith photon with polarization α that arises under the action of a sequence of n_i tensor operators of the form (4.1) and θ_j is the scattering angle of the jth scattering event.

Figure 5 represents the results of simulation of the time correlation function of field for the polarized and depolarized components for the Rayleigh scattering, $\overline{\cos\theta} = 0$, and for the case of a strongly anisotropic sys-

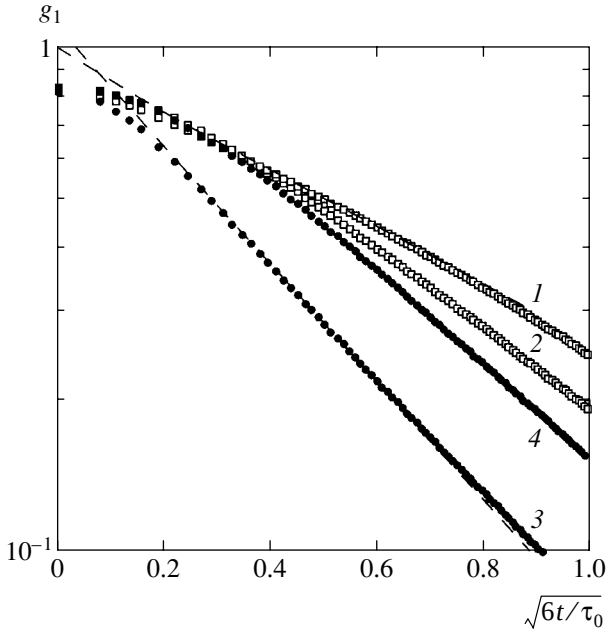

Fig. 5. Time correlation function for electromagnetic field in scattering media $g_{XX}^{(1)}$ (□), $g_{XY}^{(1)}$ (●). Curves *1* and *3* correspond to $\overline{\cos\theta} = 0$, curves *2* and *4* correspond to $\overline{\cos\theta} = 0.9$ and brocken curves determine the slopes $\gamma_{pol} \approx 1.42$ and $\gamma_{depol} \approx 2.68$ for a medium with isotropic phase function.

tem with the indicatrix stretched forward, $\overline{\cos\theta} = 0.9$. One can see that, in the case of linearly polarized light, the decay rate of the time correlation function in the units of τ appreciably depends on the anisotropy parameter $\overline{\cos\theta}$, in contrast to nonpolarized light.

The diffusive character of light propagation in the multiple scattering regime is responsible for the linear decay of the field time correlation function of the form [15]

$$g^{(1)}(t) \sim 1 - \gamma\sqrt{6t/\tau},$$

where γ is the slope ratio that determines the decay rate of the time correlation function. Note that, according to (3.5), the initial small parameter in time is $(t/\tau)(1 - \overline{\cos\theta})$. In the case of a strongly anisotropic scattering cross section, this quantity remains small for sufficiently large values of the parameter t/τ for which the dependence of $g^{(1)}(t)$ on $\sqrt{t/\tau}$ may strongly differ from a linear function.

In the case of Rayleigh scattering, the theory predicts [19, 22] the following slope ratios for the polarized and depolarized components: $\gamma_{pol} \approx 1.44$ and $\gamma_{depol} \approx 2.75$. The analysis within the diffusive approximation yields [16] $\gamma_{pol} \approx 1.6$ and $\gamma_{depol} \approx 2.7$, which are close to the experimental data $\gamma_{pol} \approx 1.6 \pm 0.1$ and $\gamma_{depol} \approx 2.8 \pm 0.2$. These experimental data were obtained for a suspension of latex particles with a diameter of $D = 0.091$ μm, which is much less than the wavelength; i.e., the scattering is close to the Rayleigh scattering.

Note that, just as in a real experiment [15], one cannot eliminate the nonlinear region for very small values of time, which is attributed to the finiteness of the aperture and the finiteness of the number of scattering events considered; the theoretically predicted linear region with the slope ratios $\gamma_{pol} \approx 1.42$ and $\gamma_{depol} \approx 2.68$ starts at $\sqrt{t/\tau} = 0.15$. As the anisotropy parameter increases, the decay rates of the polarized and depolarized components approach each other. The sum of the polarized and depolarized components, i.e., the time correlation function of nonpolarized light is described by a curve close to the curve obtained during the simulation of the time correlation function of a scalar field.

CONCLUSIONS

In the theory of coherent and correlation phenomena of multiple scattering, both the cyclic diagrams that describe the interference component of backscattering [33, 24] and the diagrams that describe the field correlations [34] can be reduced to ladder diagrams. This allows one to describe the above phenomena within the framework of the relevant Bethe–Salpeter equation. The formal difference from the original ladder diagrams that describe the transfer of the main, noncoherent, component of scattered radiation consists in the introduction of additional coefficients to the vertices of ladder diagrams. These coefficients describe the phase shift between the fields that appear in the definition of the propagator of the Bethe–Salpeter equation. In the present paper, by juxtaposing the representation of the Bethe–Salpeter equation as a ladder diagram series with the modeling of random trajectories, we have shown that the consideration of phase relations in stochastic modeling also reduces to the addition of appropriate coefficients in each scattering event experienced by a photon as it moves along a random trajectory.

The semianalytic Monte Carlo method developed allows one to compare numerical results with theoretical predictions at each step of simulation. The possibility of such a comparison allows one to significantly reduce the simulation time by using analytic results instead of numerical ones at large distances between the input and output points of radiation, where theoretical predictions are certainly correct.

The analysis carried out has shown that scattering is indeed of multiple character only in the case of a nonabsorbing semi-infinite medium. In all other cases, each scattering event gives rise to coefficients of the form $D_s\overline{q^2}\,t = 2(t/\tau)(l/l^*)$ when simulating the time correlation functions with diffusive decay of fluctuations or $kl\sin\theta_s$ for coherent backscattering; these coeffi-

cients lead to a rapid decay of a wave packet. The decay rate of coherent phenomena is determined by the parameters (t/τ) and $kl^*\sin\theta_s$, which may be substantially greater than the original parameters $(t/\tau)(l/l^*)$ and $kl\sin\theta_s$ in the case of a strongly anisotropic phase function, $1 - \overline{\cos\theta_s} \ll 1$.

The simulation error is primarily determined by the sample size, i.e., by the number N_{ph} of emitted photons and by the maximal number $n^{(max)}$ of the scattering events taken into consideration. We continued to follow up the trajectory of a photon up to 10^4 scattering events. In this case, the results become stable to within one percent even for about 1000 emitted photons; for $N_{ph} \sim 10^5$, the relative error in the scattering intensity is on the order of 10^{-4}. However, in view of the diffusive character of the propagation of radiation in an unbounded medium, the contributions of the scattering events of extremely high orders still continue to contribute when one simulates the intensity. Therefore, the statistics of photons with high-order scattering proves to be insufficient at large distances from the input point, and the result depends on the choice of the number $n^{(max)}$. Comparing, in the isotropic case, numerical results with the theoretical results that follow from the Milne solution and its generalizations, we have found that the abovementioned restriction on the number of scattering events gives an understated result for the intensity of scattered radiation: on the order of 2% for a scalar field and on the order of 5% for an electromagnetic field, even for sample sizes of up to several millions. It is this circumstance why we complement stochastic modeling with analytic calculations, applying analytic calculations within the diffusive approximation instead of modeling in the case of large distances between the input and output points (or when the number of scattering events exceeds a certain fixed number); the accuracy of these analytic calculations are easily controlled by increasing the above-mention fixed parameter.

In the case of modeling the time correlation function or the peak of coherent backscattering for the parameters $(t/\tau)(l/l^*)$ and $kl^*\sin\theta_s$ on the order of 0.01 and greater, an additional contribution (combined with the use of analytic calculations) of scattering events of extremely high orders barely affects the result. In the range of large values of $(t/\tau)(l/l^*)$ and $kl^*\sin\theta_s$ on the order of unity, the relative contribution increases again; however, in this range of parameters, the very physical model and, in particular, the ladder approximation cease to be valid.

The description of multiple scattering, including coherent phenomena, in terms of ladder diagram series is valid up to the parameter λ/l. Thus, when the time correlation function decreases as time increases or the peak of coherent backscattering decreases as the scattering angle increases by a factor greater than l/λ, these phenomena are no longer observable in real experiments against the background of contributions made by "nonladder" diagrams that take into account the phase shifts of fields in multiple rescattering processes.

Usually, multiple scattering, including coherent phenomena, is described in terms of a scalar field. Our results show that, for backward scattering, consideration of the electromagnetic nature of light leads to an essentially different quantitative description compared with that in the case of a scalar field because a considerable part of backward scattered radiation consists of low-order contributions. For example, the decay rate of the time correlation function of the polarized component is much less, while that of the depolarized component is much greater, than the decay rate in the case of nonpolarized light; the latter decay rate virtually coincides with that in the scalar case. The polarized component of backscattered light is almost twice as large as the depolarized component.

This method of numerical simulation allows one to judge the number of scattering events experienced by light transmitted through a layer of a strongly inhomogeneous opaque medium by the value of residual polarization. This information is additional to that obtained from the measurements of attenuation of nonpolarized light, from which one derives the transport mean free path. Comparative analysis makes it possible to significantly simplify the simulation of radiative transfer and coherent phenomena in randomly inhomogeneous strongly scattering media, such as liquid crystals, biological tissues, etc., and to significantly extend the application domain of these methods.

The results extend the applicability domain of the methods based on the coherent and correlation properties of diffusely scattered light.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 02-02-16577), the Royal Society (project no. 15298), and NATO (project no. PST.CLG.979652).

REFERENCES

1. M. Ospeck and S. Fraden, Phys. Rev. E **49**, 4578 (1994).
2. T. Iwai, H. Furukawa, and T. Asakura, Opt. Rev. **2**, 413 (1995).
3. K. Ishii, T. Iwai, and T. Asakura, Opt. Rev. **4**, 643 (1997).
4. S. E. Skipetrov and S. S. Chesnokov, Kvantovaya Élektron. (Moscow) **25**, 753 (1998).
5. R. Lenke and G. Maret, Eur. Phys. J. B **17**, 171 (2000).
6. S. E. Skipetrov and I. V. Meglinskiĭ, Zh. Éksp. Teor. Fiz. **113**, 1213 (1998) [JETP **86**, 661 (1998)].
7. R. Lenke, R. Tweer, and G. Maret, J. Opt. A: Pure Appl. Opt. **4**, 293 (2002).
8. D. A. Zimnyakov, Yu. P. Sinichkin, I. V. Kiseleva, and D. N. Agafonov, Opt. Spektrosk. **92**, 831 (2002) [Opt. Spectrosc. **92**, 765 (2002)].

9. V. L. Kuz'min and I. V. Meglinskiĭ, Pis'ma Zh. Éksp. Teor. Fiz. **79**, 139 (2004) [JETP Lett. **79**, 109 (2004)].
10. V. L. Kuz'min and I. V. Meglinskiĭ, Opt. Spektrosk. **97**, 108 (2004) [Opt. Spectrosc. **97**, 100 (2004)].
11. B. A. van Tiggelen and S. E. Skipetrov, *Wave Scattering in Complex Media: From Theory to Applications* (Kluwer Academic, Dordrecht, 2003).
12. M. C. W. van Rossum and Th. N. Nieuwenhuizen, Rev. Mod. Phys. **71**, 313 (1999).
13. I. M. Sobol', *The Monte Carlo Method* (Nauka, Moscow, 1985) [in Russian].
14. P. E. Wolf and G. Maret, Phys. Rev. Lett. **55**, 2696 (1985).
15. D. J. Pine, D. A. Weitz, P. M. Chaikin, and E. Herbolzheimer, Phys. Rev. Lett. **60**, 1134 (1988).
16. F. C. MacKintosh and S. John, Phys. Rev. B **40**, 2383 (1989).
17. M. I. Mishchenko, Phys. Rev. B **44**, 12597 (1991).
18. M. I. Mishchenko, J. Quant. Spectrosc. Radiat. Transf. **56**, 673 (1996).
19. E. Amic, J. M. Luck, and T. M. Nieuwenhuizen, J. Phys. I **7**, 445 (1997).
20. M. I. Mishchenko, J. M. Luck, and T. M. Nieuwenhuizen, J. Opt. Soc. Am. A **17**, 888 (2000).
21. V. L. Kuz'min, Opt. Spektrosk. **93**, 482 (2002) [Opt. Spectrosc. **93**, 439 (2002)].
22. V. L. Kuz'min and E. V. Aksenova, Zh. Éksp. Teor. Fiz. **123**, 929 (2003) [JETP **96**, 816 (2003)].
23. L. F. Rojas-Ochoa, D. Lacoste, R. Lenke, *et al.*, J. Opt. Soc. Am. A **21**, 1799 (2004).
24. E. Akkermans, P. E. Wolf, R. Maynard, *et al.*, J. Phys. (Paris) **49**, 77 (1988).
25. Yu. N. Barabanenkov and V. D. Ozrin, Zh. Éksp. Teor. Fiz. **94** (6), 56 (1988) [Sov. Phys. JETP **67**, 1117 (1988)].
26. I. V. Meglinskiĭ and S. J. Matcher, Opt. Spektrosk. **91**, 692 (2001) [Opt. Spectrosc. **91**, 654 (2001)].
27. D. Y. Churmakov, I. V. Meglinski, and D. A. Greenhalgh, Phys. Med. Biol. **47**, 4271 (2002).
28. A. Ishimaru, *Wave Propagation and Scattering in Random Media* (Academic, New York, 1978; Mir, Moscow, 1981).
29. R. Lenke and G. Maret, Eur. Phys. J. B **17**, 171 (2000).
30. E. Tinet, S. Avrillier, and J. M. Tualle, J. Opt. Soc. Am. A **13**, 1903 (1996).
31. G. I. Marchuk, G. A. Mikhailov, M. A. Nazaraliev, *et al.*, *The Monte-Carlo Methods in Atmospheric Optics* (Springer, Berlin, 1980).
32. G. Maret and P. E. Wolf, Physica B (Amsterdam) **65**, 409 (1987).
33. A. Golubentsev, Zh. Éksp. Teor. Fiz. **86**, 47 (1984) [Sov. Phys. JETP **59**, 26 (1984)].
34. M. J. Stephen, Phys. Rev. B **34**, 7564 (1986).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 33–43.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 42–53.

ATOMS, MOLECULES, OPTICS

Multiparticle Entanglement via Resonant Interaction between Light and Atomic Ensembles

V. N. Gorbachev* and A. I. Trubilko**

St. Petersburg State University of Aerospace Instrumentation, St. Petersburg, 190000 Russia
**e-mail: vn@vg3025.spb.edu*
***e-mail: Tai@at3024.spb.edu*

Received October 2, 2004

Abstract—Multiparticle entangled states generated via interaction between narrowband light and an ensemble of identical two-level atoms are considered. Depending on the initial photon statistics, correlation between atoms and photons can give rise to entangled states of these systems. It is found that the state of any pair of atoms interacting with weak single-mode squeezed light is inseparable and robust against decay. Optical schemes for preparing entangled states of atomic ensembles by projective measurement are described. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Entangled states are used as a key resource in various quantum information processing systems. Their properties and preparation methods have been the subject of extensive discussion. The desired state of a physical system can be prepared either by projective measurement or as a result of evolution. For atomic systems, both methods have already been implemented in experiments. In particular, two atomic ensembles were used in [1] to create an Einstein–Podolsky–Rosen (EPR) pair by projective measurement. The latter method was demonstrated in several studies: entangled states of alkali ions were generated via Coulomb interaction [2], neutral Rydberg atom were used to create an EPR pair in a micromaser setup [3], and resonant dipole–dipole interaction was used for entangling neutral atoms in an optical lattice [4]. The most popular methods for preparing entangled photon states are still mostly based on parametric down-conversion. For example, an entangled state equivalent to a three-state quantum system (qutrit) was prepared and examined by using quantum state tomography in [5]. These examples suggest that an entangled state of two systems can be prepared experimentally by using a certain interaction. Systems of this kind are well studied. With regard to applications, it is important to know how entanglement can be utilized and to know its robustness against decoherence. In this respect, of special interest are multiparticle systems whose entangled states are characterized by much more complicated and diverse behavior.

Previous efforts were mainly focused on analysis of entanglement between several particles. In particular, the *W* class of tripartite entanglement defined in [6] includes the symmetric three-photon polarization-entangled state implemented in the experiment reported in [7]. An extension to four qubits was proposed in [8], where nine inequivalent classes were distinguished that cannot be connected by local operations and quantum communication. Studies of multiparticle systems are relatively few, being focused on entanglement criteria and application to problems in quantum information theory. Whereas the Peres–Horodecki criterion for bipartite entanglement found in [9] was applied to a real physical system in [10], no operational criterion is known for entanglement in the general case; various approaches are used. In [11], the concept of *entanglement molecules* [12] was used to propose a classification using graphs, with particles and classical or quantum correlations represented, respectively, by vertices and edges connecting pairs of vertices. Graphs of this kind can be used to describe both pure and mixed entangled states and distinguish several classes differing by topological properties of the graphs. In [13], symmetric states (including Dicke states) were studied by using several entanglement measures (entropy of entanglement, negativity, and entanglement of formation) defined by the eigenvalues of a partial transpose of the density matrix. A numerical analysis was performed to find that symmetric states are robust to particle loss even if the number of particles is large (up to 10^3). Note that the calculation of eigenvalues is a difficult task, because the dimension of an ensemble's Hilbert space exponentially increases with the number of constituent particles. Owing to their robustness, symmetric states can be used in such applications as cloning and telecloning protocols for quantum information transmission [14], quantum key distribution [15], and quantum teleportation or dense coding [16]. The formulation of two models of a one-way quantum computer using measurements on multiparticle entangled states [17,

18] has strongly stimulated studies of the properties of multiparticle systems, in particular, Ising- and Bose–Hubbard-like models.

The present study focuses on the Dicke states arising as a result of collective interaction of many atoms with electromagnetic field [19], which has been analyzed in numerous studies (e.g., see [20]). This system exhibits many physical properties of interest for quantum information processing. Photon trapping in chain configurations of atoms was considered in [21]. When the system is placed in a cavity, this effect reduces the photon escape rate and increases the decoherence time of the cavity mode. In [22], this effect was used for generating *W* states and anticloning [23], which can be implemented with high fidelity by means of photon trapping. In those studies, only single-photon traps and single-photon initial states were analyzed. Here, we consider the more general case of multiphoton processes, assuming that the photon statistics is arbitrary.

The main questions addressed below are the following: What types of entangled states are produced by interaction between atoms and field? What states can be prepared from independent atomic ensembles entangled with a photon? How can these states be utilized? We consider resonant interaction between narrowband light and an ensemble of identical two-level atoms coupled to a common heat bath. The analysis is restricted to a simple model of radiative decay. Multiphoton processes, such as Raman scattering, are described in terms of effective Hamiltonians, which can be obtained by unitary transformation [24]. The behavior of an atomic system interacting with light characterized by arbitrary photon statistics is analyzed by using perturbation theory in the interaction strength for Gaussian, coherent, and squeezed states. We find that weak single-mode squeezed light is required to create multiparticle entanglement between atoms. As distinct to the case considered in [25], the steady state discussed here is robust against atomic decay. When decay is neglected and analysis is restricted to a single-photon initial state, simple exact solutions describing exchange of excitation between the field mode and atoms can be obtained [26]. These solutions can be used for generating and transforming symmetric Dicke states and for processing and storing quantum information. The optical schemes for projective measurement considered here can be used to generate entangled states of atomic ensembles. An EPR entangled pair of macroscopic ensembles was created in an experiment [1]. The new states produced in our schemes have hierarchical structure, thus differing from the cluster states introduced in [27] as a resource for one-way computing.

The paper is organized as follows. First, we formulate a basic model and write out the second-order perturbation solutions obtained by taking into account radiative decay. These solutions are then used to analyze the states of the atomic system corresponding to various photon statistics. Exact solutions obtained under certain initial conditions by neglecting radiative decay are used to describe generation and transformation of symmetric Dicke states. Finally, we consider optical schemes for preparing entangled states of atomic ensembles by projective measurement.

2. BASIC EQUATIONS

In the dipole approximation, the ensemble of N identical, but distinguishable, two-level atoms interacting with electromagnetic field is described by the Hamiltonian

$$H = i\hbar^{-1}\vartheta,$$

$$\vartheta = \sum_k g_k a_k S_k^\dagger - \text{H.c.},$$

where

$$g_k = \left(\frac{\hbar\omega_k}{2\varepsilon_0 L^3}\right)^{1/2} \boldsymbol{\mu}\cdot\mathbf{e}_k$$

is the coupling constant, $\boldsymbol{\mu}$ is the dipole transition matrix element, $\mathbf{e}_k$ is the polarization vector for the mode with wavevector k, a_k and $a_k^\dagger$ are photon creation and annihilation operators,

$$S_k^\dagger = \sum_a s_{10}(a)\exp(ikr_a)$$

is the collective atomic operator, $s_{xy} = |x\rangle_a\langle y|$ is the atomic operator for the atom located at a point r_a (x, y = 0, 1, where 0 and 1 denote the ground and excited states, respectively). When analysis is restricted to interaction with a single resonant mode, S_k can be replaced with $S_{k=0}$, which makes it possible to treat an atomic ensemble occupying a spatial region as a point-like object. Then,

$$\vartheta = S_{10}B - S_{01}B^\dagger, \tag{1}$$

where

$$S_{10} = \sum_a |1\rangle_a\langle 0|, \quad B = ga.$$

Effective Hamiltonian (1) is used here to describe not only interaction with a single resonant mode, but also multiphoton processes, such as Raman scattering. In the latter case, we set $B = fa_A a_S^\dagger$ and assume that the photon frequencies ω_A and ω_S satisfy the relation $\omega = \omega_A - \omega_S$, where ω is the atomic transition frequency.

Hamiltonians of this form can be obtained by unitary transformations [24].

The density matrix ρ for the N-atom system interacting with a cavity mode obeys the master equation

$$\frac{\partial}{\partial t}\rho = [\vartheta, \rho] + \mathcal{L}\rho, \tag{2}$$

where relaxation is represented by the Lindblad superoperator

$$\mathcal{L} = \sum_a \mathcal{L}_a,$$

$$\begin{aligned}\mathcal{L}_a &= -\frac{\gamma_\uparrow}{2}[s_{01}(a)s_{10}(a)\rho - s_{10}(a)s_{01}(a)] \\ &-\frac{\gamma_\downarrow}{2}[s_{10}(a)s_{01}(a)\rho - s_{01}(a)\rho s_{10}(a)] + \text{H.c.}\end{aligned} \tag{3}$$

This representation corresponds to the model of purely radiative decay with longitudinal and transverse decay rates $\gamma = \gamma_\downarrow + \gamma_\uparrow$ and $\gamma_\perp$, which satisfy the relation $\gamma_\perp = \gamma/2$. In the general case, $\gamma_\perp > \gamma/2$, since $\gamma_\perp$ should be replaced by $\gamma_\perp + \kappa$, where κ is a dephasing collision rate.

Effective Hamiltonian (1) may involve many field modes with ω_k differing from the atomic transition frequency by $\delta\omega_k$ and occupying a frequency band of width $\Delta\omega$. If $\Delta\omega, \delta\omega_k \ll \gamma_\perp$, then we can consider a narrow-band radiation field and make use of resonance approximation. Otherwise, the field must be described in terms of multiple-time correlation functions.

Solution of Eq. (2) is a difficult task. To describe the interaction between an individual atom and a cavity mode, the following equation for the density matrix $\rho_a = \mathrm{Tr}'_a\rho$ is derived from (2) by taking the trace Tr'_a over all atoms except for one:

$$\frac{\partial}{\partial t}\rho_a = [\vartheta_a, \rho_a] + \mathcal{L}_a\rho_a + N(N-1)\mathrm{Tr}'_a[\vartheta_{a'}, \rho_{aa'}], \tag{4}$$

where $\vartheta_a = s_{10}(a)B - \text{H.c.}$ and $\rho_{aa'} = \mathrm{Tr}'_{aa'}\rho$ is a two-particle density matrix. The right-hand side of (4) contains a multiparticle contribution proportional to $N(N-1)$, because the density matrix $\rho_{aa'}$ does not commute with the field operators. This leads to the Bogolyubov–Born–Green–Kirkwood–Yvon chain of equations for the multiparticle density matrices ρ_a, $\rho_{aa'}$, $\rho_{aa'a''}$, …. In physical terms, this means that fluctuations of quantized electromagnetic field induce correlation between atoms. If the field is assumed to be classical and noise-free, for example, a coherent state is considered, then the interaction will not give rise to any correlation, and the initially uncorrelated atoms will remain mutually independent. In what follows, we use (2) to analyze interactions that can be used to generate symmetric Dicke states.

3. DICKE STATES

First, we define symmetric Dicke states and introduce a representation of symmetric Dicke states that demonstrates their relation to the collective interaction processes. The Dicke states are eigenstates of the operators J_z and $J^2 = J_x^2 + J_y^2 + J_z^2$:

$$\begin{aligned}J_z|jma\rangle &= m|jma\rangle, \\ J^2|jma\rangle &= j(j+1)|jma\rangle,\end{aligned} \tag{5}$$

where

$$J_s = \frac{1}{2}\sum_k \sigma_{sk},$$

σ_{sk} ($s = x, y, z$) are the Pauli operators for a two-level atom labeled by the index k ($k = 1, \ldots, N$), and j and m are integer and half-integer numbers such that $|m| \le j$ and $\max j = N/2$. If $j = N/2$, then the states are symmetric, and the quantum number a introduced to lift degeneracy can be omitted. For h excited atoms ($h = m + N/2$), the states can be represented as

$$\begin{aligned}|j = N/2;\, m\rangle &\equiv |h;\, N\rangle \\ &= \sum_z P_z|1_1, 1_2, \ldots, 1_h, 0_{h+1}, \ldots, 0\rangle,\end{aligned} \tag{6}$$

where P_z is one of the

$$C_h^N = \frac{N!}{h!(N-h)!}$$

distinguishable permutations of particles. The vector $|h;\, N\rangle$ represents an atomic ensemble of h excited atoms normalized by the condition $\langle h;\, N|h,\, N\rangle = C_h^N$. Symmetric states of a multiparticle system arise when interaction is described by collective operators of the form

$$S_{10} = \sum_a^N |1\rangle_a\langle 0|.$$

In particular, the following representation is valid:

$$|h;\, N\rangle = \frac{1}{h!}S_{10}^h|0;\, N\rangle. \tag{7}$$

If $h = 1$, then it holds that

$$|1;\, N\rangle = |10\ldots0\rangle + \ldots + |00\ldots1\rangle. \tag{8}$$

Since the wavefunction $|h; N\rangle$ is not factorizable, it represents an entangled state. In terms of correlation between particles, it is substantially different from other entangled states. For example, in the Greenberger–Horne–Zeilinger (GHZ) state

$$|\mathrm{GHZ}\rangle = \frac{1}{2}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}),$$

the correlation of any group of M particles ($M < N$) is classical. In particular, the density matrix corresponding to the state $|1; N\rangle$ of a group of particles ($M \le N$) is

$$\rho(M \le N) = \frac{1}{N}|1; M\rangle\langle 1; M| + \frac{N-M}{N}|0; N\rangle\langle 0; N|.$$

The corresponding von Neumann entropy depends on the relative particle number $p = M/N$:

$$S(\rho(M \le N)) = -p\log_2 p - (1-p)\log_2(1-p).$$

When $p = 1/2$, the entropy has a maximum value of unity. If $M = 2$, we can apply the necessary and sufficient separability criterion proposed in [9]. According to this criterion, the state is inseparable (entangled) if the transpose of the density matrix with respect to the variables associated with one of the atoms has at least one negative eigenvalue. In the case considered here, one of the four eigenvalues

$$\left\{\frac{1}{N}; \frac{1}{N}; \frac{N-2}{2N}; \left[1 \pm \sqrt{1 + \frac{4}{(N-2)^2}}\right]\right\}$$

is negative. Note that the behavior of correlation between M particles depends on $p = M/N$. As the total particle number N increases, $p \longrightarrow 0$ and the correlation vanishes, since their state becomes pure as $\rho(M \le N) \longrightarrow |0; N\rangle\langle 0; N|$. In what follows, we make use of the following equalities:

$$\begin{aligned} S_{01}|0; N\rangle &= 0, \\ S_{10}|h; N\rangle &= (h+1)|h+1; N\rangle, \\ S_{01}|h; N\rangle &= (N-h+1)|h-1; N\rangle, \\ S_{01}S_{10}|h; N\rangle &= (h+1)(N-h)|h; N\rangle, \\ S_{10}S_{01}|h; N\rangle &= h(N-h+1)|h; N\rangle. \end{aligned} \tag{9}$$

4. SECOND-ORDER PERTURBATION SOLUTIONS

To solve Eq. (2), we use perturbation theory in the interaction strength:

$$\rho = \rho^{(0)} + \rho^{(1)} + \rho^{(2)} + \ldots. \tag{10}$$

Here, the zeroth-order approximation $\rho^{(0)}$ is the time-independent solution to (2) with $\vartheta = 0$: $\rho^{(0)} = |0\rangle\langle 0| \otimes \rho_f$, where the density matrix ρ_f represents the cavity mode and $|0\rangle = |0\rangle^{\otimes N}$ corresponds to the ground state of all atoms. The operators $\rho^{(k)}$ ($k = 1, \ldots$) satisfy the equations

$$\frac{\partial}{\partial t}\rho^{(k)} = [\vartheta, \rho^{(k-1)}] + \mathscr{L}\rho^{(k)} \tag{11}$$

subject to the initial conditions $\rho^{(k)}(0) = 0$.

The analysis that follows is restricted to second-order perturbation theory, which is sufficient to obtain statistical characteristics of the excitation field. The matrix equation for $\rho^{(2)}$ is

$$\begin{aligned} &\langle 1_k; 1_m; N|\frac{\partial}{\partial t}\rho^{(2)}|0; N\rangle \\ &= -2\gamma_\perp\langle 1_k; 1_m; N|\rho^{(2)}|0; N\rangle + \langle 1_k; 1_m; N|R|0; N\rangle, \\ &\langle 1_k; N|\frac{\partial}{\partial t}\rho^{(2)}|1_m; N\rangle \\ &= -2\gamma_\perp\langle 1_k; N|\rho^{(2)}|1_m; N\rangle \\ &+ \langle 1_k; N|R|1_m; N\rangle, \quad k \ne m, \\ &\langle 1_k; N|\frac{\partial}{\partial t}\rho^{(2)}|1_k; N\rangle = -\gamma\langle 1_k; N|\rho^{(2)}|1_k; N\rangle \\ &+ \langle 1_k; N|R|1_k; N\rangle, \\ &\langle 0; N|\frac{\partial}{\partial t}\rho^{(2)}|0; N\rangle = \gamma\sum_k\langle 1_k; N|\rho^{(2)}|1_k; N\rangle \\ &+ \langle 0; N|R|0; N\rangle, \end{aligned} \tag{12}$$

where the vectors

$$\begin{aligned} s_{10}(k)|0; N\rangle &= |1_k; N\rangle, \\ s_{10}(k)s_{10}(p)|0; N\rangle &= |1_k, 1_p; N\rangle \end{aligned}$$

represent the states in which only the kth atom is excited and only the kth and pth atoms are excited, respectively. The nonzero matrix elements of the operator $R = [\vartheta, \rho^{(1)}]$ are

$$\begin{aligned} \langle 1_k, 1_m; N|R|0; N\rangle &= 2\kappa(t)B^2\rho_f, \\ \langle 0; N|R|0; N\rangle &= -\kappa(t)N(B^\dagger B\rho_f + \rho_f B^\dagger B), \\ \langle 1_k; N|R|1_m; N\rangle &= 2\kappa(t)B\rho_{sf}B^\dagger, \end{aligned} \tag{13}$$

where

$$\kappa(t) = \frac{1}{\gamma_\perp}(1 - \exp(-\gamma_\perp t)).$$

For purely radiative decay, $\gamma_\perp = \gamma/2$ and the second-order perturbation theory yields

$$\begin{aligned}\rho &= |0\rangle\langle 0| \otimes \rho_f + \kappa[|1; N\rangle\langle 0; N| \otimes B\rho_{sf} + \text{H.c.}] \\ &+ \kappa^2[|2; N\rangle\langle 0; N| \otimes B^2\rho_f + \text{H.c.}] \\ &- N\gamma\mathcal{H}|0; N\rangle\langle 0; N| \otimes [B^\dagger B\rho_f - B\rho_f B^\dagger + \text{H.c.}] \\ &- (1/2)N\kappa^2|0; N\rangle\langle 0; N| \otimes [B^\dagger B\rho_f + \text{H.c.}] \\ &+ \kappa^2|1; N\rangle\langle 1; N| \otimes B\rho_f B^\dagger,\end{aligned} \quad (14)$$

where

$$\mathcal{H} = \frac{1}{\gamma_\perp}\left\{\frac{1}{\gamma^2}[\gamma t + 1 - \exp(-\gamma t)] - \frac{\kappa^2}{2}\right\}.$$

This expression is valid to second order if the field is relatively weak:

$$N\kappa^2\langle B^\dagger B\rangle \ll 1. \quad (15)$$

In the case of interaction with a single resonant cavity mode, we have $B = ga$ and $\kappa^2\langle B^\dagger B\rangle = n/n_s$, where $n_s = (\gamma_\perp/g)^2$ is a saturation parameter and $n = \langle a^\dagger a\rangle$ is the mean photon number. Then, (15) reduces to the standard condition imposed in the case of resonant coupling between the field and two-level atoms: $Nn/n_s \ll 1$. Solution (14) describes the joint evolution of the atomic ensemble and field starting from an ensemble of ground-state atoms and an arbitrary state of the field.

5. MIXED ENTANGLED ATOMIC STATES

Second-order perturbation theory predicts correlation between atoms depending on photon statistics, i.e., providing a framework for describing entangled (inseparable) atomic states. To analyze the properties of the atomic system, we use second-order perturbation theory to find the density matrix for a group of $M \le N$ atoms, $\rho_A(M \le N)$, obtained by taking the trace of (14) over the field states represented by ρ_f and over $N - M$ particles. The result has the form

$$\begin{aligned}\rho_A(M \le N) &= |0\rangle\langle 0|[1 - M\kappa^2\langle B^\dagger B\rangle] \\ &+ \kappa[\langle B\rangle|1; M\rangle\langle 0| + \text{H.c.}] \\ &+ \kappa^2[\langle B^2\rangle|2; M\rangle\langle 0| + \text{H.c.}] \\ &+ \kappa^2\langle B^\dagger B\rangle|1; M\rangle\langle 1; M|.\end{aligned} \quad (16)$$

Note that the density matrix $\rho_A(M \le N)$ describes a mixed state of the atomic ensemble. Unlike the density matrices for symmetric Dicke states (6), $\rho_A(M \le N)$ is independent of both N and $p = M/N$. Therefore, the correlations between $M < N$ atoms are identical and are independent of the total particle number N. This implies that the state is robust to particle loss.

The atomic density matrix cannot be factorized because of the correlation depending on photon statistics. Consider two atoms described in terms of their respective observables c_1 and c_2 such that $[c_1; c_2] = 0$. Setting $M = 2$ in (16), we have the two-atom density matrix

$$\begin{aligned}\rho_A(2) &= |00\rangle\langle 00|(1 - 2\kappa^2\langle B^\dagger B\rangle) \\ &+ \kappa\langle B\rangle(|10\rangle\langle 00| + |01\rangle\langle 00| + \text{H.c.}) \\ &+ \kappa^2\langle B^2\rangle(|11\rangle\langle 00| + \text{H.c.}) \\ &+ \kappa^2\langle B^\dagger B\rangle(|10\rangle + |01\rangle)(\langle 10| + \langle 01|).\end{aligned} \quad (17)$$

Using (17), we find that the covariance of the operators c_1 and c_2 is determined by the electromagnetic field variance:

$$\begin{aligned}\langle c_1c_2\rangle - \langle c_1\rangle\langle c_2\rangle &= \kappa^2[(\langle B^2\rangle - \langle B\rangle^2)\langle 0|c_1|1\rangle\langle 0|c_2|1\rangle \\ &+ (\langle B^\dagger B\rangle - \langle B^\dagger\rangle\langle B\rangle)\langle 1|c_1|0\rangle\langle 0|c_2|1\rangle + \text{c.c.}].\end{aligned} \quad (18)$$

If the field is not fluctuating in the sense that its variances are zero, i.e., $\langle B^2\rangle - \langle B\rangle^2 = 0$ etc. (which is true in the present case, e.g., for a coherent state), then there is no correlation between atoms. Suppose that c_k ($k = 1, 2$) are dipole operators:

$$c_k = d_k = \mu(s_{01}(k) + s_{10}(k)),$$

where the matrix element μ is real. Then the correlation between two dipole moments depends on photon statistics. We define the quadrature operator

$$X_f = B^\dagger\exp(i\theta) + \text{H.c.}$$

Then, (18) implies that the covariance of the dipole moments is determined by the variance of the quadrature operator normally ordered with respect to the field operators B and $B^\dagger$ at $\theta = 0$:

$$\langle d_1d_2\rangle - \langle d_1\rangle\langle d_2\rangle = \mu^2\kappa^2D_N,$$

where

$$D_N = \langle X_f^2\rangle - \langle X_f\rangle^2 - \langle[B, B^\dagger]\rangle.$$

For coherent states, the variance is $D_N = 0$. The dipole moments are correlated both for a squeezed-state field (with $D_N < 0$) and for field in a classical state (with $D_N > 0$).

The necessary and sufficient condition for inseparability of a mixed state is provided by the Peres–Horodecki criterion [9], which is valid for systems with Hilbert spaces of dimension 2×2 and 2×3. In the case considered here, the state of a two-atom system described by $\rho_A(2)$ is inseparable (entangled) if at least one eigenvalue of the (partial) transpose of the density matrix with respect to the variables associated with one of the atoms (e.g., $\rho_A^{T_1}(2)$) is negative. As example, we consider light in Gaussian and squeezed states.

For a Gaussian field ($\langle B\rangle = \langle B^2\rangle = 0$), expression (17) reduces to the density matrix describing a superposition of the ground and mixed states:

$$\rho_A(2) = a|00\rangle\langle 00| + b[(|01\rangle + |10\rangle)(\langle 01| + \langle 10|)],$$

where $a + 2b = 1$ and $a = 1 - 2\kappa^2\langle B^\dagger B\rangle$. The eigenvalues of $\rho_A^{T_1}(2)$ are

$$\lambda = \left\{b, b, \frac{a}{2} \pm \sqrt{\frac{a^2}{4} + b^2}\right\}.$$

Since $\sqrt{a^2/4 + b^2} \approx a/2$ in the approximation considered here, we have the eigenvalues $\{b, b, a, 0\}$, i.e., a separable state.

Consider the case of resonant interaction with single-mode squeezed light ($B = ga$) generated, for example, by a parametric oscillator. A simple model of the oscillator is defined by the effective Hamiltonian

$$H = i\hbar\frac{f}{2}(a^{\dagger 2} - \text{H.c.}).$$

The solution is

$$a = a_0\cosh r + a_0^\dagger \sinh r,$$

where $r = f\tau$ is the squeezing parameter, τ is the normalized length of the nonlinear medium, and a_0 and $a_0^\dagger$ denote the input field operators. For the initial vacuum state,

$$\langle a\rangle = 0, \quad \langle a^2\rangle = \langle a^{\dagger 2}\rangle = \cosh r\sinh r,$$
$$\langle a^\dagger a\rangle = \sinh^2 r.$$

In this case, (17) reduces to the following density matrix of a two-atom system:

$$\rho_A(2) = |00\rangle\langle 00|[1 - 2\kappa^2\langle B^\dagger B\rangle]$$
$$+ \kappa^2[\langle B^2\rangle(|11\rangle\langle 00| + |00\rangle\langle 11|) + \text{H.c.}] \qquad (19)$$
$$+ \kappa^2\langle B^\dagger B\rangle(|10\rangle\langle 10| + |01\rangle\langle 10| + |10\rangle\langle 01| + |01\rangle\langle 01|).$$

The four eigenvalues of $\rho_A^{T_1}(2)$ are

$$\lambda = \left\{0, 1 - \frac{2}{n_s}\sinh^2 r, \pm\frac{1}{n_s}\exp(\pm r)\sinh r\right\}. \qquad (20)$$

To be specific, we set $r > 0$; i.e., we consider the state squeezed with respect to canonical momentum or phase. In this case, $(-1/n_s)\sinh r\exp(-2r) < 0$. However, it is clear that the degree of squeezing is low, because the approximations used here imply that

$$\frac{\sinh^2 r}{n_s} \ll 1. \qquad (21)$$

Thus, the state of the atomic system is inseparable. This behavior is explained as follows. Fluctuations of light give rise to correlation between atoms, which leads to two-atom coherence. When condition (21) holds, this coherence plays the key role. Since absorption is weak, the system is almost entirely in the ground state. As distinct from the case of Gaussian statistics, the density matrix has the form

$$\rho_A(2) \approx |00\rangle\langle 00| + \kappa^2[\langle B^2\rangle|11\rangle\langle 00| + \text{H.c.}].$$

Note that the following two observations can be inferred from this example. First, a steady entangled atomic state can be created by using weak squeezed light, which looks promising from an experimental perspective. Second, the entire ensemble cannot be interpreted as separable, because any pair in a group of $M \le N$ atoms is entangled, i.e., the quantum correlation of the ensemble as a whole is robust to particle loss.

Since no reliable universally applicable criterion is known for multiparticle entanglement, we apply the Peres–Horodecki criterion to two two-level subsystems and find that any pair of atoms in the ensemble can be inseparable, which gives reason to interpret the state of the entire system as inseparable.

Note also that spurious entanglement may be predicted by perturbation theory [28]. In that study, an example of expansion of the product of two wavefunctions in terms of a common classical parameter was considered in which individual summands represent entangled states. However, if entropy of entanglement is used as a measure, then we have initially independent systems, because the entropy is either quadratic in the small parameter or zero in arbitrary-order perturbation theory. Note that physical implementation of such entangled states, i.e., preparation of an independent state of a pair of entangled particles, requires projective measurement in an entangled basis. The present analysis also relies on perturbation theory, but we deal with a different situation in both physical and formal sense, in which interaction between particles gives rise to correlation. The wavefunction obtained in first-order perturbation theory is not factorizable, and the corresponding entropy of entanglement is zero to the corresponding accuracy. This result is physically plausible, because there is no correlation in first-order perturbation theory. In our analysis, entanglement is predicted by second-order perturbation theory, which describes real emission and absorption processes conducive to

correlation. In this order of perturbation theory, the existence of quantum correlation is substantiated by entanglement criteria consistent with approximation accuracy.

6. EXACT SOLUTIONS

Radiative decay can be neglected in (2) when evolution over a time $t \ll \gamma^{-1}$ is considered, and the behavior of the entire system is described by the wavefunction

$$\phi(t) = \exp(-i\hbar^{-1}Ht)(\phi_A \otimes \phi_f),$$

where the initial states of the atoms and field are assumed to be uncorrelated. Then, simple solutions can be obtained under certain initial conditions.

Consider the mixing of modes a and b described by

$$H = i\hbar f(a^\dagger bS - ab^\dagger S^\dagger), \tag{22}$$

where $S = S_{10}$ and $S^\dagger = S_{01}$. If analysis is restricted to single-photon Fock states of the modes $\phi_f = c|01\rangle_{ab} + e|10\rangle_{ab}$, exact solutions can be written as

$$\begin{aligned}
&\exp(-i\hbar^{-1}Ht)(c|01\rangle_{ab} + e|10\rangle_{ab}) \otimes \phi_A \\
&\quad = c\Bigg\{|01\rangle\cos(tf\sqrt{SS^\dagger}) \\
&\quad + |10\rangle S^\dagger \frac{1}{\sqrt{SS^\dagger}}\sin(tf\sqrt{SS^\dagger})\Bigg\} \otimes \phi_A \\
&\quad + e\Bigg\{-|01\rangle S\frac{1}{\sqrt{S^\dagger S}}\sin(tf\sqrt{S^\dagger S}) \\
&\quad + |10\rangle\cos(tf\sqrt{S^\dagger S})\Bigg\} \otimes \phi_A.
\end{aligned} \tag{23}$$

In the case of a single-photon process described by the Hamiltonian

$$H = i\hbar g(aS - a^\dagger S^\dagger), \tag{24}$$

there also exist simple solutions. For example,

$$\begin{aligned}
&\exp(-i\hbar^{-1}Ht)(c|1\rangle \otimes |0;\ N\rangle + e|0\rangle \otimes |1;\ N\rangle) \\
&\quad = c\Bigg\{\cos(gt\sqrt{N})|1\rangle \otimes |0;\ N\rangle \\
&\quad + \frac{1}{\sqrt{N}}\sin(gt\sqrt{N})|0\rangle \otimes |1;\ N\rangle\Bigg\} \\
&\quad + e\{-\sqrt{N}\sin(gt\sqrt{N})|1\rangle \otimes |0;\ N\rangle \\
&\quad + \cos(gt\sqrt{N})|0\rangle \otimes |1;\ N\rangle\},
\end{aligned} \tag{25}$$

where $|h;\ N\rangle$ ($h = 0, 1$), represents the ground state $|0\rangle^{\otimes N}$ of the atomic ensemble and a symmetric Dicke state defined in accordance with (6). These solutions are valid only under the restrictions imposed above on the initial states. They describe exchange of excitation between the cavity mode and the atoms.

7. GENERATION AND TRANSFORMATION OF SYMMETRIC STATES

Now, we use the exact solutions written out above to analyze the evolution of symmetric Dicke states $|h;\ N\rangle$ in single-photon and wave-mixing processes.

First, consider the case when the spatial inhomogeneity of the field within the region occupied by the atomic ensemble can be neglected. Setting $\phi_A = |h;\ N\rangle$ in (25), we use (9) to obtain

$$\begin{aligned}
&(\alpha|01\rangle + \beta|10\rangle) \otimes |h;\ N\rangle \\
&\quad \longrightarrow \alpha\Bigg\{\cos\theta_h|01\rangle \otimes |h;\ N\rangle \\
&\quad + \sqrt{\frac{h+1}{N-h}}\sin\theta_h|10\rangle \otimes |h+1;\ N\rangle\Bigg\} \\
&\quad + \beta\Bigg\{-\sqrt{\frac{N-h+1}{h}}\sin\theta'_h|01\rangle \otimes |h-1;\ N\rangle \\
&\quad + \cos\theta'_h|10\rangle \otimes |h;\ N\rangle\Bigg\},
\end{aligned} \tag{26}$$

where

$$\theta_h = tf\sqrt{(h+1)(N-h)},$$
$$\theta'_h = tf\sqrt{h(N-h+1)}.$$

Relation (26) entails possibilities of preparation of an entangled from ground-state atoms ($|0;\ N\rangle \longrightarrow |1;\ N\rangle$) and transformation of entangled states by changing the number of excited atoms ($|h;\ N\rangle \longrightarrow |h \pm 1;\ N\rangle$), including disentanglement ($|h;\ N\rangle \longrightarrow |h-1;\ N\rangle \longrightarrow \ldots \longrightarrow |0;\ N\rangle$).

Note that exact solutions (25) and (26) describe state swapping, which can be used to map the state of light onto atoms in order to store it in a long-lived atomic ensemble, i.e., to implement quantum memory. In particular, an unknown superposition of photons can be transferred to atoms and back by using the following transformation entailed by (25):

$$\begin{aligned}
&(\alpha|1\rangle + \beta|0\rangle) \otimes |0;\ N\rangle \\
&\quad \longleftrightarrows |0\rangle \otimes \left(\alpha\frac{1}{\sqrt{N}}|1;\ N\rangle + \beta|0;\ N\rangle\right).
\end{aligned} \tag{27}$$

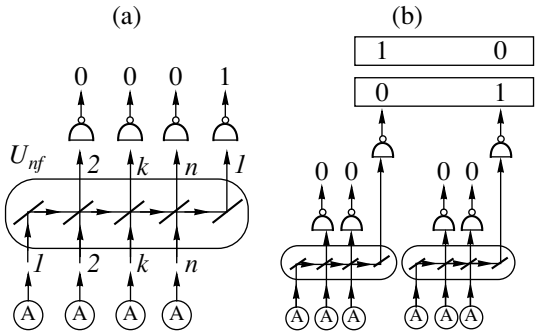

Fig. 1. (a) Scheme for generating entangled states of atomic ensembles. (b) Preparation of entangled states by correlation of photocounts recorded by two schemes.

Similarly, (26) can be used to map an entangled state of photons onto the atomic ensemble:

$$(\alpha|01\rangle + \beta|10\rangle) \otimes |0;\ N\rangle \rightleftarrows |10\rangle \otimes \left(\alpha\frac{1}{\sqrt{N}}|1;\ N\rangle + \beta|0;\ N\rangle\right). \tag{28}$$

Solutions (25) and (26) make it possible to take into account the spatial configuration of atoms in the ensemble. For example, consider the interaction between a one-dimensional array of atoms located at points $x_1, \ldots, x_N$ and a single photon described by Hamiltonian (24) with

$$S = \sum_p s_{10}(p)\exp(ikx_p),$$

where the operator $s_{10}(p) = |1\rangle_p\langle 0|$ corresponds to the atom located at x_p ($p = 1, \ldots, N$). Using (25), we can show that

$$|1\rangle \otimes |0;\ N\rangle \longrightarrow \cos\theta|1\rangle \otimes |0;\ N\rangle + \sin\theta|0\rangle \otimes \eta_N(1), \tag{29}$$

where $\theta = tg\sqrt{N}$, and

$$\eta_N = \frac{1}{\sqrt{N}}[\exp(ikx_1)|10\ldots0\rangle + \ldots + \exp(ikx_N)|0\ldots01\rangle]. \tag{30}$$

Expression (30) implies that an array of entangled atoms is created when $\theta = \pi/2$. Note that η_N is the Dicke state with $j = m = N/2 - 1$ only if

$$\sum_p \exp[ikx_p] = 0.$$

8. ENTANGLED ATOMIC ENSEMBLES

Solutions (23) and (25) imply that a photon and an atomic ensemble are entangled via interaction. If photons are entangled (e.g., by projective measurement) in a combination of such independent systems, then the atomic ensembles will become entangled. We consider optical measurement schemes based on this method, known as entanglement swapping. The key resources used in these schemes are sets of atomic ensembles correlated with respective photons, beamsplitters, and single-photon detectors. The analysis that follows is restricted to schemes in which only specific single-photon output is recorded.

As an initial state, we use the EPR pair

$$Z(W) = a|0\rangle_\mathrm{f} \otimes |0\rangle + b|1\rangle_\mathrm{f} \otimes |W\rangle, \tag{31}$$

where Fock states are denoted by the subscript "f," $|W\rangle = |1;\ N\rangle/\sqrt{N}$, $|0\rangle = |0;\ N\rangle$. It is generated by the mode mixing described by (22), where the mode b is a classical wave. The state of n independent identical ensembles entangled with respective photons is represented by the product

$$Z_n(W) = Z(W)^{\otimes n} = a^{n-1}b[|10\ldots0\rangle_\mathrm{f} \otimes |W0\ldots0\rangle + \ldots + |00\ldots1\rangle_\mathrm{f} \otimes |00\ldots W\rangle] + \ldots. \tag{32}$$

As illustrated by Fig. 1a, the photons associated with atomic ensembles are injected into a system of $n - 1$ beamsplitters with n input ports and n output ports. Each beamsplitter performs the transformation

$$|01\rangle_\mathrm{f} \longrightarrow c_k|01\rangle_\mathrm{f} + s_k|10\rangle_\mathrm{f},$$
$$|10\rangle_\mathrm{f} \longrightarrow -s_k|01\rangle_\mathrm{f} + c_k|10\rangle_\mathrm{f},$$

where $c_k^2 + s_k^2 = 1$ ($k = 1, \ldots, n - 1$). The scheme is described by a unitary operator U_{nf} and characterized by the following property. There exist an input port optically coupled to every output port and an output port optically coupled to every input port. In Fig. 1a, the latter is output port *1*. We call it the optical output port, and the corresponding detector is called the output detector. The scheme performs the transformation

$$U_{nf}|1\ldots0\rangle_f = t_1|1\ldots0\rangle_\mathrm{f} + \ldots + t_n|0\ldots1\rangle_\mathrm{f},$$
$$U_{nf}^{-1}|1\ldots0\rangle = \tau_1|1\ldots0\rangle + \ldots + \tau_n|0\ldots1\rangle_\mathrm{f}, \tag{33}$$

where the coefficients t_k and τ_k ($k = 1, \ldots, n$) are determined by the transmittances and reflectances of the beamsplitters, and

$$\sum_k t_k^2 = \sum_k \tau_k^2 = 1.$$

If the output detector detects a photon (which corresponds to the state $|1_\mathrm{f}\rangle = |1\ldots0\rangle$, then there is a probability

$$\mathrm{Prob}(1) = |a^{n-1}b|^2 \tag{34}$$

that an entangled state of atomic ensembles will be prepared:

$$\frac{\langle 1_f|U_{nf}Z_n(W)}{\sqrt{\text{Prob}(1)}} = \eta_n(W),$$
$$\eta_n(W) = q_1|W\ldots 0\rangle + q_n|0\ldots W\rangle. \quad (35)$$

This scheme has the following property. Since the coefficients $q_1, \ldots, q_n$ are completely determined by the transmittances and reflectances of the beamsplitters, weakly entangled states $Z(W)$ can be used to prepare highly entangled states atomic ensembles.

Let us consider several particular cases. If $n = 2$, then $q_1 = c_1$ and $q_2 = s_1$, and we have an EPR pair of the form

$$\eta_2(W) = \text{EPR}(W) = c_1|W0\rangle + s_1|0W\rangle.$$

When $c_1 = s_1 = 1/\sqrt{2}$, it is maximally entangled. If $n = 3$ and $q_1 = c_1c_2$, $q_2 = -s_1c_2$, and $q_3 = s_2$, then we have a W state. If $c_1 = -s_1 = 1/\sqrt{2}$ and $c_2 = \sqrt{2/3}$, then

$$\eta_3(W) = W(W) = \frac{1}{\sqrt{3}}(|W00\rangle + |0W0\rangle + |00W\rangle). \quad (36)$$

In particular, one can prepare the asymmetric state

$$\tilde{W}(W) = \frac{1}{\sqrt{2}}|W00\rangle + \frac{1}{2}|0W0\rangle + \frac{1}{2}|00W\rangle.$$

When $N = 1$, it is unitarily equivalent to the GHZ state and can be used as a quantum channel for teleportation or dense coding [16].

Using correlation between photocounts in a combination of schemes considered above, mixed states of atomic ensembles can be prepared, including inseparable ones. For example, consider two independent identical schemes $S_2(X)$ combined as shown in Fig. 1b, with three single-photon detectors in each scheme. If a photon is detected by either scheme, then we have the pair of states

$$\langle 1_f|S_2(X) \otimes \langle 0_f|S_2(X)w = |\eta_2(X), 0\rangle$$

and

$$\langle 0_f|S_2(X) \otimes \langle 1_f|S_2(X)w = |0, \eta_2(X)\rangle.$$

Suppose that the detector outputs are connected so that a single photon produced by either scheme is counted. This measurement is described by the projector $|1_f0_f\rangle\langle 1_f0_f| + |0_f1_f\rangle\langle 0_f1_f|$. The resulting mixed state is represented by a density matrix of the form

$$\rho(X) = \frac{1}{2}[|\eta_2(X), 0\rangle\langle\eta_2(X), 0| + |0, \eta_2(X)\rangle\langle 0, \eta_2(X)|]. \quad (37)$$

Its separability is an open question, because a necessary and sufficient condition is known only for mixed systems of dimension 2×2 and 2×3. However, if we assume that $N = 1$, i.e., consider a combination of four atoms instead of ensembles, then

$$\eta_2(X) = \Psi^+ = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$

and density matrix (37) describes a four-particle state:

$$\rho(4) = \frac{1}{2}(|\Psi^+00\rangle\langle\Psi^+00| + |00\Psi^+\rangle\langle 00\Psi^+|). \quad (38)$$

Taking the state of the pair of atoms in the first scheme defined by the two-particle reduced density matrix

$$\rho(2) = \frac{1}{2}[|\Psi^+\rangle\langle\Psi^+| + |00\rangle\langle 00|],$$

we can apply the separability criterion. Since the transpose of the density matrix with respect to the variables associated with one atom has four eigenvalues one of which is negative, $\{1/4, 1/4, (1 \pm \sqrt{2})/4\}$, the state of this pair is inseparable (entangled). Therefore, the density matrix $\rho(4)$ is inseparable.

9. HIERARCHICAL STRUCTURE OF STATES

Note that expression (35) is hierarchically structured. To illustrate this property, we consider a combination of schemes generating states of this type. As distinct to schemes using correlation of photocounts, we consider optically connected schemes. If an elementary scheme that performs the transformation $S_n(X) = U_{nf}Z_n(W)$ with $X = W$ (see Fig. 1a) records single-photon output, then the resulting state has the form of (35):

$$\langle 1_f|S_n(X)w = \eta_n(X) = \tau_1'|X0\ldots 0\rangle + \ldots + \tau_n'|00\ldots X\rangle, \quad (39)$$

where $w = 1/\sqrt{\text{Prob}(1)}$. We define the optical output port of the scheme $S_n(X)$ as the one optically coupled to every input port. In Fig. 1a, it is output port *1*. The input port of the scheme $S_n(X)$ is defined as the optical input port of the system of beamsplitters. Then, we can take, for example, p independent schemes represented as $(S_n(X))^p$ and use their optical outputs as the input of the scheme S_p. As a result, we have a new scheme $S_p((S_n(X))^p)$. If it records single-photon output, we have an entangled state that consists of lower level entangled states:

$$\langle 1_f|S_p((S_n(X))^p)w = \eta_p(\eta_n(X)) = t_1|\eta_n(X), 0\ldots 0\rangle + \ldots + t_p|0, 0, \ldots, \eta_n(X)\rangle. \quad (40)$$

By virtue of (39), it can be represented as

$$\eta_p(\eta_n(X)) = \eta_{pn}(X). \quad (41)$$

Thus, we can formulate the following rule. The state $\eta_n(X)$ defined by (39) with $n = n_1 n_2 \ldots n_p$ can be represented as

$$\eta_n(X) = \eta_{n_1}(\eta_{n_2}(\ldots(\eta_{n_p})). \quad (42)$$

This implies that the vector $\eta_n(X)$ has the structure of an entangled state with respect to any group of *s* particles, where *s* is such that *n*/*s* is a natural number greater than unity.

When the wavefunctions in state (35) or (40) are symmetric, a hierarchically structured representation can be obtained by using the permanent expansion defined as a determinant with a summation rule for permutations depending on symmetry [29]. In particular, successive decomposition of a determinant with respect to rows or columns and subsequent association of summands can be used to represent a permanent in terms of permanents of lower dimension, which reflects hierarchical structure.

For example, when $n = 6$, it holds that

$$\eta_6(X) = \eta_3(\eta_2(X)) = \eta_2(\eta_3(X))). \quad (43)$$

This state has the structure of an EPR pair or a *W* state:

$$\eta_3(\eta_2(X)) = W(\mathrm{EPR}) = \mathrm{EPR}(W).$$

This example demonstrates that the same state exhibits a structure characteristic of entangled states of two different types. This property can be used in different applications: the EPR pair can serve as a quantum channel for teleportation or dense coding, while the symmetric *W* state can be used for cloning.

To choose a particular structure defined by the dimension of the Hilbert space of its element, appropriate basis vectors and observables should be introduced. In physical terms, this is equivalent to a two-level approximation. Indeed, any group of *s* particles, where *s* is such that *n*/*s* is a natural number greater than unity, is represented in $\eta_n(X)$ by two states, $|0\rangle = 0_s$ and $\eta_s(X) = 1_s$. The group can be treated as a two-level particle (qubit) with basis vectors 0_s and 1_s. Such qubits and hierarchically structured states $\eta_n(X)$ can be used in quantum information processing. By analogy with (30), the vector $\eta_n(X)$ represents a Dicke state only if $\sum_k \tau_k = 0$.

10. CONCLUSIONS

A model describing resonant interaction of identical two-level atoms with a narrow-band radiation field is used to analyze multiparticle entanglement. The interaction is described by an effective Hamiltonian that allows for various multiphoton processes. The statistics of radiation and atoms are characterized by a density matrix, for which solutions are calculated in second-order perturbation theory in the interaction strength and exact solutions are found in the case of negligible decay. It is shown that the state of any pair of atoms interacting with weak single-mode squeezed light is inseparable and robust against decay. It is demonstrated that symmetric entangled multiparticle states can be generated by using optical schemes based on single-photon projection. An optical scheme is described that can be used to prepare highly entangled states of atomic ensembles from weakly entangled states by projective measurement.

ACKNOWLEDGMENTS

This work was supported, in part, by Delzell Foundation, Inc.

REFERENCES

1. B. Julsgaard, A. Kozhekin, and E. S. Polzik, Nature **413**, 400 (2001).
2. C. Monroe, D. M. Meekhof, B. E. King, *et al.*, Phys. Rev. Lett. **75**, 4714 (1995).
3. E. Hagley, X. Maître, G. Nogues, *et al.*, Phys. Rev. Lett. **79**, 1 (1997).
4. G. K. Brennen, C. M. Caves, P. S. Jessen, and I. H. Deutsch, Phys. Rev. Lett. **82**, 1060 (1999).
5. A. V. Burlakov, M. V. Chekhova, and O. A. Karabutova, Phys. Rev. A **60**, 4209 (1999); A. V. Burlakov and M. V. Chekhova, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 505 (2002) [JETP Lett. **75**, 432 (2002)]; A. V. Burlakov, L. A. Krivitskiĭ, S. P. Kulik, *et al.*, Opt. Spektrosk. **94**, 743 (2003) [Opt. Spectrosc. **94**, 684 (2003)]; Yu. I. Bogdanov, L. A. Krivitskiĭ, and S. P. Kulik, Pis'ma Zh. Éksp. Teor. Fiz. **78**, 804 (2003) [JETP Lett. **78**, 352 (2003)]; L. A. Krivitskiĭ, S. P. Kulik, A. N. Penin, and M. V. Chekhova, Zh. Éksp. Teor. Fiz. **124**, 943 (2003) [JETP **97**, 846 (2003)].
6. W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).
7. M. Eibl, N. Kiesel, M. Bourennane, *et al.*, Phys. Rev. Lett. **92**, 077901 (2004).
8. F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde, Phys. Rev. A **65**, 052112 (2002).
9. A. Peres, Phys. Rev. Lett. **77**, 1413 (1996); M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
10. A. M. Basharov, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 151 (2002) [JETP Lett. **75**, 123 (2002)]; Zh. Éksp. Teor. Fiz. **121**, 1249 (2002) [JETP **94**, 1070 (2002)].
11. M. Plesch, J. Novotny, Z. Dzurakova, and V. Buzek, J. Phys. A: Math. Gen. **37**, 1843 (2004).
12. W. Dür, Phys. Rev. A **63**, 020303 (2001).
13. J. K. Stockton, J. M. Geremia, A. C. Doherty, and H. Mabuchi, Phys. Rev. A **67**, 022112 (2003).

14. V. Buzek and M. Hillery, Phys. Rev. A **54**, 1844 (1996); M. Murao, D. Jonathan, M. B. Plenio, and V. Vedral, Phys. Rev. A **59**, 156 (1999).
15. J. Joo, Y.-J. Park, J. Lee, and J. Jang, quant-ph/0204003.
16. V. N. Gorbachev, A. I. Trubilko, A. A. Rodichkina, and A. I. Zhiliba, Phys. Lett. A **314**, 267 (2003).
17. E. Knill, R. Laflamme, and G. J. Milburn, Nature **409**, 46 (2001).
18. R. Raussendorf and H. J. Briegel, Phys. Rev. Lett. **86**, 5188 (2001).
19. R. Dicke, Phys. Rev. **93**, 99 (1954).
20. A. V. Andreev, V. I. Emel'yanov, and Yu. A. Il'inskiĭ, *Cooperative Phenomena in Optics* (Nauka, Moscow, 1988) [in Russian].
21. H. Hammer, quant-ph/0407094.
22. A. Olaya-Castro, N. F. Johnson, and L. Quiroga, Phys. Rev. A **70**, 020301 (2004).
23. N. Gisin and S. Popescu, Phys. Rev. Lett. **83**, 432 (1999).
24. A. M. Basharov, *Method of Unitary Transformation in Nonlinear Optics* (Mosk. Inzh.–Fiz. Inst., Moscow, 1990) [in Russian]; Zh. Éksp. Teor. Fiz. **102**, 1126 (1992) [Sov. Phys. JETP **75**, 611 (1992)].
25. D. V. Kupriyanov, I. M. Sokolov, and A. V. Slavgorodskii, Phys. Rev. A **63**, 063811 (2001); **68**, 043815 (2003).
26. V. N. Gorbachev, A. I. Zhiliba, A. A. Rodichkina, and A. I. Trubilko, Phys. Lett. A **323**, 339 (2004).
27. H. J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001); R. Raussendorf, D. E. Browne, and H. J. Briegel, quant-ph/0301052.
28. A. M. Basharov and É. A. Manykin, Opt. Spektrosk. **96**, 91 (2004) [Opt. Spectrosc. **96**, 81 (2004)].
29. S. S. Schweber, *An Introduction to Relativistic Quantum Field Theory* (Row, Peterson, Evanston, Ill, 1961; Inostrannaya Literatura, Moscow, 1963).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 44–55.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 54–66.

ATOMS, MOLECULES, OPTICS

Metastable Helium Cluster He_4^*

V. F. Elesin, N. N. Degtyarenko, N. V. Matveev, A. I. Podlivaev, and L. A. Openov*

Moscow Engineering Physics Institute (State University), Moscow, 115409 Russia
**e-mail: opn@supercon.mephi.ru*
Received January 27, 2005

Abstract—The existence of a metastable cluster He_4^* with total spin $S = 2$ is predicted. The cluster consists of two covalently bound excited spin-polarized triplet He_2^* molecules and is rectangular in shape. The electron wavefunctions, the dependence of the energy of the He_4^* system on the distance between the He_2^* triplet molecules, the atomic spacing, the frequency spectrum of natural oscillations of the cluster, and other characteristics are calculated from first principles. It is shown that the metastable state is formed if one of the excited He_2^* molecules is in the $^3\Sigma_u^+$ state, while the other is in the $^3\Pi_g$ state. The radiation lifetime τ of the metastable cluster He_4^* is calculated; it is found to range from 100 to 200 s, which is much longer than the lifetime $\tau \approx 20$ s of the triplet molecule He_2^* ($^3\Sigma_u^+$). The height $U \approx 0.5$ eV of the potential barrier preventing the departure from the local energy minimum is determined. The energy $E_{acc} \approx 9$ eV/atom accumulated in the He_4^* cluster is calculated; this energy considerably exceeds the energy of known chemical energy carriers. It is shown that the accumulated energy is released virtually completely during decomposition of the He_4^* cluster into individual helium atoms. This means that helium clusters are a promising material with a high accumulated energy density (HEDM). © *2005 Pleiades Publishing, Inc.*

INTRODUCTION

One of the main goals in nanophysics and nanotechnologies is the synthesis of fundamentally new substances with required properties from the known elements of the Periodic Table. The "building block" of such a substance is a cluster consisting of a small number of atoms (and having a size on the order of a nanometer) and having the structure and properties radically differing from those of the conventional condensed substance.

If clusters are combined into an ensemble so that they preserve their individual properties while remaining bound to one another, they form a new material (nanosubstance). A brilliant example is the carbon cluster fullerene C_{60} discovered in 1985 [1]. An ensemble of fullerenes (fullerite) possesses a large number of unique properties. For example, fullerites exhibit ferromagnetism [2] and superconductivity [3], while familiar carbon structures such as graphite and diamond are semiconductors.

The question arises: is it possible to synthesize a nanosubstance capable of storing and releasing large amounts of energy? Large demand existing at present in the new type of energy carriers is due to a considerable gap between the characteristic energies of chemical and nuclear energy carriers. Another question: which chemical elements should be used to obtain HEDM? In this study, we give the following answer to these questions: it is possible in principle to synthesize clusters of helium, viz., an element that does not form a condensed substance under normal conditions (i.e., at room temperature and under atmospheric pressure).

We predict the existence of a metastable He_4^* cluster consisting of four helium atoms and prove that the energy accumulated in the cluster exceeds the energy of available chemical energy carriers by more than an order of magnitude (preliminary results were published in [4]). The structure and energy parameters of the He_4^* cluster and its stability and lifetime are studied in detail from first principles. It is shown that the energy accumulated in the cluster is released virtually completely during its decomposition into individual helium atoms.

It should be noted that helium is a representative of a group of elements that do not form a condensed substance under normal conditions. This group also includes nitrogen, oxygen, and some other elements. It was shown in [4] that such elements may form metastable clusters whose binding energy decreases with increasing number of atoms in a cluster (type II clusters

according to terminology proposed in [4]). The energy stored in type II clusters is released during decomposition of clusters into molecules or individual atoms. Note that most chemical elements form type I clusters (see [4]) whose binding energy increases with the number of atoms in a cluster, while the accumulated energy is released upon merging of small clusters into large ones.

It was noted earlier in [4] that type II clusters are formally analogous to heavy metastable nuclei (e.g., of uranium), while type I clusters are analogous to light nuclei (e.g., deuterium). Indeed, the binding energy of heavy nuclei decreases upon an increase in the number N of nucleons in the nuclei, while the energy is released during nuclear fission. On the contrary, the binding energy in light nuclei increases with N, making nuclear fusion advantageous from the energy point of view. We studied the properties of type I and II clusters theoretically in [4] using carbon and nitrogen clusters, respectively, as examples.

As motivation for such a choice of chemical elements, let us consider the total energy $E(\{\mathbf{R}_i\})$ of a cluster consisting of N atoms as a function of coordinates of all atoms, $\{\mathbf{R}_i\}$, $i = 1, \ldots, N$. The minima of $E(\{\mathbf{R}_i\})$ correspond to different structures that can be formed by these atoms. The global minimum with the lowest energy E_0 corresponds to the so-called ground state of the system. This minimum is restricted by an infinitely long barrier ensuring an infinitely long lifetime in the ground state (in the absence of external effects). Most of the substances existing in nature are in the ground state.

The remaining (local) minima with energy $E_n > E_0$ correspond to metastable structures. The lifetime of a metastable structure is finite since there is a finite probability of transition of the system to the ground state with release of energy $E_{acc} = E_n - E_0$ stored in the metastable state. It should be noted that the lifetime τ of a metastable structure may be very long (e.g., many years in the case of diamond) since the quantity τ is an exponential function of temperature, $\tau(T) = \tau_0 \exp(U/k_B T)$, where U is the height of the energy barrier separating the metastable state from the ground state, τ_0 is a microscopic value on the order of 1 fs–1 ps, and k_B is the Boltzmann constant.

Since nitrogen, oxygen, and hydrogen, as well as noble gases (like helium), do not form a condensed substance under normal conditions (we disregard condensation due to very week Van der Waals forces), we can assume that clusters of these elements do not possess a global minimum of the total energy $E(\{\mathbf{R}_i\})$. At the same time, local minima (if they exist) must obviously be formed for relatively high energies (otherwise they would have been observed experimentally). It follows hence that, first, metastable structures of helium, nitrogen, and other elements of this group must accumulate large amounts of energy and, second, release of energy during a transition from a metastable to the ground state must take place during decomposition of clusters into atoms (e.g., He) or molecules (e.g., N_2). This leads to virtually complete release of accumulated energy, which is a considerable advantage of such structures as prospective energy carriers.

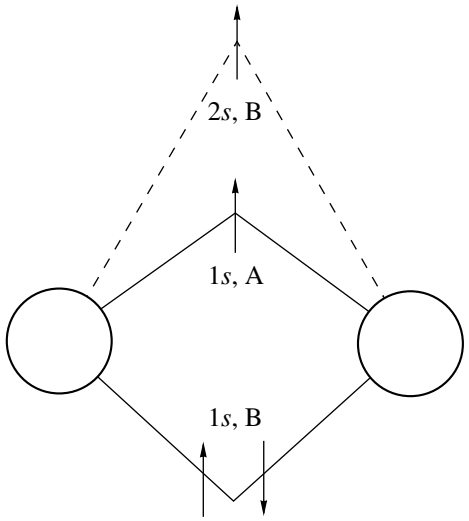

Fig. 1. Diagram of electron filling of orbitals of the He_2^* molecule in the triplet state $^3\Sigma_u^+$ with total spin $S = 1$.

The available experimental and theoretical data indicate that helium atoms in the ground state do not form clusters He_n ($n = 2, 3, 4, \ldots$) with covalent bonds. However, it was demonstrated experimentally in [5] that a metastable covalently bound cluster (triplet molecule) He_2^* in the excited state $^3\Sigma_u^+$, which was studied theoretically in [6, 7], exists. Figure 1 shows schematically the filling of orbitals of the He_2^* molecule with electrons. Three electrons occupy "inner" orbitals formed by atomic 1*s* orbitals, while the fourth electron is on the "outer" excited orbital formed by atomic 2*s* orbitals. It is important to note that the He_2^* molecule is in the triplet state with total spin $S = 1$ (see Fig. 1), which increases its stability and lifetime due to the prohibition on recombination imposed by the Pauli exclusion principle.

The triplet He_2^* molecule accumulates energy approximately equal to 9 eV/atom and releases this energy during decomposition into helium atoms. This was confirmed experimentally in [5]. Thus, the He_2^* molecule may play the role of a building block whose properties radically differ from the properties of conventional helium. The question arises: can an ensemble

of He_2^* molecules be formed? In such an ensemble, the building blocks He_2^* can be treated as *quasi-atoms* [4].

In the present paper, a search for such ensembles is carried out on the basis of ab initio calculations. As the first step, we theoretically analyze a He_4^* cluster consisting of two He_2^* quasi-atoms. It is shown that the formation of such a cluster is possible in principle due to overlapping of the wavefunctions of outer electrons. The total spin of this cluster is $S = 2$; i.e., it is in the "ferromagnetic" state. It is found that the energy accumulated in the He_4^* cluster is approximately twice as high as the energy stored in the excited He_2^* molecule; i.e., the specific energy $E_{acc} \approx 9$ eV/atom accumulated in the He_4^* cluster virtually does not change upon fusion of two triplet molecules (quasi-atoms) He_2^* into the He_4^* cluster. The radiation lifetime of the He_4^* cluster is estimated ($\tau \sim 100$ s) and is found to be an order of magnitude longer than the lifetime $\tau \sim 10$ s of the triplet He_2^* molecule.

The paper has the following structure. The computational methods used by us are described in Section 2. In Section 3, the results of calculation of the electron structure of the triplet He_2^* molecules and the energy accumulated in them are considered. The radiation lifetime of the He_2^* molecule in the $^3\Sigma_u^+$ state is calculated in Section 4. The structural and energy characteristics of the metastable He_4^* cluster are determined in Section 5, where the accumulated energy of the cluster is calculated. The radiation lifetime of the He_4^* cluster is obtained in Section 6. The results are briefly discussed in the Conclusions.

2. COMPUTATIONAL METHODS

In theoretical analysis of metastable states of small atomic clusters, even most advanced ab initio methods of calculation often lead to contradictory results (see, for example, [8] and the literature cited therein). This is observed when the energy difference for two or several states of a cluster is comparable to the error of the method, which amounts to not less than 0.1–1 eV as a rule. For describing known structures, one chooses a method which leads to correct results for physical quantities whose values are already known from experiments (e.g., the bond lengths and the binding energy). A completely different situation takes place when characteristics of new physical objects which have not been discovered or experimentally studied, must be sought and calculated. The results can be treated as authentic only if the data obtained by different methods coincided to a reasonable degree of accuracy.

In view of the above arguments, we calculated the energies and wavefunctions of helium clusters using several different ab initio methods, including the conventional unrestricted Hartree–Fock method (UHF) [9], the Hartree–Fock method with the second-order Møller–Plessett correction (MP2) [9], the density functional method [10, 11] in the local density approximation (LDA) [12], and the density functional method with the exchange-correlation functional proposed in [13, 14] (B3LYP). Some calculations were made using the GAMESS program (General Atomic and Molecular Electronic Structure Systems) [15, 16]. We performed calculations with various sets of basis functions to attain the convergence upon an increase in the number of functions in the basis. For the triplet molecule He_2^*, we strove to attain correspondence to the data obtained by the configuration interaction method in the framework of the self-consistent field theory (SCF-CI) [6]. In most cases, admissible accuracy was attained by using a basis (7*s*5*p*2*d*) formed by the *s*, *p*, and *d* orbitals of each atom.

3. METASTABLE He_2^* MOLECULES

It is well known that two helium (He) atoms, each of which is in the ground state, do not form a covalently bound He_2 molecule. If, however, one of the helium atoms is in the excited state He*, He_2^* metastable molecules can be formed. Henceforth, we will be interested in triplet molecules $He_2^*(^3\Sigma_u^+)$ and $He_2^*(^3\Pi_g)$ with total spin $S = 1$. It is these molecules that claim to be building blocks for spin-polarized ensembles $(He_2^*)_N$ (see Introduction). In the $He_2^*(^3\Sigma_u^+)$ molecule, electrons with spin up occupy three molecular α orbitals ($1s\sigma_g$, $2p\sigma_u$, $2s\sigma_g$), while one electron with spin down occupies the β orbital ($1s\sigma_g$), which corresponds to configuration $|\uparrow\downarrow_1, \uparrow_2, \uparrow_3\rangle$ (see Fig. 1). Orbitals $1s\sigma_g$ and $2p\sigma_u$ are bonding and antibonding orbitals, respectively, which are formed upon overlap of atomic 1*s* wavefunctions, while the $2s\sigma_g$ orbital is a bonding orbital that is formed from atomic 2*s* functions. The $He_2^*(^3\Sigma_u^+)$ molecule is long-lived on atomic time scale (its lifetime is $\tau \sim 10$ s [5–7]) and has been studied in detail both theoretically and experimentally.

In the triplet molecule $He_2^*(^3\Pi_g)$, the excited electron occupies the orbital formed from atomic 2*p* functions. This molecule is also metastable [6], but its radiation lifetime is very small with respect to the transition to the $He_2^*(^3\Sigma_u^+)$ state.

To verify and finalize our computational techniques, we performed detailed calculations of wavefunctions, energy, and other characteristics of He_2^* molecules by various methods (see Section 2). Self-consistent single-

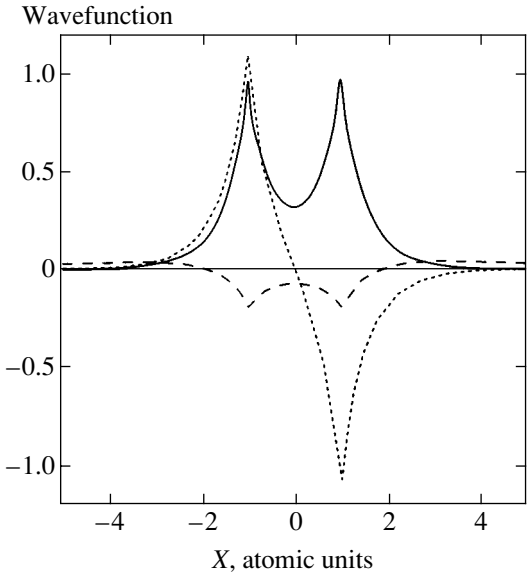

Fig. 2. Self-consistent single-electron wavefunctions of the metastable He_2^* molecule in the triplet state $^3\Sigma_u^+$ with total spin $S = 1$ along the line passing through the atoms ($y = z = 0$). Helium atoms are located at points with coordinates ($\pm a_B$, 0, 0). The solid curve corresponds to the first α orbital, the dotted curve is the second α orbital, and the dashed curve describes the third α orbital (the wavefunctions of the β orbital and the first α orbital are almost indistinguishable).

electron wavefunctions of the He_2^* ($^3\Sigma_u^+$) molecule are shown in Fig. 2. The wavefunctions calculated by different methods are almost identical and have the same shape as the corresponding single-electron wavefunctions of the molecular ion He_2^{3+} [9].

Figure 3a shows the dependences of the total energy E of the He_2^* ($^3\Sigma_u^+$) system on distance R between the atoms, which were calculated by different methods (including those described in the literature [6]). It can be seen from the figure that the $E(R)$ curve exhibits a local minimum for the atomic spacing $R_0 \approx 2a_B$ for all the methods used. The curves depicted in the figure have almost identical shapes and differ only in a shift along the energy axis. The triplet molecule He_2^* ($^3\Sigma_u^+$) has a single vibrational mode. Its frequency, which is determined by the curvature of the $E(R)$ curve at the point of minimum, is real-valued. The presence of the local minimum on the $E(R)$ curve indicates that the He_2^* ($^3\Sigma_u^+$) molecule is metastable. The height U of the potential barrier preventing the escape from the local minimum depends on the computational method and amounts to 1.1–1.7 eV.

To calculate the energy accumulated in the He_2^* ($^3\Sigma_u^+$) molecule, we calculated the energy of the

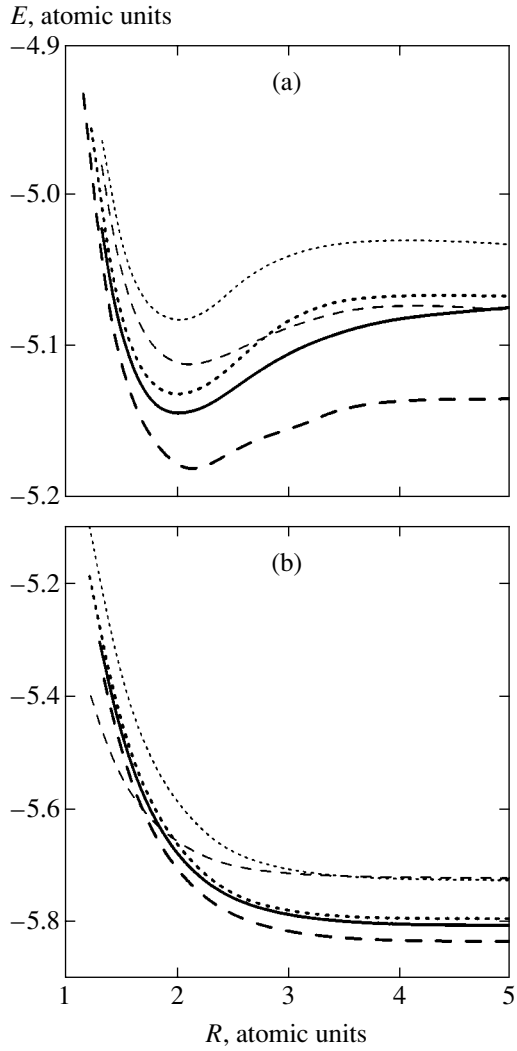

Fig. 3. Dependence of total energy E on distance R between atoms for the He_2^* system consisting of two helium atoms in the triplet state $^3\Sigma_u^+$ with total spin $S = 1$ (a) and for a system of two helium atoms in the singlet state $^1\Sigma_g^+$ with total spin $S = 0$ (b). Quantities E and R are expressed in atomic units ($me^4/\hbar^2$ and $\hbar^2/me^2$, respectively). Calculations are performed by the UHF (fine dotted curve), MP2 (bold dotted curve), LDA (fine dashed curve), and B3LYP (bold dashed curve) methods. The results of calculations based on the SCF-CI method [6] (solid curve) are shown for comparison.

He_2 system formed by two helium atoms in the singlet ground state $^1\Sigma_g^+$ with total spin $S = 0$ (see Fig. 3b). Each of the two lower α (spin up) orbitals and β (spin down) orbitals are occupied by two electrons, which corresponds to configuration $|\uparrow\downarrow_1, \uparrow\downarrow_2\rangle$. It can be seen from Fig. 3b that the energy as a function of atomic spacing has no local minimum; i.e., atoms in the ground state do not form a He_2 molecule. We can now estimate

the energy difference in the triplet excited He_2^* ($^3\Sigma_u^+$) state and the singlet ground He_2($^1\Sigma_g^+$) state for $R \approx 2a_B$. This value depends insignificantly on the computational method and amounts to 14–15 eV. The end product of decomposition of the He_2^* ($^3\Sigma_u^+$) molecule are two He atoms in the singlet $1s^2$ states, separated from each other. The difference $E_{acc} = E(He_2^*) - 2E(He) \approx 18$ eV ≈ 9 eV/atom is the energy accumulated in the He_2^* ($^3\Sigma_u^+$) molecule. The accumulated energy can be released, for example, due to emission of a photon by the electron occupying the third α orbital upon its transition to the second β orbital (see Fig. 1) with spin flip.

Analogous calculations were performed for the triplet molecule He_2^* ($^3\Pi_g$). Like the above data, the results of these calculations are also in good agreement with the familiar experimental and theoretical results. Thus, we confirmed the results available in the literature and verified the reliability of our computational methods.

4. THE LIFETIME OF THE TRIPLET MOLECULE He_2^* ($^3\Sigma_u^+$)

The lifetime of the metastable state is the question of fundamental importance. The experimentally measured lifetime of the triplet molecule He_2^* ($^3\Sigma_u^+$) is $\tau = 13 \pm 2$ s [5], which is in accordance with the results of numerical theoretical calculations ($\tau \approx 18$ s [6] and 10 s [7]). The long (by atomic standards) lifetime of the spin-polarized state is associated with the prohibition imposed by the Pauli exclusion principle on electron transitions.

To verify and finalize the computational technique, we calculated the lifetime of the metastable triplet molecule He_2^* ($^3\Sigma_u^+$). The total-spin prohibition imposed on the transition $^3\Sigma_u^+ \longrightarrow {}^1\Sigma_g^+$ is removed if we take into account the spin-dependent relativistic corrections in the Hamiltonian of the interaction of He_2^* with the electromagnetic field. As a result, the probability of the radiative transition $^3\Sigma_u^+ \longrightarrow {}^1\Sigma_g^+$ becomes nonzero, although contains a small factor on the order of α^4 as compared to the probability of the conventional (spin-allowed) dipole transition. Here, $\alpha = e^2/\hbar c$ is the fine structure constant, e is the electron charge, $\hbar$ is the Planck constant, and c is the velocity of light.

Retaining the terms of the same powers in α in the Hamiltonian $\hat{V}_{int}$ describing the interaction of electrons with the photon field [17], we obtain

$$\hat{V}_{int} = -\frac{e\hbar}{2mc}\sum_j \boldsymbol{\sigma}_j \cdot \mathbf{H}_j - \frac{e\hbar}{4m^2c^2}\sum_j \boldsymbol{\sigma}_j \cdot [\mathbf{E}_j \times \hat{\mathbf{p}}_j]$$
$$- \frac{Ze^3\hbar}{4m^2c^3}\sum_j\sum_n \frac{\boldsymbol{\sigma}_j \cdot [\mathbf{r}_{jn} \times \mathbf{A}_j]}{r_{jn}^3}$$
$$- \frac{e^3\hbar}{2m^2c^3}\sum_j\sum_{j'\neq j} \frac{\boldsymbol{\sigma}_j \cdot [\mathbf{r}_{jj'} \times \mathbf{A}_{j'}]}{r_{jj'}^3} \quad (1)$$
$$+ \frac{e^3\hbar}{4m^2c^3}\sum_j\sum_{j'\neq j} \frac{\boldsymbol{\sigma}_j \cdot [\mathbf{r}_{jj'} \times \mathbf{A}_j]}{r_{jj'}^3},$$

where m is the electron mass; $\boldsymbol{\sigma}_j$, $\mathbf{r}_j$, and $\hat{\mathbf{p}}_j$ are the Pauli matrix, coordinate, and momentum operator of the jth electron, respectively; $\mathbf{r}_{jk} = \mathbf{r}_j - \mathbf{r}_k$, $\mathbf{r}_{jn} = \mathbf{r}_j - \mathbf{R}_n$, $\mathbf{R}_n$ being the coordinate of the nth helium atom; $Z = 2$; and $\mathbf{H} = \mathrm{curl}\,\mathbf{A}$ and $\mathbf{E} = -(1/c)(\partial\mathbf{A}/\partial t)$ are the magnetic and electric fields of a photon, respectively. Vector potential $\mathbf{A}$ at the locus of the jth electron has the form of a linear combination of operators of creation, $\hat{a}_{k\lambda}^+$, and annihilation, $\hat{a}_{k\lambda}$, of a photon in a state with wave vector $\mathbf{k}$ and polarization λ:

$$\mathbf{A}_j = \sum_{\mathbf{k}\lambda}\left(\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}\lambda}}\right)^{1/2}$$
$$\times \mathbf{e}_{\mathbf{k}\lambda}[\hat{a}_{\mathbf{k}\lambda}\exp(i\mathbf{k}\cdot\mathbf{r}_j - i\omega_{\mathbf{k}\lambda}t) \quad (2)$$
$$+ \hat{a}_{\mathbf{k}\lambda}^+\exp(-i\mathbf{k}\cdot\mathbf{r}_j + i\omega_{\mathbf{k}\lambda}t)],$$

where V is the normalization volume; $\omega_{\mathbf{k}\lambda} = kc$ is the photon frequency; $\mathbf{e}_{\mathbf{k}\lambda}$ are the unit vectors of photon polarization ($\lambda = 1, 2$); and $\mathbf{k}\cdot\mathbf{e}_{\mathbf{k}\lambda} = 0$ in the calibration $\mathrm{div}\,\mathbf{A} = 0$. Summation over j and j' in Eq. (1) is carried out over all electrons.

In the initial state $|i; 0_{\mathbf{k}\lambda}\rangle$, photons are absent and the electron subsystem is described by the wavefunction $\Psi_i(\{\mathbf{r}_j\}, \{\sigma_j\})$ corresponding to the configuration $|\uparrow\downarrow_1, \uparrow_2, \uparrow_3\rangle$ ($^3\Sigma_u^+$) and depending on the coordinates and spin projections of all electrons. In the final state $|f; 1_{\mathbf{k}\lambda}\rangle$, one photon ($\mathbf{k}$, λ) is present, and the electron subsystem is described by the wavefunction $\Psi_f(\{\mathbf{r}_j\}, \{\sigma_j\})$ corresponding to the configuration $|\uparrow\downarrow_1, \uparrow\downarrow_2\rangle$ ($^1\Sigma_g^+$).

The probability of the $|i\rangle \longrightarrow |f\rangle$ transition per unit time is determined by the Fermi golden rule,

$$W_{i\to f} = \frac{2\pi}{\hbar}\sum_{\mathbf{k}\lambda}\left|\langle f; 1_{\mathbf{k}\lambda}|\hat{V}_{\rm int}|i; 0_{\mathbf{k}\lambda}\rangle\right|^2 \times \delta(E_i - E_f - \hbar\omega_{\mathbf{k}\lambda}), \quad (3)$$

where E_i and E_f are the total energies of the initial and final states of the electron subsystem, respectively, and the δ function reflects the energy conservation law in the case of emission of a photon, $E_i = E_f + \hbar\omega_{\mathbf{k}\lambda}$. The lifetime of the metastable state is given by

$$\tau = \frac{1}{W_{i\to f}}. \quad (4)$$

The calculation of $W_{i\to f}$ is reduced to the calculation of matrix elements

$$\langle f|\hat{V}_m|i\rangle = \int\Psi_f^*(\{\mathbf{r}_j\},\{\sigma_j\})\hat{V}_m(\{\mathbf{r}_j\},\{\hat{\mathbf{p}}_j\}) \times\Psi_i(\{\mathbf{r}_j\},\{\sigma_j\})\prod_j d\mathbf{r}_j, \quad (5)$$

where operators $\hat{V}_m(\{\mathbf{r}_j\},\{\hat{\mathbf{p}}_j\})$ are determined by the form of the terms in the interaction Hamiltonian (1). Since the energy difference between the initial and final states is $E_i - E_f \sim 10$ eV, the wavelength

$$\frac{2\pi}{k} = \frac{2\pi c}{\omega} = \frac{2\pi\hbar c}{E_i - E_f} \sim 100 \text{ nm}$$

of the emitted photon is much larger than the size of the He_2^* molecule; consequently, we can expand the exponent $\exp(-i\mathbf{k}\cdot\mathbf{r})$ in formula (2) into a Taylor series in powers of $\mathbf{k}\cdot\mathbf{r}$. Each subsequent term in this expansion introduces an additional small factor α. Taking into account the orthogonality of the electron wavefunctions of the initial and final states, retaining in expression (1) only the terms of the same order in α after expanding $\exp(-i\mathbf{k}\cdot\mathbf{r})$ in powers of $\mathbf{k}\cdot\mathbf{r}$, and taking into consideration the identity

$$\langle f|\hat{\mathbf{p}}_j|i\rangle = -\frac{im}{\hbar}(E_i - E_f)\langle f|\hat{\mathbf{r}}_j|i\rangle,$$

we obtain the following expression from (3) after evaluating the sum over photon polarizations, integrating over angles, and averaging over the initial spin states and summing over the final spin states:

$$W_{i\to f} = \frac{1}{9}\frac{c}{a_B}\alpha^8\Delta E_{if}^5(\mathbf{r}_{if}^2 + \mathbf{a}_{if}^2). \quad (6)$$

Here, a_B is the Bohr radius, $\Delta E_{if} = E_i - E_f$,

$$\mathbf{r}_{if} = \langle f|\mathbf{r}_j|i\rangle,$$
$$\mathbf{a}_{if} = \langle f|\frac{Z}{\Delta E_{if}^2}\sum_n\frac{\mathbf{r}_{jn}}{r_{jn}^3} + \frac{1}{\Delta E_{if}^2}\sum_{j'\neq j}\frac{\mathbf{r}_{jj'}}{r_{jj'}^3}|i\rangle, \quad (7)$$

and all quantities with dimensions of energy and length are expressed in atomic units ($me^4/\hbar^2$ and $\hbar^2/me^2$, respectively).

Numerical calculation of matrix elements (7) using the wavefunctions obtained by the UHF, MP2, and LDA methods leads to values of radiative transition probability $W_{i\to f} = 3.0\times10^{-2}$ s^{-1}, 4.1×10^{-2} s^{-1}, and 4.3×10^{-2} s^{-1}, respectively; this corresponds to lifetimes $\tau = 33$, 24, and 23 s, respectively. These values of τ are in qualitative agreement with experiment ($\tau = 13\pm2$ s [5]), as well as with the available results of numerical calculations ($\tau\approx18$ s [6] and $\tau\approx10$ s [7]). A certain spread in theoretical results is apparently due to the difference in the method of numerical solution of the multielectron Schrödinger equation. Thus, we have confirmed the known results and proved the reliability of our approach to calculating the radiation lifetimes of metastable helium systems.

5. METASTABLE CLUSTER He_4^*

It was shown in Section 3 that the triplet molecule He_2^* in excited state possesses properties radically differing from the properties of helium in the conventional ground state. Indeed, the He_2^* molecule is a covalently bound object with a large amount of accumulated energy, which is subsequently released during decomposition into helium atoms. This molecule has quite a long radiation lifetime and is thermally stable. The question arises: can triplet molecules He_2^* form ensembles such that the above-mentioned properties will be preserved? In other words, can the He_2^* molecule play the role of a building block in the formation of metastable helium structures (see Introduction)?

As the first step, it is natural to consider an ensemble of two He_2^* molecules. The following physical considerations (which were given for the first time in [4]) raise hopes for such a formation. First of all, let us pay attention to the spatial distribution of electron density in the He_2^* molecule (see Fig. 2). Three electrons on the lower orbitals are localized near the nuclei, while the fourth (excited) electron has a much larger localization radius. For this reason, the He_2^* molecule is analogous to an atom with partly filled lower orbitals and a single "outer" electron. For brevity (and better visualization), following [4], we will refer to He_2^* as a quasi-atom.

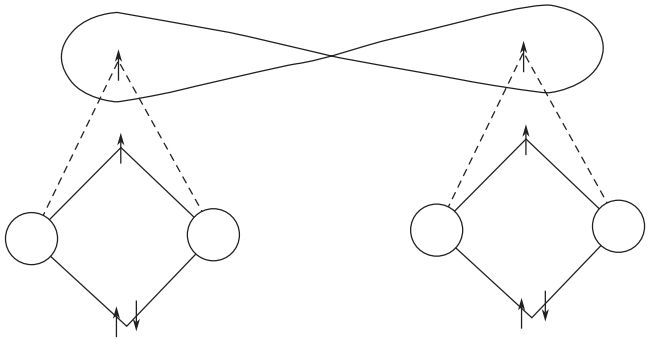

Fig. 4. Diagram of electron filling of orbitals of the He_4^* cluster with total spin $S = 2$.

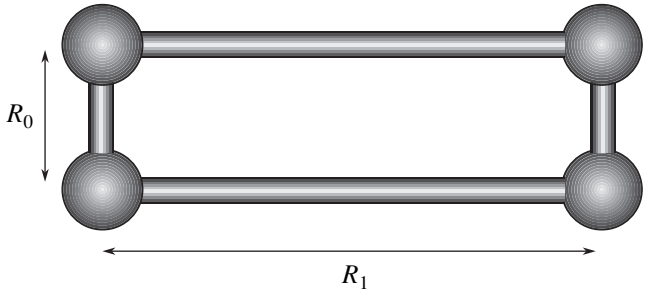

Fig. 5. Metastable He_4^* cluster with total spin $S = 2$, formed from two triplet He_2^* molecules (quasi-atoms); R_0 is atomic spacing in He_2^* quasi-atoms; R_1 is the distance between He_2^* quasi-atoms.

The problem is thus reduced to the formation of an ensemble (cluster) He_4^* of two quasi-atoms He_2^* due to the covalent bond formed by two excited electrons (see Fig. 4), electron configuration $|\uparrow\downarrow_1, \uparrow\downarrow_2, \uparrow_3, \uparrow_4, \uparrow_5, \uparrow_6\rangle$. There are reasons to believe that the spins of the outer bonding electrons will be aligned parallel to each other so that the total spin of the He_4^* cluster becomes $S = 2$ and the "ferromagnetic" state will appear (see Fig. 4). This is very important for improving stability and for increasing the lifetime since the Pauli exclusion principle will prevent recombination (as in the case of He_2^*). This is the case, for example, in the O_2 molecule, in which the spins of the outer electrons are parallel (in accordance with Hund's rule, calculations, and experiment; see [18]). Finally, we can expect that He_2^* quasi-atoms (building blocks) preserve their individual properties in the He_4^* ensemble (cluster) since the orbits of the outer and inner electrons differ substantially.

In this section, we consider the results of calculations, which prove the existence of the He_4^* cluster whose properties virtually coincide with those described above. It should be noted above all that we can expect various geometrical configurations of helium atoms in the He_4^* cluster, e.g., in the form of a one-dimensional chain, rectangle (see Fig. 5), etc. Using the theoretical approaches described in Section 2, we performed numerous calculations of the total energy of various configuration of the He_4^* system with spin $S = 2$ as a function of the distance between quasi-atoms He_2^*. In addition, we calculated the vibrational spectrum which can be used to judge on the stability of a certain atomic configuration.

A large number of studied configurations proved to be unstable (in particular, the configuration in the form of a 1D chain). In this case, it was shown that the energy as a function of spacing between the He_2^* quasi-atoms has a minimum. However, such an atomic configuration is not metastable since it is not stable to transverse displacements of the atoms (the vibrational spectrum contains the corresponding imaginary frequencies). In other words, this configuration is a saddle point on the energy surface (dependence $E(\{\mathbf{R}_i\})$ of the total energy of the system on the coordinates of all atoms).

We have found only one metastable configuration in which the He_4^* cluster has the shape of a rectangle (see Fig. 5). The spacing between helium atoms in each He_2^* quasi-atom is $R_0 \approx 2a_B$ and is virtually the same as the distance between helium atoms in an isolated He_2^* quasi-atom. Figure 6 shows the dependence of the total energy E of the He_4^* system on distance R between the He_2^* quasi-atoms.

It can be seen from Fig. 6 that, for all computational methods used, quantity E has a minimum for the same distance $R_1 \approx 6a_B$ between quasi-atoms. All frequencies in the vibrational spectrum of the He_4^* cluster are real-valued (the minimal frequency is 174 cm^{-1}). It can be concluded that the configuration of He_4^* depicted in Fig. 5 corresponds to a local minimum of $E(\{\mathbf{R}_i\})$ and is metastable. The height of the potential barrier restrict the local minimum is $U \approx 0.5$ eV and depends on the method of calculation only slightly.

Figures 7a–7f show the self-consistent single-electron wavefunctions of six α orbitals (spin up) of the He_4^* cluster. The wavefunctions of two β orbitals (spin down) cannot be visually distinguished from the corresponding wavefunctions of the α orbitals

(see Figs. 7a and 7b). In accordance with the physical pattern described above, Fig. 7 clearly demonstrates that six "lower" electrons (four α orbitals and two β orbitals) are localized at quasi-atoms, while the two "upper" electrons are delocalized between the quasi-atoms and form the covalent bond between them.

Our analysis shows that the state of the He_4^* cluster with the set of self-consistent one-electron wavefunctions depicted in Fig. 7 (we will call it the $|A\rangle$ state) is not the only eigenstate of the He_4^* system with total spin $S = 2$. This follows, in particular, from the presence of a kink on the $E(R)$ curves for $\tilde{R} \approx 8.5a_B$ (see Fig. 6), which indicates that, with increasing R, the system passes from state $|A\rangle$ to a certain state $|B\rangle$ with another symmetry of the multielectron wavefunction. This is due to the fact that the energy $E_A(R)$ of state $|A\rangle$ is smaller than the energy $E_B(R)$ of state $|B\rangle$ for $R < \tilde{R}$, while $E_B(R) < E_A(R)$ for $R > \tilde{R}$.

It can be shown that the He_4^* cluster in state $|A\rangle$ is formed from quasi-atoms He_2^* ($^3\Sigma_u^+$) and He_2^* ($^3\Pi_g$) (see Section 3), while in state $|B\rangle$ it is formed from two quasi-atoms He_2^* ($^3\Sigma_u^+$). To be more precise, the wavefunction of the sixth α orbital in state $|A\rangle$ is a symmetric superposition of atomic $2p_z$ functions (the z axis is perpendicular to the plane of the cluster; see Fig. 7f) or, alternatively, a symmetric superposition of molecular $2p\pi_u$ functions; it is equal to zero in the plane of the cluster ($z = 0$) and has different signs for $z > 0$ and $z < 0$. In state $|B\rangle$, the wavefunction of the sixth α orbital is an antisymmetric superposition of molecular $2s\sigma_g$ functions (the wavefunctions of the remaining orbitals in states $|A\rangle$ and $|B\rangle$ are almost identical).

To verify the above idea, we calculated the energy of the system in states $|A\rangle$ and $|B\rangle$. Figure 8 shows the dependences of E_A and E_B on R, which were calculated by the MP2 method. The shape of the $E_A(R)$ and $E_B(R)$ curves and their intersection for $R = \tilde{R}$ explain the reason for the kink on the $E(R)$ curve in Fig. 6. Another fundamental result following from the shape of the $E_{A,B}(R)$ dependences is that the deepest local minimum (and, hence, the highest stability) of the He_4^* cluster is realized in state $|A\rangle$. It should be noted that the depth of the minimum in state $|B\rangle$ is very small and is within the computational error. Thus, the He_4^* cluster should be formed from He_2^* quasi-atoms in states $^3\Sigma_u^+$ and $^3\Pi_g$.

Let us estimate the energy accumulated in the He_4^* cluster. The end products of the decay of the metastable

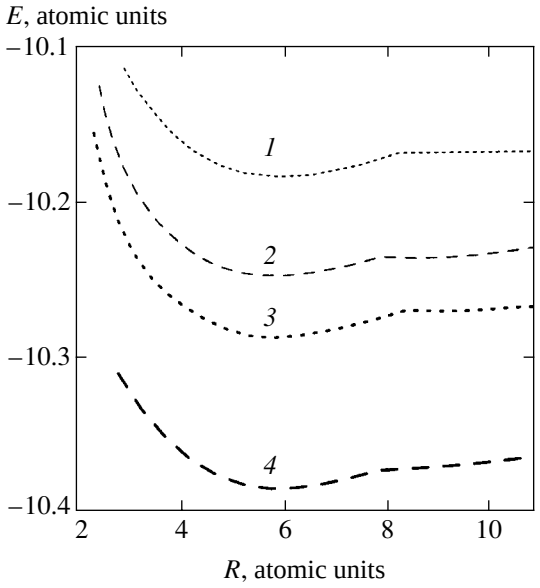

Fig. 6. Total energy E of the He_4^* system of four helium atoms with total spin $S = 2$ as a function of distance R between He_2^* quasi-atoms (see Fig. 5). Calculations were made using the UHF (*1*), LDA (*2*), MP2 (*3*), and B3LYP (*4*) methods.

He_4^* cluster are four He atoms in the singlet $1s^2$ states. The difference $E_{acc} = E(He_4^*) - 4E(He)$ is the energy accumulated in the metastable state. This energy can be released in the form of photons emitted during successive spin flips of two electrons and the passage of these electrons from excited α orbitals to the lower β orbitals as well as in the form of the kinetic energy of helium atoms (as the total spin changes from $S = 2$ to $S = 1$ and further to $S = 0$, the system of four helium atoms appears to be in the states which do not correspond to the local minimum of the total energy as a function of the coordinates of the atoms; i.e., these states are not metastable and hence decay into individual atoms). Our calculations give $E_{acc} = 33$–36 eV for the energy accumulated in the He_4^* cluster, which is approximately twice as large as the energy $E_{acc} = E(He_2^*) - 2E(He) \approx 18$ eV accumulated in the triplet He_2^* molecule. Thus, the specific value of accumulated energy (recalculated per helium atom) $E_{acc} \approx 9$ eV/atom is virtually the same for He_2^* and He_4^*. This is due to the fact that the interaction between quasi-atoms He_2^* in the He_4^* cluster is much weaker than the interaction between two He atoms in the He_2^* quasi-atom (which is manifested in the

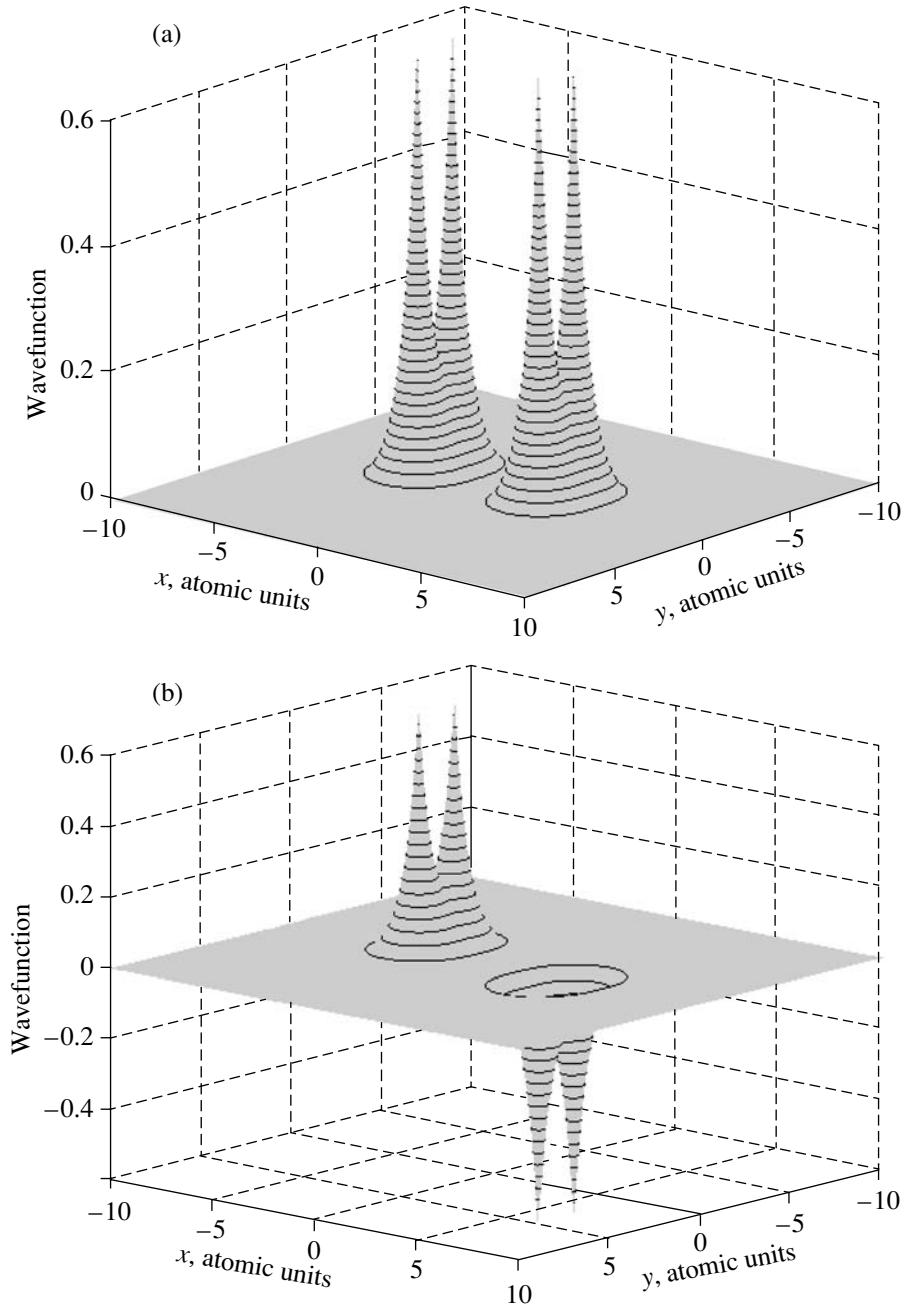

Fig. 7. Self-consistent single-electron wavefunctions of the six lower α orbitals of the metastable He_4^* cluster with total spin $S = 2$ (state $|A\rangle$) (see Figs. 4 and 5). Helium atoms are located at points with coordinates $(\pm 3a_B, \pm a_B, 0)$; $z = 0$ (a–e) and $0.25a_B$ (f) (see text). The wavefunctions of the two β orbitals are almost indistinguishable from the wavefunctions of the corresponding two lower α orbitals (a, b).

strong difference between the corresponding bond lengths; see Fig. 5).

6. LIFETIME OF THE He_4^* CLUSTER

Let us determine the lifetime of the He_4^* cluster. To our knowledge, this question has not been considered in the literature so far. A metastable He_4^* cluster with $S = 2$ may pass to the final state of the system of four helium atoms with $S = 0$ by emitting photons. This transition occurs via an intermediate state with spin $S = 1$, in which five electrons occupy α orbitals and three electrons occupy β orbitals (electron configuration $|\uparrow\downarrow_1, \uparrow\downarrow_2, \uparrow\downarrow_3, \uparrow_4, \uparrow_5\rangle$). Such a transition occurs upon spin flip of one of the excited electrons. Since the intermediate state for the configuration of four helium atoms

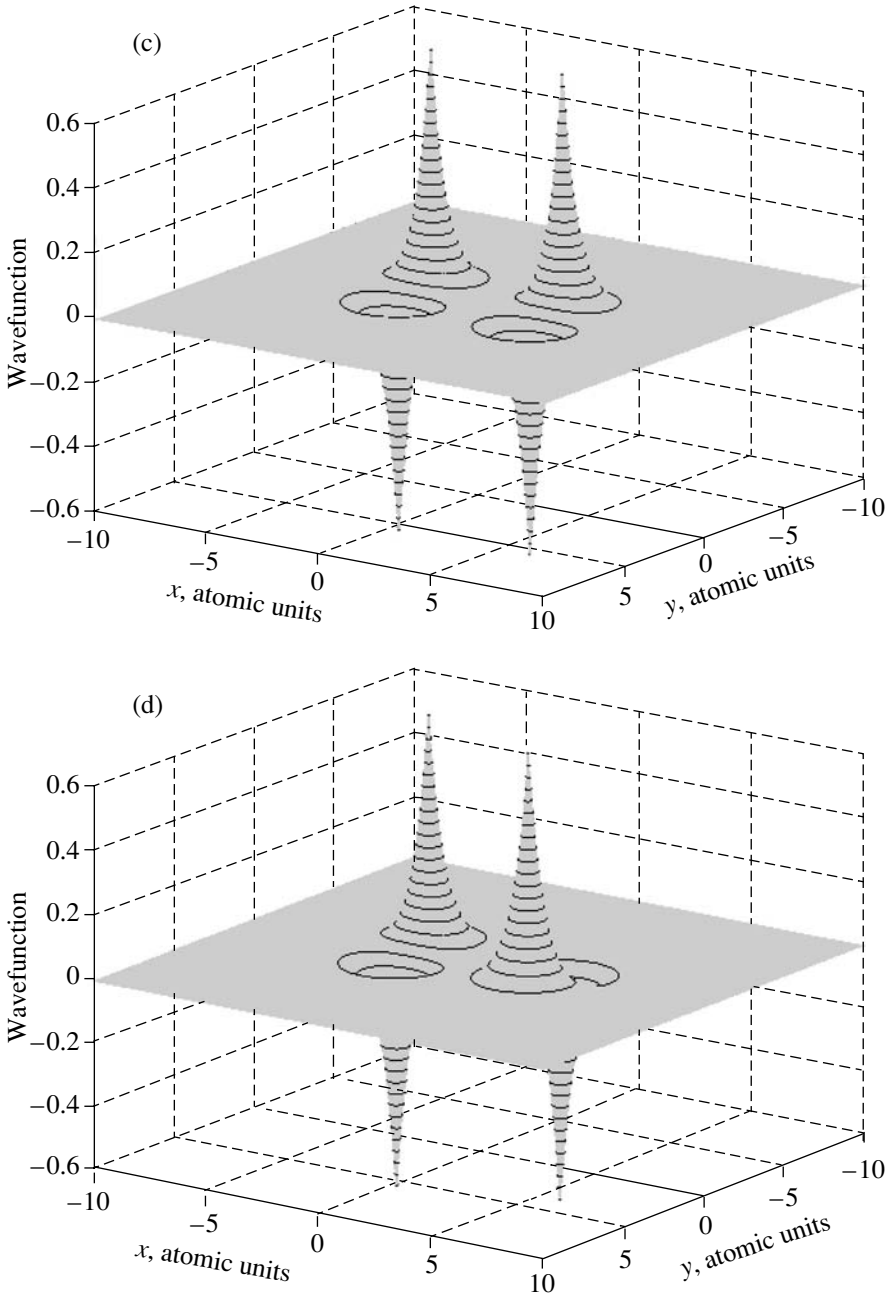

Fig. 7. (Contd.)

depicted in Fig. 5 has no local minimum, this state decays very rapidly. Thus, the lifetime τ of the He_4^* cluster is in fact determined by the time of its transition to the state with $S = 1$.

To calculate τ, we used formulas (6) and (7), in which the initial state is state $|A\rangle$ (see Section 5), while the state with $S = 1$ plays the role of state $|f\rangle$. There exist several intermediate states with spin $S = 1$. For illustration, we consider here two of such states, $|\mathrm{I}\rangle$ and $|\mathrm{II}\rangle$. State $|\mathrm{I}\rangle$ has the lowest possible energy for $S = 1$. It can be obtained as a result of transition of an excited electron from the atomic p_z orbital; consequently, the p_z component of all orbitals in state $|\mathrm{I}\rangle$ is zero (in this case, the self-consistent single-electron wavefunctions have the same form as the wavefunctions of the lower five α orbitals of the He_4^* cluster; see Fig. 7). For this reason, matrix elements $\mathbf{r}_{if}$ and $\mathbf{a}_{if}$ in formula (7) are equal to zero. Consequently, the probability $W_{i\to f}$ of the electron transition to state $|\mathrm{I}\rangle$ vanishes due to different symmetries of the wavefunctions of the initial and final states.

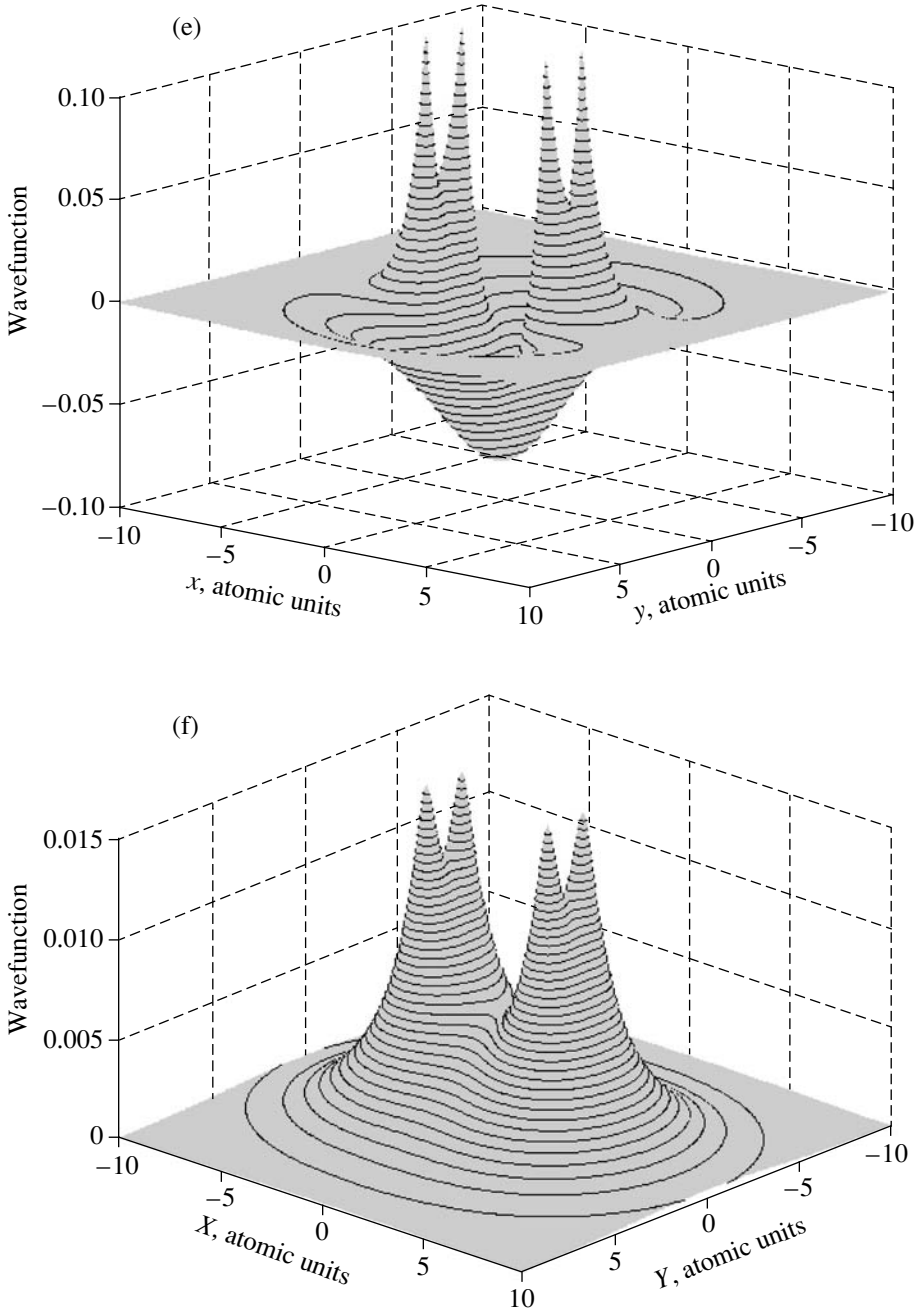

Fig. 7. (Contd.)

In the intermediate state $|II\rangle$, one of the orbitals has a p_z component. The energy of this state is higher than the energy of state $|I\rangle$. We calculated the multielectron wavefunction of state $|II\rangle$ and proved that the self-consistent single-electron wavefunctions of the four lower α orbitals and three β orbitals have the same form as the wavefunctions of the corresponding orbitals of the He_4^* cluster (see Fig. 7), while the wavefunction of the fifth α orbital is a superposition of the atomic $2p_z$ functions like the sixth α orbital of the He_4^* cluster.

As in the case of the triplet He_2^* molecule, we calculated matrix elements $\mathbf{r}_{if}$ and $\mathbf{a}_{if}$ (see formula (7)) using the wavefunctions determined by different methods. For the radiative transition probability $W_{i\to f}$, we obtained values of 4.7×10^{-3} s^{-1}, 6.9×10^{-3} s^{-1}, and 1.05×10^{-2} s^{-1} using the UHF, MP2, and LDA methods, respectively; this corresponds to lifetimes of $\tau = 213$, 144, and 95 s, respectively. Although the results obtained by different methods differ substantially, it can be concluded that the radiation lifetime $\tau = 100$–200 s of the

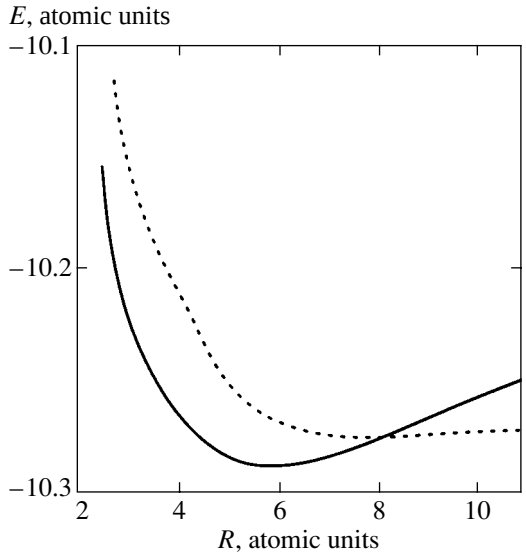

Fig. 8. Energies E_A (solid curve) and E_B (dotted curve) of two eigenstates of the He_4^* system with total spin $S = 2$ as a function of distance R between the He_2^* quasi-atoms (see text). Calculations were performed by the MP2 method.

metastable spin-polarized He_4^* cluster is approximately an order of magnitude longer than the lifetime $\tau \approx 20$ s of the triplet He_2^* molecule. The possible reason for this increase in the lifetime is associated with the emergence of a new type of symmetry (such as the translational symmetry in large systems).

7. CONCLUSIONS

The helium He_4^* cluster studied here is the simplest ensemble, consisting of two quasi-atoms He_2^* ($^3\Sigma_u^+$) and He_2^* ($^3\Pi_g$). We also obtained preliminary results indicating the possibility of the existence of large ensembles $(He_2^*)_N$ up to $N = 8$. The specific energy $E_{acc} \approx 9$ eV/atom accumulated in such ensembles is the same as in the He_4^* cluster.

Using He_2^* and He_4^* as examples, we proved that the radiation lifetime τ of the metastable state increases with the number of atoms in the system. Consequently, there is hope that ensembles $(He_2^*)_N$ with $N > 2$ will be even more long-lived. In addition, our calculations indicate that the charged $(He_4^*)^+$ cluster is also metastable. This makes it possible to detect experimentally metastable helium clusters using standard mass-spectroscopic methods [19, 20].

It should be noted that a certain analogy exists between the $(He_2^*)_N$ ensembles and the so-called Rydberg systems, which were studied in detail in [21] for alkali metals. However, the fundamental difference is that alkali metals are condensed under normal conditions and form type I clusters. On the contrary, spin-polarized helium clusters (type II clusters) are formed only in excited states from blocks (quasi-atoms); such systems have no global energy minimum as a function of atomic coordinates and accumulate large amounts of energy.

ACKNOWLEDGMENTS

This study was supported by the CRDF under the project "Fundamental Properties of Matter under Extreme Conditions."

REFERENCES

1. H. W. Kroto, J. R. Heath, S. C. O'Brien, *et al.*, Nature **318**, 162 (1985).
2. T. L. Makarova, B. Sundqvist, R. Hohne, *et al.*, Nature **413**, 716 (2001).
3. A. F. Hebard, M. J. Rosseinsky, R. C. Haddon, *et al.*, Nature **350**, 600 (1991).
4. V. F. Elesin, N. N. Degtyarenko, and L. A. Openov, Inzh. Fiz., No. 3, 2 (2002).
5. D. N. McKinsey, C. R. Brome, J. S. Butterworth, *et al.*, Phys. Rev. A **59**, 200 (1999).
6. C. F. Chabalowski, J. O. Jensen, D. R. Yarkony, and B. H. Lengsfield III, J. Chem. Phys. **90**, 2504 (1989).
7. A. V. Konovalov and G. V. Shlyapnikov, Zh. Éksp. Teor. Fiz. **100**, 521 (1991) [Sov. Phys. JETP **73**, 286 (1991)].
8. S. Sokolova, A. Lüchow, and J. B. Anderson, Chem. Phys. Lett. **323**, 229 (2000).
9. S. Fujinaga, *Method of Molecular Orbitals* (Iwanami Shoten, Tokyo, 1980; Mir, Moscow, 1983).
10. P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
11. W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
12. O. Gunnarsson and B. I. Lundqvist, Phys. Rev. B **13**, 4274 (1976).
13. A. D. Becke, Phys. Rev. A **38**, 3098 (1988).
14. C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).
15. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, *et al.*, J. Comput. Chem. **14**, 1347 (1993).
16. G. D. Fletcher, M. W. Schmidt, and M. S. Gordon, Adv. Chem. Phys. **110**, 267 (1999).
17. C. W. H. Drake, Phys. Rev. A **3**, 908 (1971).
18. C. A. Coulson, *Valence* (Oxford Univ. Press, Oxford, 1961; Mir, Moscow, 1965).
19. H. Buchenau, J. P. Toennies, and J. A. Northby, J. Chem. Phys. **95**, 8134 (1991).
20. B. von Issendorff, H. Haberland, R. Fröchtenicht, and J. P. Toennies, Chem. Phys. Lett. **233**, 23 (1995).
21. É. A. Manykin, M. I. Ozhovan, and P. P. Poluéktov, Khim. Fiz. **18**, 88 (1999).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 56–63.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 67–76.

ATOMS, MOLECULES, OPTICS

Wakefield Generation as the Mechanism behind Spectral Shift of a Short Laser Pulse

N. E. Andreev and M. V. Chegotov

Institute for High Energy Densities, Joint Institute for High Temperatures, Russian Academy of Sciences, Moscow, 127412 Russia

e-mail: chegotov@ihed.ras.ru

Received January 28, 2005

Abstract—An equation describing the dynamics of plasma wave generation by a short intense laser pulse is analyzed to find a relation between the difference in mean-square pulse frequency before and after laser–matter interaction and the electric field amplitude in the wakefield plasma wave generated by the laser pulse. This relation can be effectively used in systems for wakefield diagnostics. The relation is applied to several geometries of interaction between a pulse and an ionizing gas or preformed plasma. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Analysis of spectral characteristics of electromagnetic radiation after interaction with matter has important advantages as a method for investigating the processes taking place in the medium, in comparison with other diagnostic methods. In this optical method, the effect of the diagnostic intervention required to obtain the desired information about the processes in question is relatively weak (e.g., see [1]). In view of the current progress in ultrashort pulse optics, the rapid development of the processes associated with laser–matter interactions is particularly important for practical applications. The only approach that can ensure a high temporal resolution in studies of these processes is optical measurement (e.g., see [2–4]). Furthermore, the most promising optical diagnostic methods would be those in which spectral characteristics of an intense laser pulse are measured after interaction with matter without using additional electromagnetic field (probe pulse).

In the method proposed in [5], the mean-square pulse frequencies before and after interaction with matter measured in the wave zone relative to the interaction region are used to evaluate integral characteristics of electric field and current in that region for arbitrary focusing geometry and incident laser intensity. In particular, the difference in mean-square pulse frequency before and after interaction with an ionizing medium is related to the pulse field strength in the ionization region. In [6], the method was extended to an arbitrary even power of frequency averaged over an appropriate spectral density. Since a power of frequency averaged over an appropriate spectral density is a moment of the spectral density, the spectral density of a beam that has passed through the interaction region is completely determined by the integral characteristics of electric field and current in that region. Accordingly, the method developed in [5, 6] was called *moment method.* Note that moments of a spectral density, being integrals of frequency, provide a spatiotemporal characterization of the interaction between electromagnetic field and matter in this method.

A short laser pulse interacting with a preformed plasma generates relatively long-lived wakefield plasma waves, which can be used to accelerate electrons to ultrarelativistic energies [7]. Due to plasma wave generation, the pulse propagates in a rapidly varying medium and its spectrum changes. In [8], the moment method proposed in [5, 6] was extended to the case of weakly damped plasma waves. It was found that the difference in mean-square pulse frequency before and after interaction with the generated plasma wave is determined by the electron-density disturbance and laser field intensity in the region of plasma wave generation, as well as by the plasma wave amplitude after the interaction.

In this paper, an equation describing the dynamics of plasma wave generation by a laser pulse is analyzed to find a relation between the difference in mean-square pulse frequency before and after interaction and the electric field amplitude in the wakefield plasma wave generated by the laser pulse. We analyze generation of wakefield plasma waves in several geometries ("free" and channeled Gaussian laser pulse and pulse propagation in a capillary tube), assuming that the pulse duration is nearly resonant [9]. Since no satellites are generated at multiples of the plasma frequency in this case (e.g., see Fig. 6 below), treatment of the effect of plasma wave generation on the pulse based on the conventional theory of parametric processes [10, 11] cannot be used for diagnosing the wakefield plasma wave. However, the present theoretical analysis relying on general relations between moments makes it possible to determine the frequency shift for a moderate-intensity

laser pulse due to wakefield generation as a function of both material and pulse parameters. In particular, we show that the frequency shift is proportional to the pulse propagation distance in a channel or capillary for a weakly damped mode. Note also that the moment-based theory is consistent with general conservation laws and can therefore be used to verify the accuracy of numerical simulations.

Since self-consistent computations of nonlinear propagation of short intense laser pulses agree with the analytical results obtained here, these results can be used as a basis for an effective optical technique for diagnosing the plasma wakefields generated by such pulses in experiments on acceleration of electrons to ultrarelativistic energies.

2. MODEL

Self-consistent equations describing the dynamics of ultrashort laser pulse propagation in a gas, including ionization and ensuing plasma wave generation in the weakly relativistic regime, were derived in [12]. Laser pulse guiding and generation of wakefield plasma waves in the strongly relativistic regime were analyzed in [13]. The present study relies on the results obtained in [12, 13].

We represent the electric field $\mathscr{E}_L$ of a laser pulse propagating in a medium as

$$\mathscr{E}_L = \frac{1}{2}\mathbf{E}_L \exp(ik_0\xi) + \text{c.c.},$$

where $\xi = z - ct$, the z axis is aligned with the propagation direction, c is the speed of light, $k_0 = \omega_0/c$, ω_0 is the carrier frequency of the laser pulse, and $\mathbf{E}_L$ is the complex amplitude slowly varying over times and lengths on the order of $2\pi/\omega_0$ and $2\pi/k_0$. In the comoving frame $\xi z \mathbf{r}_\perp$, the equation for the dimensionless electric field amplitude of the laser pulse, $\mathbf{a} = e\mathbf{E}_L/m\omega_0 c$ (e and m denote the electron charge and mass, respectively), is written as

$$\left\{2ik_0\frac{\partial}{\partial z} + 2\frac{\partial^2}{\partial z \partial \xi} + \Delta_\perp\right\}\mathbf{a} = k_0^2\left(\frac{n}{n_c\gamma}\mathbf{a} - i\mathbf{G}^{(\text{ion})}\right). \quad (1)$$

Here, $\Delta_\perp$ is the transverse Laplace operator ($\mathbf{r}_\perp = \mathbf{e}_x x + \mathbf{e}_y y$, where $\mathbf{e}_x$ and $\mathbf{e}_y$ denote the unit basis vectors in the x and y directions, respectively); $n_c = m\omega_0^2/4\pi e^2$ is the critical electron density;

$$\gamma = \left[1 + \left(\frac{\mathbf{p}}{mc}\right)^2 + \frac{|\mathbf{a}|^2}{2}\right]^{1/2}$$

is the relativistic factor, with $\mathbf{p}$ denoting the slowly varying momentum of electrons in the plasma wake;

$$\mathbf{G}^{(\text{ion})} = \frac{2}{n_c\omega_0}\sum_{k=0}^{Z_n-1}\Gamma_0^{(k)}\left(\frac{\mathbf{a}}{|\mathbf{a}|^2}\frac{U_k}{mc^2} - \mu\frac{\mathbf{a}^*}{4}\right) \quad (2)$$

represents effects due to optical field ionization, which are described in terms of the ionization potential U_k corresponding to the increase in ion charge number from k to $k+1$, the slowly time-varying ionization rate per unit volume $\Gamma_0^{(k)}$, the nuclear charge Z_n of gas atoms, and the coefficient μ characterizing the contribution of the second harmonic to the ionization rate ($\mu = 0$ and 0.7–1.0 for circularly and linear polarized laser beams, respectively, the latter value depending on the gas and laser pulse parameters); and n is the electron density slowly varying over times on the order of $2\pi/\omega_0$. Details of the calculation of $\Gamma_0^{(k)}$ can be found in [14]. Note that the ionization current $\mathbf{G}^{(\text{ion})}$ given by (2) ensures correct balance of beam energy and momentum in the process of optical field ionization [14].

The evolution of free-electron density is governed by the equations of ionization kinetics and by the equations of motion for the electron plasma in a strong electromagnetic field. When $n \ll n_c$, the ratio n/γ in the nonlinear-response term in Eq. (1) can be expressed in terms of the ion charge density $|e|n_0$ as

$$\frac{n}{\gamma} = n_0\frac{1 + k_p^{-2}\Delta_\perp\Phi}{\Phi + \delta\Phi_S}, \quad (3)$$

where

$$k_p = \frac{\omega_p}{c}, \quad \omega_p = \left(\frac{4\pi e^2 n_0}{m}\right)^{1/2},$$

and the wakefield potential Φ depends on the z-component of dimensionless electron momentum $q_z = p_z/mc$ and laser pulse characteristics:

$$\Phi = \gamma - q_z + \int_\infty^\xi \frac{d\xi'}{n_0 c}$$
$$\times \sum_{k=0}^{Z_n-1}\Gamma_0^{(k)}\left(q_z - \frac{|\mathbf{a}|^2}{2} - \frac{\mu}{4}\text{Re}(\mathbf{a}^* \cdot \mathbf{a}^*)\right) = \gamma - q_z - \delta\Phi_S.$$

Assuming that the pulse intensity distribution is cylindrically symmetric and its characteristic radius is larger

than k_p^{-1}, then the equation for Φ is

$$\left\{(\Delta_\perp - k_p^2)\frac{\partial^2}{\partial\xi^2} - \frac{\partial \ln n_0}{\partial r}\frac{\partial^3}{\partial r\partial\xi^2} + k_p^2\Delta_\perp\right\}\Phi - \frac{k_p^4}{2}\left[1 - \frac{1 + |\mathbf{a}|^2/2}{(\Phi + \delta\Phi_S)^2}\right] = \frac{k_p^2}{4}\Delta_\perp|\mathbf{a}|^2, \quad (4)$$

and the slowly varying electric and magnetic field strengths in the plasma, $\mathbf{E}_p$ and $\mathbf{B}_p$, can be expressed in terms of Φ:

$$\frac{\partial\Phi}{\partial\xi} = \frac{eE_{p,z}}{m_e c^2}, \quad \frac{\partial\Phi}{\partial r} = \frac{eE_{p,r}}{m_e c^2} - \frac{eB_{p,\varphi}}{m_e c^2}, \quad (5)$$

where $B_{p,\varphi}$ is the azimuthal component of $\mathbf{B}_p$.

The total ion density n_0 is expressed in terms of the ion densities N_k ($k = 0$ corresponds to neutral particles) as follows:

$$n_0 = \sum_{k=1}^{Z_n} kN_k,$$

where the number densities N_k either are governed by local equations of ionization kinetics written for immobile ions and neutrals (e.g., see [15]) or are prescribed distributions determining the background plasma electron density n_0 (in the case of a preformed plasma). When optical field ionization is taken into account, the derivative

$$\frac{\partial n_0}{\partial t} = \sum_{k=0}^{Z_n - 1}\Gamma_0^{(k)}$$

and $|\mathbf{a}|^2$ determine the pulse frequency shift due to ionization [5, 6].

Equations (1)–(5) must be supplemented with boundary conditions. For a preformed plasma, we assume that the ion charge density $|e|n_0$ remains time-independent during the interaction between the pulse and the medium. For an ionizing gas ($N_k(\xi = \infty) = 0$ for $k \neq 0$), we assume that the gas concentration $N_0(\xi = \infty)$ is uniform. The boundary condition for Φ in the unperturbed gas is $\Phi = 1$ ($\xi \longrightarrow \infty$, $r \longrightarrow \infty$).

The incident laser pulse is assumed to have a Gaussian profile:

$$\mathbf{a}(\xi, r, z = +0) = \mathbf{a}_0\hat{A}(r)\exp\left[-\frac{(\xi - \xi_0)^2}{c^2\tau^2}\right], \quad (6)$$

where laser-pulse full width at half maximum (FWHM) τ_L reached at ξ_0 is related to τ as follows: $\tau_L = \sqrt{2\ln 2}$. The function $\hat{A}(r)$ characterizes the transverse distribution of the pulse amplitude **a**. For a laser pulse propagating in a laterally unbounded plasma, we set

$$\hat{A}(r) = A(r) \equiv \exp(-r^2/r_0^2),$$

which corresponds to the asymptotic boundary condition $a(\xi, r, z) \longrightarrow 0$ as $r \longrightarrow \infty$.

The initial radial profile $\hat{A}(r)$ of a laser pulse propagating in a capillary tube can be approximated by a finite superposition of hybrid modes:

$$\hat{A}(r) = \sum_{m=1}^{M} C_m J_0(k_{\perp m} r),$$

$$k_{\perp m} = \frac{b_m}{R_{cap}} - i\frac{b_m}{k_{w\perp}R_{cap}^2}, \quad (7)$$

$$k_{w\perp} = \frac{2(\varepsilon_w - 1)^{1/2}}{\varepsilon_w + 1}k_0,$$

$$C_m = \frac{2}{[R_{cap}J_1(b_m)]^2}\int_0^{R_{cap}} A(r)J_0\left(\frac{b_m r}{R_{cap}}\right)r dr,$$

where J_0 and J_1 are the first- and second-order Bessel functions, respectively; b_m is the mth root of the equation $J_0(b_m) = 0$; M is the number of eigenmodes $J_0(k_{\perp m}r)$ retained in the expansion of $\hat{A}(r)$; R_{cap} is the radius of the capillary tube; and ε_w is the dielectric constant of the capillary-tube wall. For a linearly polarized pulse ($\mathbf{a} = \mathbf{e}_0 a(\xi, r, z)$, where $\mathbf{e}_0$ is the unit polarization vector), the wall boundary condition (at $r = R_{cap}$) is

$$\frac{\partial\mathbf{a}}{\partial r} = ik_{w\perp}\left(1 - \frac{i}{k_0}\frac{\partial}{\partial\xi}\right)\mathbf{a}. \quad (8)$$

Note that boundary condition (6) with a real function $A(r)$ implies that the pulse is focused on the boundary of the medium, whereas wall boundary condition (8) leads to damping of the eigenmodes in (7) characterized by spatial damping rates $\delta k''_{zm}$:

$$\delta k''_{zm} = \frac{b_m^2}{k_0 R_{cap}^3}\frac{\mathrm{Re}\,k_{w\perp}}{|k_{w\perp}|^2} = \frac{b_m^2}{2k_0^2 R_{cap}^3}\frac{1 + \varepsilon_w}{\sqrt{\varepsilon_w - 1}}.$$

The damping is due to electromagnetic field penetration into the capillary wall and is described by the exact expression for the damping rates of the eigenmodes of the capillary tube obtained by taking into account the distributions of the field vectors [16].

3. FREQUENCY SHIFT

General relations for the shift in the mean-square frequency of the outgoing pulse were obtained in [5, 6]. For the pulse envelope $\mathbf{a}(\xi, r, z)$ defined above, the mean-square frequency of the laser pulse that has propagated to a distance z in the medium is expressed as

$$\langle\omega^2\rangle(z) = \left(\int_0^\infty\int_0^\infty |\mathbf{a}(\omega, r, z)|^2 r dr d\omega\right)^{-1}$$
$$\times \int_0^\infty (\omega+\omega_0)^2 \int_0^\infty |\mathbf{a}(\omega, r, z)|^2 r dr d\omega,$$

where $\mathbf{a}(\omega, r, z)$ is a Fourier component of the dimensionless electric field of the pulse. If that incident radiation has a narrow spectrum, then the mean-square pulse frequency prior to interaction is

$$\sqrt{\langle\omega^2\rangle(z=0)} \approx \omega_0.$$

Note that the mean-square frequency $\langle\omega^2\rangle(z)$ is an integral spectral property, whereas the function

$$I(\omega+\omega_0, z) \equiv \left(\max_\omega \int_0^\infty |\mathbf{a}(\omega, r, z)|^2 r dr\right)^{-1} \times \int_0^\infty |\mathbf{a}(\omega, r, z)|^2 r dr \quad (9)$$

is the integral of the normalized spectrum of the propagating electromagnetic field centered at the carrier frequency ω_0 over the transverse cross-sectional plane.

The frequency shift in the outgoing pulse is

$$\delta\omega(z) \equiv \frac{\langle\omega^2\rangle(z) - \langle\omega^2\rangle(0)}{2\sqrt{\langle\omega^2\rangle(0)}} \approx \frac{\langle\omega^2\rangle(z) - \omega_0^2}{2\omega_0}. \quad (10)$$

In [8], an expression for the frequency shift in terms of the electron-density disturbance and the pulse intensity in the region of wakefield generation was derived by assuming that $\delta\omega(z)$ is smaller than ω_0 and using the weakly relativistic approximation. The analysis presented below makes use of results obtained in [5, 6, 8].

When the peak pulse intensity is much higher than that corresponding to the optical ionization threshold, ionization and wakefield generation are separated in space and time. In this case, it can be shown that the total frequency shift due to both processes is the sum of the shift $\delta\omega_{ion}/\omega_p$ due to ionization and the shift $\delta\omega_{wf}/\omega_p$ associated with wakefield generation:

$$\frac{\delta\omega}{\omega_p} = \frac{\delta\omega_{ion}}{\omega_p} + \frac{\delta\omega_{wf}}{\omega_p}. \quad (11)$$

Since the ionization-induced frequency shift is smaller than that caused by wakefield generation (see below), we first consider the relations that determine $\delta\omega_{wf}/\omega_p$.

In the weakly relativistic approximation ($|\mathbf{a}| < 1$, $|\delta\Phi| \equiv |\Phi - 1| \ll 1$), assuming that the contribution due to ionization is negligible, we derive the following expression for the plasma density disturbance $\delta n = n - n_0$ from (3):

$$\frac{dn}{n_0} = (k_{p0}^{-2}\Delta_\perp - 1)\delta\Phi + \frac{1}{4}|\mathbf{a}|^2. \quad (12)$$

Then, if the background ion charge density $|e|n_0$ is assumed to be uniform in space (or weakly nonuniform over lengths on the order of k_p^{-1}), Eq. (4) for the potential reduces to

$$\left(\frac{\partial^2}{\partial\xi^2} + k_p^2\right)\delta\Phi = \frac{1}{4}|\mathbf{a}|^2. \quad (13)$$

Combined with (5), this equation yields the wakefield amplitude $\mathbf{E}_p$:

$$\mathbf{E}_{p,\max}^2 = \frac{m^2c^2\omega_p^2}{e^2}\frac{1}{16}$$
$$\times \left\{k_p^2\left|\int_{-\infty}^\infty d\xi \exp(-ik_p\xi)|\mathbf{a}|^2\right|^2\right.$$
$$\left. \times \left|\frac{\partial}{\partial r}\int_{-\infty}^\infty d\xi \exp(-ik_p\xi)|\mathbf{a}|^2\right|^2\right\}. \quad (14)$$

The pulse frequency shift can be expressed in terms of this amplitude (by using Eqs. (12), (13), and (3) from [8]) as

$$\frac{\delta\omega_{wf}}{\omega_p} = -\frac{\omega_0}{\omega_p}\frac{1}{\mathscr{E}_{out}}\frac{1}{8\pi}\int_V \mathbf{E}_{p,\max}^2 d\mathbf{r}, \quad (15)$$

where $\mathscr{E}_{out}$ is the total energy of the outgoing pulse. Note that an expression similar to (14) was obtained for the electron density disturbance $n - n_0$ in [17] (see also [9]).

When a laser pulse interacts with a preformed plasma or an ionizing medium, electrons are driven by

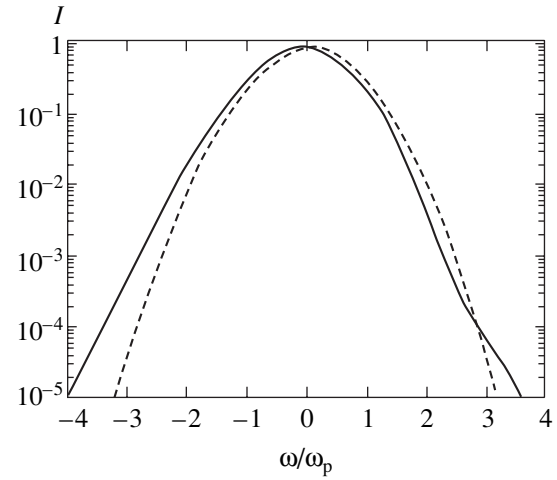

Fig. 1. Spectral density $I(\omega + \omega_0, z)$ defined by (9) versus the normalized frequency ω/ω_p at pulse propagation distances in plasma $z = 0$ (dashed curve) and $6z_R$ (solid curve) for $n_0 = 0.98 \times 10^{17}$ cm^{-3}, $a_0 = 0.14$, $k_p r_0 = 16$, $\tau_L = 100$ fs, $P_L/P_{cr} = 0.17$, and $\Omega = 1.5$.

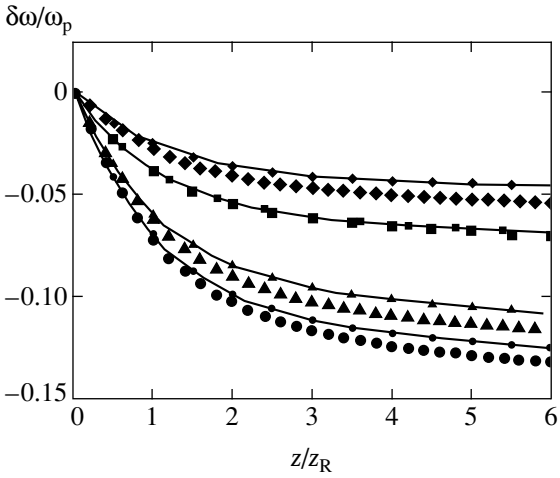

Fig. 2. Normalized frequency shift (10) versus pulse propagation distance for several values of pulse duration: Ω = 1.5 (●), 2.0 (▲), 0.5 (■), and 3.0 (◆). The curves corresponding to these values of Ω are plotted by using (14). The pulse and plasma parameters are $n_0 = 0.98 \times 10^{17}$ cm^{-3}, a_0 = 0.14, $k_p r_0 = 16$, and $P_L/P_{cr} = 0.17$.

ponderomotive force out of the region of high pulse intensity. As a result, a wakefield plasma wave is generated, and the pulse propagates in a rapidly varying medium whose reciprocal effect on the pulse leads to (relativistic and ponderomotive) self-focusing and changes in the pulse spectrum. This study is focused on the change in the pulse spectrum due to its interaction with the medium.

In what follows, we discuss numerical solutions of nonlinear problem (1)–(8) and compare computed characteristics of a laser pulse with analytical expressions (15) derived by using linearized model (12)–(14) for several geometries (semi-infinite preformed plasma, a channel in a semi-infinite preformed plasma, a capillary tube filled with a homogeneous neutral gas) for different values of pulse and material parameters. To obtain closed-form analytical expressions for frequency shift (15), simplified expressions for the laser pulse amplitude are used in (14), where the reciprocal effect of the wakefield plasma wave on pulse propagation is neglected.

3.1. Gaussian Pulse Propagation in a Semi-infinite Preformed Plasma

Consider a Gaussian pulse (with $\hat{A}(r) = A(r) \equiv \exp(-r^2/r_0^2)$ in (6)) focused onto a semi-infinite preformed plasma (with boundary at $z = 0$) characterized by an initial electron density n_0. When n_0 is negligible as compared to n_0, the time-dependent Gaussian pulse amplitude can be represented as

$$\mathbf{a}(\xi, r, z) = \frac{\mathbf{a}_0}{1 + iz/z_R} \times \exp\left[-\frac{r^2/r_0^2}{1 + iz/z_R} - \frac{(\xi - \xi_0)^2}{c^2\tau^2}\right]. \quad (16)$$

Assuming that this approximation is valid, we can estimate squared wakefield amplitude (14). Then, the frequency shift given by (15) is a function of the propagation distance z of the pulse:

$$\frac{\delta\omega_{wf}}{\omega_p} = -\frac{1}{64}\left(\frac{\pi}{2}\right)^{1/2} \times a_0^2 k_p^2 r_0^2 \left(1 + \frac{2}{k_p^2 r_0^2}\right) D(\Omega) \arctan\frac{z}{z_R}, \quad (17)$$

where $z_R = \omega_0 r_0^2/2c$ is the Rayleigh length and $D(\Omega) = \Omega\exp(-\Omega^2/4)$ is a function of $\Omega = \omega_p\tau$ reaching its maximum at $\Omega_{max} = \sqrt{2}$. Note that the resonant behavior of frequency shift (17) reflects the resonant nature of wakefield plasma wave generation, with a maximum amplitude reached at $\Omega = 2$ [9].

In Fig. 1, the spectral density $I(\omega + \omega_0)$ defined by (9) is shown as a function of the normalized frequency ω/ω_p obtained in self-consistent computations of (1)–(6) at $z = 0$ and $6z_R$ for $n_0 = 0.98 \times 10^{17}$ cm^{-3}, $a_0 = 0.14$, $k_p r_0 = 16$, $P_L/P_{cr} = 0.17$, and $\Omega = 1.5$ corresponding to $\tau_L = 100$ fs (P_L is the laser pulse power, and P_{cr} is the relativistic self-focusing threshold). It is clear that the pulse spectrum shifts toward lower frequencies as the pulse propagates into the plasma; i.e., the fre-

quency shift $\delta\omega(z)/\omega_p$ defined by (10) becomes increasingly negative (see Fig. 2).

Figure 2 compares the integral frequency shift $\delta\omega_{wf}/\omega_p$ obtained in self-consistent computations with $\delta\omega_{wf}/\omega_p$ calculated by using (14)–(16) as a function of z for several values of Ω (both peak intensity and radius of the incident pulse and the background plasma electron density are as in Fig. 1). It is obvious that the numerical results and analytical predictions are in good agreement. The relatively small discrepancy is explained by the fact that nonlinear reciprocal effects on the pulse amplitude were neglected in calculating $\delta\omega_{wf}/\omega_p$ for freely diffracting Gaussian beams. Indeed, Fig. 3a demonstrates that the effect of self-focusing on the pulse field amplitude on the axis $r = 0$ builds up as the pulse propagates deeper into the plasma, while Fig. 3b shows that the discrepancy between $\delta\omega(z)/\omega_p$ and $\delta\omega_{wf}/\omega_p$ increases with z. The insignificant discrepancy obtained for relatively small Ω (see Fig. 2) is explained by mutual compensation of nonlinear ponderomotive and relativistic effects [13]. Our computations have also shown that the discrepancy between $\delta\omega(z)/\omega_p$ and $\delta\omega_{wf}/\omega_p$ decreases with P_L/P_{cr}.

3.2. Pulse Propagation in a Plasma Channel

For a pulse propagating in a preformed plasma channel of radius R_{ch} such that $R_{ch} = k_p r_0^2/2$, the pulse shape calculated by neglecting the nonlinear effects is independent of z:

$$\mathbf{a}(\xi, r, z) = \mathbf{a}_0 \exp\left[-\frac{(\xi-\xi_0)^2}{c^2\tau^2} - \frac{r^2}{r_0^2}\right], \qquad (18)$$

and the transverse electron-density profile is

$$n_0(r) = n_0(0)\left(1 + \frac{r^2}{R_{ch}^2}\right).$$

Substituting (18) into (14) and (15), we find that the integral frequency shift is a linear increasing function of the pulse propagation distance:

$$\frac{\delta\omega_{wf}}{\omega_p} = -\frac{1}{16\pi}\left(\frac{\pi}{2}\right)^{3/2} a_0^2 \frac{\omega_p}{\omega_0}\left(1 + \frac{4}{k_p^2 r_0^2}\right) D(\Omega) k_p z. \qquad (19)$$

Figure 4 demonstrates that the frequency shift given by (19) is close to the value of $\delta\omega(z)/\omega_p$ obtained in a self-consistent computation for relatively low P_L ($P_L/P_{cr} = 0.082$). Note that $\delta\omega(z)/\omega_p$ and $\delta\omega_{wf}/\omega_p$

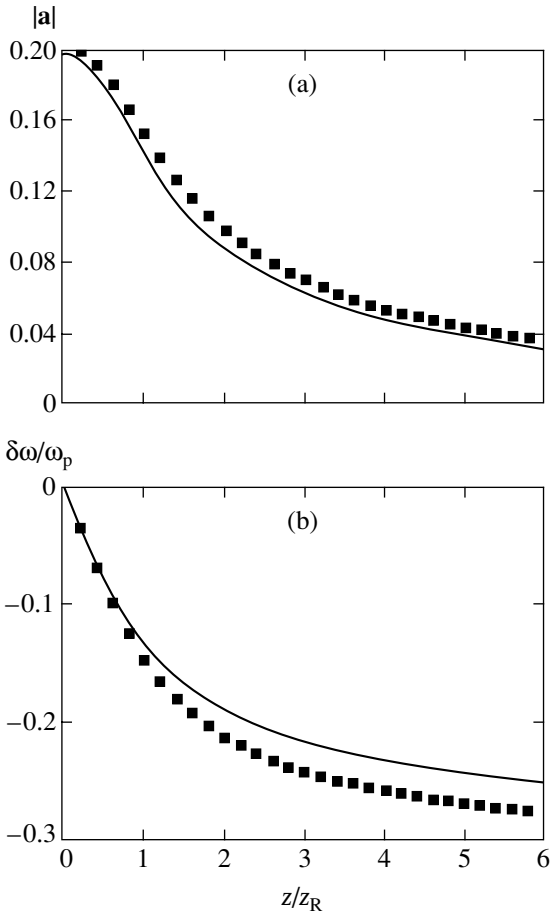

Fig. 3. (a) Pulse electric-field amplitude $|\mathbf{a}|$ on the axis $r = 0$ and (b) frequency shift (10) versus the pulse propagation distance z. The pulse and plasma parameters correspond to Fig. 1. Curves and symbols are predicted by Eq. (14) and numerical calculation, respectively.

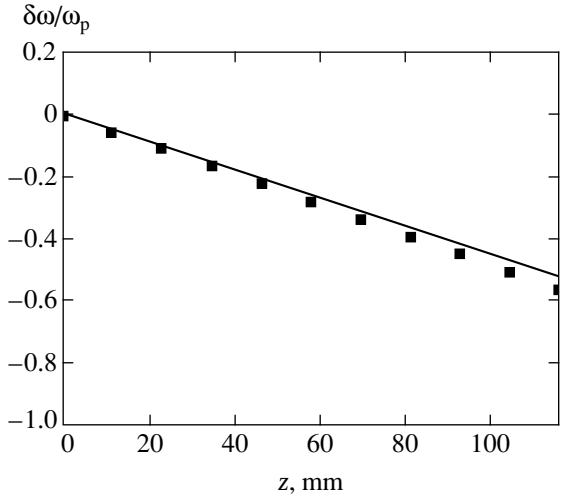

Fig. 4. Plots of $\delta\omega/\omega_p$ (■) and $\delta\omega_{wf}/\omega_p$ predicted by (19) (curve) versus the plasma-channel length z. The pulse and channel parameters are $n_0 = 0.98 \times 10^{17}$ cm^{-3}, $P_L = 10$ TW, $k_p r_0 = 4$, and $R_{ch} = 136$ μm.

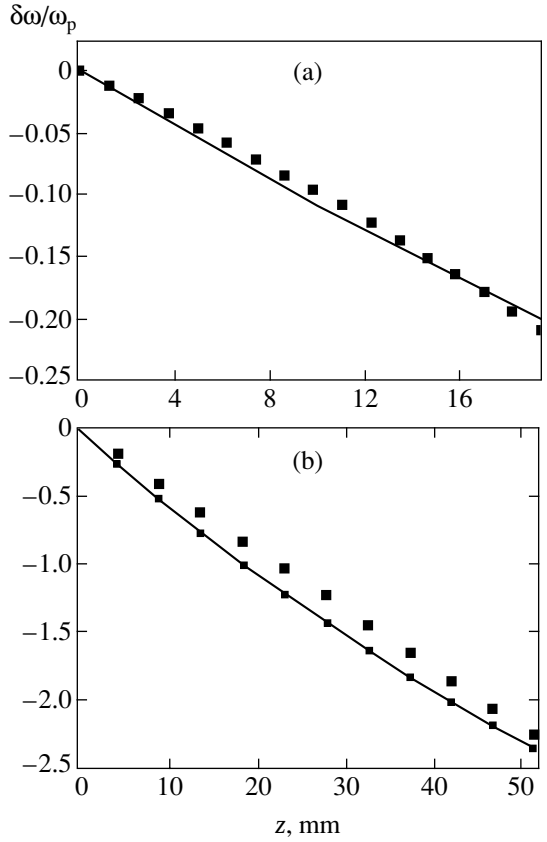

Fig. 5. Plots of $\delta\omega/\omega_p$ (■) and $\delta\omega_{wf}/\omega_p$ predicted by (18) (curves) versus the length z of a channel filled with ionizing hydrogen: (a) capillary tube diameter $D_{cap} = 75$ μm, initial hydrogen concentration $N_0 = 4.1 \times 10^{18}$ cm^{-3}, $P_L = 0.8$ TW, $k_p r_0 = 9.5$, $P_L/P_{cr} = 0.12$; (b) $D_{cap} = 77$ μm, $N_0 = 6.7 \times 10^{17}$ cm^{-3}, $\lambda_0 = 0.82$ μm, $\tau_L = 50$ fs, $k_p r_0 = 3.8$, $P_L = 9.6$ TW, $P_L/P_{cr} = 0.23$.

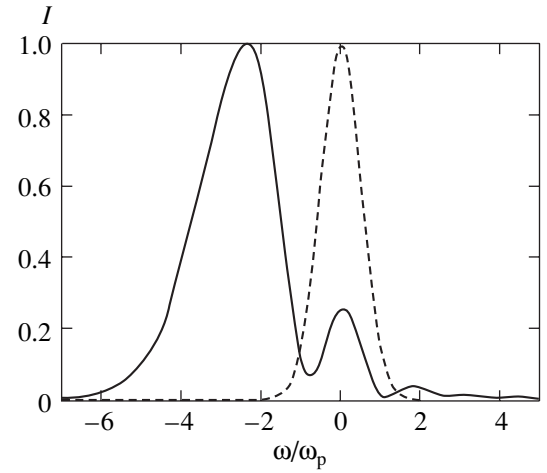

Fig. 6. Spectral density $I(\omega + \omega_0, z)$ defined by (9) versus the normalized frequency ω/ω_p at the pulse propagation distances in a capillary tube $z = 0$ (dashed curve) and 5.2 cm for $n_0 = 6.7 \times 10^{17}$ cm^{-3}, $\lambda_0 = 0.82$ μm, $\tau_L = 50$ fs, $k_p r_0 = 3.8$, $P_L = 9.6$ TW, $P_L/P_{cr} = 0.23$, $D_{cap} = 77$ μm, and $r_0/R_{cap} = 0.645$.

remain nearly equal as the pulse propagates to a distance of 10 cm and beyond.

3.3. Pulse Propagation in an Ionizing Gas in a Capillary Tube

Let us discuss the generation of a wakefield plasma wave in a capillary dielectric tube filled with an ionizing gas. When the focal spot size of a Gaussian pulse focused onto the capillary entrance is matched with the inner radius of the capillary tube, about 98% of the incident energy is transferred to the fundamental capillary mode, which corresponds to $m = 1$ in (7) [13, 16]. Accordingly, we calculate frequency shift (15) by using (14) with

$$|\mathbf{a}(\xi, z, r)|^2 = |\mathbf{a}_0|^2 \times \exp\left[-2\delta k_{z1}'' z - 2\frac{(\xi - \xi_0)^2}{c^2\tau^2}\right] J_0^2\left(b_1 \frac{r}{R_{cap}}\right), \quad (20)$$

where $b_1 \approx 2.405$ is the first root of the Bessel function J_0. When the expression for $\delta k_{z1}''$ presented at the end of Section 2 is used to characterize the attenuation of a pulse propagating in a capillary tube due to energy leakage through the capillary wall, formula (15) yields

$$\frac{\delta\omega_{wf}}{\omega_p} = -\frac{1}{64}\sqrt{\frac{\pi}{2}}\frac{\omega_p}{\omega_0} a_0^2 D(\Omega)\frac{k_p}{\delta k_{z1}''} \times [1 - \exp(-4\delta k_{z1}'' z)]\left(C_1 + \frac{C_2}{k_p^2 R_{cap}^2}\right), \quad (21)$$

where the constants C_1 and C_2 are defined as follows:

$$C_1 \equiv \int_0^{b_1} x J_0^4(x)dx \left(\int_0^{b_1} x J_0^2(x)dx\right)^{-1} \approx 0.5655,$$

$$C_2 \equiv 4b_1^2 \int_0^{b_1} x J_0^2(x) J_1^2(x)dx \left(\int_0^{b_1} x J_0^2(x)dx\right)^{-1} \approx 4.361.$$

In addition to frequency shift (21) associated with wakefield generation, the pulse spectrum exhibits a shift due to ionization, and the total shift in the mean-square frequency is the sum in (11). The value of $\delta\omega_{ion}/\omega_p$ can be estimated by combining formulas obtained in [5, 6] with the field prescribed by (20). In particular, the ionization-induced frequency shift calculated for the pulse and capillary parameters corresponding to Fig. 5b is

$$\frac{\delta\omega_{ion}}{\omega_p} = 0.635 \times 10^{-2} z,$$

where the propagation distance z is measured in centimeters. Note that the ionization-induced frequency

shift is much smaller than that due to wakefield plasma wave generation under the conditions discussed here. The resultant frequency shift shown in Fig. 5 for different pulse and capillary parameters demonstrates good agreement between numerical solutions to Eqs. (1)–(8) and the frequency shifts predicted analytical expression (21).

Note that the conventional theory of parametric processes [10, 11] cannot be applied when both laser pulse and plasma have parameters corresponding to Fig. 5b, because the pulse is too short as measured in units of the plasma wave period ($\Omega = 2$). In particular, it should not be expected that the spectrum of the outgoing pulse would contain satellites corresponding to harmonics of the plasma frequency ω_p. In Fig. 6, the spectral density $I(\omega + \omega_0, z)$ defined by (9) is shown as a function of the normalized frequency ω/ω_p obtained for the pulse and capillary parameters corresponding to Fig. 5b at the capillary entrance and at the propagation distance $z = 5.2$ cm. The spectrum contains pronounced peaks at $\omega = 0$ and $-2.4\omega_p$. Note that the latter peak frequency does not match any harmonic of ω_p.

4. CONCLUSIONS

The change in the frequency spectrum of a short intense laser pulse propagating in a medium is analyzed by taking into account wakefield plasma wave generation. Since wakefield generation by a laser pulse leads to a change in electron density, the pulse propagates in a rapidly varying medium. The reciprocal effect of the electron-density disturbance on the pulse changes the pulse spectrum as the pulse propagates deeper into the medium.

A linearized model is used to find expression (15) for the frequency shift in terms of the wakefield amplitude. For a pulse of nearly resonant duration, closed-form analytical expressions (17), (19), and (21) corresponding to different geometries are obtained for the shift in the mean-square frequency. In particular, it is shown that the frequency shift is proportional to the pulse propagation distance in a channel or capillary for a weakly damped mode.

Since self-consistent computations of nonlinear propagation of short intense laser pulses agree with the analytical results obtained here, these results can be used as a basis for an effective optical technique for diagnosing plasma wakefields generated by such pulses in experiments on acceleration of electron to ultrarelativistic energies.

ACKNOWLEDGMENTS

This work was supported by the Presidium of the Russian Academy of Sciences (Basic Research Program nos. 2 and 16) and by the Russian Foundation for Basic Research (project no. 04-02-17055).

REFERENCES

1. *Radiation Processes in Plasmas*, Ed. by G. Bekefi (Wiley, New York, 1966; Mir, Moscow, 1971).
2. F. Amiranoff, D. Bernard, B. Cros, *et al.*, IEEE Trans. Plasma Sci. **24**, 296 (1996); G. W. Siders, S. P. Le Blanc, A. Babine, *et al.*, IEEE Trans. Plasma Sci. **24**, 301 (1996).
3. P. Gibbon, F. Jakober, P. Monot, and T. Auguste, IEEE Trans. Plasma Sci. **24**, 343 (1996).
4. A. A. Solodov, P. Mora, and P. Chessa, Phys. Plasmas **6**, 503 (1999).
5. M. V. Chegotov, Zh. Tekh. Fiz. **72** (8), 71 (2002) [Tech. Phys. **47**, 1002 (2002)].
6. M. V. Chegotov, Kvantovaya Élektron. (Moscow) **33**, 370 (2003).
7. E. Esarey, P. Sprangle, J. Krall, and A. Ting, IEEE Trans. Plasma Sci. **224**, 252 (1996); N. E. Andreev and L. M. Gorbunov, Usp. Fiz. Nauk **169**, 53 (1999) [Phys. Usp. **42**, 49 (1999)].
8. M. V. Chetogov, in *Proceedings of Scientific Session of Moscow Institute of Engineering Physics, MIFI-2004* (Mosk. Inzh.–Fiz. Inst., Moscow, 2004), p. 163.
9. L. M. Gorbunov and V. I. Kirsanov, Zh. Éksp. Teor. Fiz. **93**, 509 (1987) [Sov. Phys. JETP **66**, 290 (1987)]; Tr. Fiz. Inst. im. P. N. Lebedeva, Ross. Akad. Nauk **219**, 3 (1992).
10. N. Bloembergen, *Nonlinear Optics* (Benjamin, New York, 1965; Mir, Moscow, 1966).
11. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984; Nauka, Moscow, 1989).
12. N. E. Andreev, M. V. Chegotov, M. C. Downer, *et al.*, IEEE Trans. Plasma Sci. **28**, 1090 (2000); N. E. Andreev, M. V. Chegotov, and A. A. Pogosova, Zh. Éksp. Teor. Fiz. **123**, 1006 (2003) [JETP **96**, 885 (2003)].
13. N. E. Andreev, Y. Nishida, and N. Yugami, Phys. Rev. E **65**, 056407 (2002); N. E. Andreev and S. V. Kuznetsov, Plasma Phys. Controlled Fusion **45**, A39 (2003); N. E. Andreev, B. Cros, L. M. Gorbunov, *et al.*, Phys. Plasmas **9**, 3999 (2002).
14. N. E. Andreev, M. E. Veĭsman, M. G. Keĭdzhyan, and M. V. Chegotov, Fiz. Plazmy **26**, 1010 (2000) [Plasma Phys. Rep. **26**, 947 (2000)]; N. E. Andreev, M. V. Chegotov, and M. E. Veisman, IEEE Trans. Plasma Sci. **28**, 1098 (2000).
15. V. P. Kandidov, O. G. Kosareva, and S. A. Shlenov, Kvantovaya Élektron. (Moscow) **21**, 971 (1994); N. E. Andreev, M. V. Chegotov, M. E. Veĭsman, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 566 (1998) [JETP Lett. **68**, 592 (1998)].
16. F. Dorchies, J. R. Marque's, B. Cros, *et al.*, Phys. Rev. Lett. **82**, 4655 (1999); C. Courtois, A. Couairon, B. Cros, *et al.*, Phys. Plasmas **8**, 3445 (2001); B. Cros, C. Courtois, G. Matthieussent, *et al.*, Phys. Rev. E **65**, 026405 (2002).
17. W. P. Leemans, C. W. Siders, E. Esarey, *et al.*, IEEE Trans. Plasma Sci. **24**, 331 (1996).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 64–72.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 77–87.

ATOMS, MOLECULES, OPTICS

Sub-Poissonian Statistics of Fluorescence from a Single Atom Driven by a Continuous Wave Laser Field

I. S. Osad'ko

Lebedev Physical Institute, Russian Academy of Sciences, Moscow, 119991 Russia
e-mail: osadko@sci.lebedev.ru

Received February 4, 2005

Abstract—Statistics of fluorescence photons emitted by a single two-level atom interacting with a continuous wave laser field are analyzed. The photon-counting distribution is calculated for the so-called intermediate fluorescence photons, i.e., those counted during the time interval between instants at which photons are emitted. The result is a sub-Poissonian (narrower than Poisson) distribution, which agrees with experimental observations. This intermediate-photon distribution is used to calculate the average number of fluorescence photons, the second factorial moment of the photon-counting distribution, and Mandel's *Q* parameter commonly used to evaluate the deviation of photon statistics from the Poisson distribution. The theoretical expressions obtained for moments of the intermediate-photon distribution are different from well-known Mandel's formulas. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Generally, light is emitted by an ensemble of excited atoms and consists of many photons. Statistics of photons in a beam and their dependence on the nature of the beam were discussed in a number of books focused on quantum characteristics of the laser field, fluctuations of the field, and beam coherence [1–5].

In particular, the light emitted by a thermal source was represented in [2] as a superposition of random contributions due to many atoms, where each contribution varies with time and depends on the location of the corresponding atom in the source. Mathematically, light waves of this kind are treated as fluctuating random fields. Depending on the source, the phases, amplitudes, and frequencies of waves can fluctuate independently or simultaneously. According to the authors of [2], their approach is essentially similar to that used in analysis of background noise in a radio receiver.

The development of lasers, which generate coherent light, motivated studies of optical coherence. It was established in [1] that "classical" coherent light is characterized by the Poisson distribution of the photon number P_n:

$$P_n(\alpha) = \frac{\alpha^{2n}}{n!}\exp(-\alpha^2), \tag{1}$$

where the parameter α is proportional to the electric-field amplitude.

However, there exists an aspect of radiation never mentioned in these books: the radiation emitted by a single atom interacting with a continuous wave laser field. Light emission by a single atom interacting with a continuous wave laser field is an essentially quantum process. It cannot be described by methods of classical physics, which were frequently used as the starting points of analyses in [1–5]. Therefore, one is led to a natural question: what distribution function describes the radiation emitted by a single atom interacting with a continuous wave laser field?

This question was posed in Mandel's pioneering study [6], where the probability of counting *n* photons within a time interval *T* was expressed as the quantum statistical average of a normally ordered time-ordered product of intensity operators *I* [5–8]:

$$N(n, T) = \left\langle T: \frac{1}{n!}\left[\int_t^{t+T} dx\hat{I}(x)\right]^n \times \exp\left[-\int_t^{t+T} dx\hat{I}(x)\right]: \right\rangle. \tag{2}$$

This probability can be used to obtain the first and second factorial moments:

$$\langle n(T)\rangle = \sum_{n=0}^{\infty} nN(n, T) = \left\langle : \left[\int_t^{t+T} dx\hat{I}(x)\right]: \right\rangle = \langle I\rangle T, \tag{3}$$

$$\begin{aligned}\langle n^{(2)}(T)\rangle &= \sum_{n=0}^{\infty} n(n-1)N(n,T)\\ &= \left\langle T:\left[\int_t^{t+T} dx\hat{I}(x)\right]^2:\right\rangle\\ &= \int_0^T dx\int_0^T dy\langle T:\hat{I}(x)\hat{I}(y)\rangle\\ &= \langle I\rangle^2\int_0^T dx\int_0^x dy g^{(2)}(x-y),\end{aligned} \tag{4}$$

where the fluorescence autocorrelation function $g^{(2)}(t)$ can be measured in an experiment. Since these expressions are derived for a stationary field, time is set to zero.

Formulas (2)–(4) apply both to an atomic ensemble and to a single atom. If the number $N_n(T, t)$ of sample intervals of length T within which n fluorescence photons are counted is measured during a total observation time t, then the probability of such an event is

$$w_n(T) = \lim_{t\to\infty}\frac{N_n(T,t)}{N_0(T,t)+\sum_{n=1}^{\infty}N_n(T,t)}. \tag{5}$$

This probability can be compared to theoretical formula (2): the measured probability $w_n(T)$ must be equal to the probability $N(n, T)$ predicted by (2) for a particular physical model.

Formula (2) was used to calculate the photon-number probability distribution for several simple models in [5]. In particular, the photon number for a coherent state of monochromatic field was found to be described by the Poisson function. The Poisson distribution was also obtained for a single-mode laser field with random phase, whereas a binomial distribution was found for a single-mode Fock state. The polarized light emitted by a thermal source was shown to be described by the Bose–Einstein distribution.

It may seem surprising, but the probability defined by (2) is more difficult to calculate for a single atom than for an atomic ensemble. This explains why this calculation has never been performed. In [6], a simpler calculation of moments of the distribution was presented. Using these moments, Mandel introduced the following parameter as a measure of deviation of photon-counting statistics from the Poisson distribution:

$$Q(T) = \frac{\langle n^{(2)}(T)\rangle - \langle n(T)\rangle^2}{\langle n(T)\rangle}, \tag{6}$$

where $\langle n^{(2)}(T)\rangle$ is the second factorial moment and $\langle n(T)\rangle$ is the mean number of photocounts per sample time T. By using (3) and (4), Mandel's Q parameter is expressed as

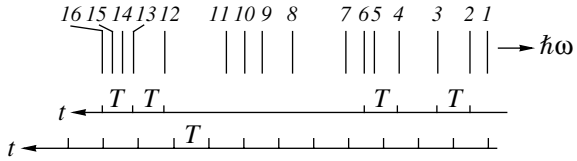

Fig. 1. Sequence of photons emitted by a single atom interacting with a continuous wave laser field.

$$Q(T) = \langle I\rangle\left[\frac{2}{T}\int_0^T dx\int_0^x dy g^{(2)}(y) - T\right]. \tag{7}$$

The autocorrelation function proved very convenient for estimating deviations of photon-counting statistics from the Poisson distribution for various systems [9–11]. For the Poisson distribution, Mandel's Q parameter is zero. For sub- and super-Poissonian (narrower than Poisson) distributions, its values are negative and positive, respectively. Formula (7) is generally used to estimate the deviations from the Poisson photon-counting distribution in single-molecule spectroscopy.

The approach used in this paper to analyze the statistics of fluorescence photons emitted by a single atom is essentially different from that applied in [1–5]. Whereas the analyses presented in these books were focused on the radiation field, the present study deals with the quantum dynamics of a single atom interacting with a continuous wave laser field, and fluorescence dynamics are derived from the dynamics of radiative de-excitation to the ground state of the atom. In this approach, the probability $N(n, T)$ can be calculated directly by using a formula derived below. In what follows, this calculation is presented and the resulting probability is used to calculate the first two moments of the distribution and Mandel's Q parameter for a two-level atom interacting with a continuous wave laser field.

2. FLUORESCENCE PHOTON COUNTING METHODS

Suppose that a two-level atom interacting with a continuous wave laser field emits fluorescence photons at random instants. Figure 1 shows numbered fluorescence photons emitted by the atom. Mandel's formula (2) corresponds to the photon counting method in which the time scale is divided into equal sample intervals (lower scale in Fig. 1) and photons are counted within each interval. The start of the sample interval is neither specified by any definition nor related to any

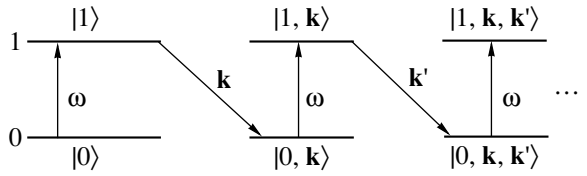

Fig. 2. Transitions between quantum states of the atom–field system.

instant of photon emission. There may exist sample intervals when no photons are counted, as illustrated by Fig. 1.

However, a different photon counting method can be used in single-molecule spectroscopy. It is advantageous in that the probability distribution, in contrast to (2), can be not only derived theoretically, but also calculated. Let us consider this method in some detail.

Suppose that a fluorescence photon is emitted at some instant and another photon is emitted after a time T has elapsed. The former and latter instants define the start and end of the sample interval, respectively. Thus, only intervals of length T containing at least two photocounts are considered, as indicated on the upper time scale in Fig. 1. Two cases are possible.

In one case, no photons are counted within the sample time T between the start and the end, as between photons *2* and *3* or *12* and *13* in Fig. 1. When only such photon pairs are counted, the measurement is actually performed in a start–stop mode (the two-photon start–stop correlator is measured). This measurement mode was called detection of successively emitted photons in [12].

However, a number of photons (called intermediate here) may be emitted during a sample interval of length T between two photons, as photon *5* between photons *4* and *6* or photons *14* and *15* between photons *13* and *16*. When all photon pairs separated by a time T are counted irrespective of the number of intermediate photons emitted within this time, the full two-photon correlator is measured. In the next section, a mathematical expression is found for the probability of counting any number of intermediate photons within a sample time T.

3. QUANTUM DYNAMICS OF A TWO-LEVEL ATOM

Figure 2 schematizes the transitions between the quantum states of the atom–field system: each absorption of a photon of frequency ω is followed by the emission of a fluorescence photon **k**, and the process repeats indefinitely. In detailed studies of the dynamics of the system schematized in Fig. 2 by several independent groups [13–17], the following system of equations was derived for the elements $\rho_{ij}^{(N)}$ of the density matrix describing the state of the atom–field system in the case of N intermediate fluorescence photons:

$$\begin{aligned}
\dot\rho_{10}^{(N)} &= -i(\Delta - i/2T_1)\rho_{10}^{(N)} + \chi[\rho_{11}^{(N)} - \rho_{00}^{(N)}],\\
\dot\rho_{01}^{(N)} &= (\dot\rho_{10}^{(N)})^*,\\
\dot\rho_{11}^{(N)} &= -\chi[\rho_{10}^{(N)} + \rho_{01}^{(N)}] - \frac{1}{T_1}\rho_{11}^{(N)},\\
\dot\rho_{00}^{(N)} &= \chi[\rho_{10}^{(N)} + \rho_{01}^{(N)}] + \frac{1}{T_1}\rho_{11}^{(N-1)}.
\end{aligned} \tag{8}$$

Here, $\chi = \mathbf{E}\cdot\mathbf{d}/\hbar$ is the Rabi frequency, Δ is the difference between the laser and atomic-transition frequencies, and T_1 is the energy relaxation time. This system resembles, but differs from, the optical Bloch equations: it is not closed, since the fourth equation contains the function $\rho_{11}^{(N-1)}$. System (8) describes the first pair of states in Fig. 2 when $N = 0$ (without intermediate photons), the second pair when $N = 1$, and so on.

It can readily be shown that the four matrix elements

$$\rho_{ij} = \sum_{N=0}^{\infty}\rho_{ij}^{(N)} \tag{9}$$

satisfy the Bloch equations with the phase relaxation rate $1/T_2 = 1/2T_1$:

$$\begin{aligned}
\dot\rho_{10} &= -i\left(\Delta - \frac{i}{2T_1}\right)\rho_{10} + \chi(\rho_{11} - \rho_{00}),\\
\dot\rho_{01} &= i\left(\Delta + \frac{i}{2T_1}\right)\rho_{01} + \chi(\rho_{11} - \rho_{00}),\\
\dot\rho_{11} &= -\chi(\rho_{10} + \rho_{01}) - \frac{\rho_{11}}{T_1},\\
\dot\rho_{00} &= \chi(\rho_{10} + \rho_{01}) + \frac{\rho_{11}}{T_1}.
\end{aligned} \tag{10}$$

The analysis of conversion of laser photons into fluorescence photons developed in [14, 15] has shown that the first two equations in system (10) written for a two-level atom interacting with phonons contain a constant $1/T_2$ that is greater than $1/2T_1$ and is determined by electron–phonon interaction. In other words, the Bloch equations contain independent energy and phase relaxation times T_1 and T_2.

Consider Eqs. (8) with $N = 0$. By using the fact that $\rho_{11}^{(-1)} = 0$, denoting $\rho_{ij}^{(0)} = W_{ij}$, and performing the Laplace transform

$$(G(t))_\omega \equiv G(\omega) = \int_0^\infty G(t)e^{i(\omega + i0)t}dt,$$
$$G(t) = \int_{-\infty}^{\infty} G(\omega)e^{-it\omega}\frac{d\omega}{2\pi} = \begin{cases} 0, & t < 0, \\ G(t), & t > 0 \end{cases} \quad (11)$$

of system (8), the following system of equations is obtained for the Laplace components $W(\omega)$:

$$\left(i\omega - \frac{1}{T_1} - k\right)W_{11} + kW_{00} = 0,$$
$$kW_{11} + (i\omega - k)W_{00} = -1, \quad (12)$$

where

$$k(\omega) = i2\chi^2\frac{\omega + i\Gamma}{(\omega + i\Gamma)^2 - \Delta^2} \quad (13)$$

with $\Gamma = 1/T_2 = 1/2T_1$. By virtue of (11), the Bloch equations (10) can be rewritten as equations for the Laplace components $\rho(\omega)$:

$$\left(i\omega - \frac{1}{T_1} - k\right)\rho_{11} + k\rho_{00} = 0,$$
$$\left(\frac{1}{T_1} + k\right)\rho_{11} + (i\omega - k)\rho_{00} = -1. \quad (14)$$

Equations (12) and (14) are obtained under the initial condition

$$W_{00}(t = 0) = \rho_{00}(t = 0) = 1.$$

Systems (12) and (14) yield

$$\rho_{11}(\omega) = \frac{k}{D_p}, \quad W_{11}(\omega) = \frac{k}{D_s}, \quad (15)$$

with the respective determinants

$$D_p = i\omega\left(i\omega - \frac{1}{T_1} - 2k\right), \quad D_s = D_p + \frac{k}{T_1}. \quad (16)$$

Hence, the Laplace components are related as follows:

$$\rho_{11}(\omega) = W_{11}(\omega) + \frac{W_{11}(\omega)}{T_1}\rho_{11}(\omega). \quad (17)$$

By virtue of the equality

$$\int_{-\infty}^{\infty} W(\omega)\rho(\omega)e^{-it\omega}\frac{d\omega}{2\pi} = \int_0^t W(t-x)\rho(x)dx, \quad (18)$$

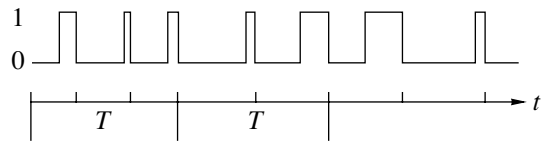

Fig. 3. Dynamics of transitions between the ground and excited electronic states.

the inverse Laplace transform of (17) yields the equation

$$\rho_{11}(t) = W_{11}(t) + \int_0^t \frac{dx}{T_1}W_{11}(t-x)\rho_{11}(x)dx. \quad (19)$$

Iterative solution of Eq. (19) leads to a series expansion of the probability $\rho_{11}(t)$ in terms of the number of intermediate photons:

$$\rho_{11}(t) = W_{11}(t) + \sum_{N=1}^{\infty}\rho_{11}^{(N)}(t), \quad (20)$$

where

$$\rho_{11}^{(N)}(t) = \int_0^t \frac{dt_1}{T_1}\int_0^{t_1}\frac{dt_2}{T_1}\cdots\int_0^{t_{N-1}}\frac{dt_N}{T_1}W_{11}(t-t_1)$$
$$\times W_{11}(t_1 - t_2)\ldots W_{11}(t_{N-1} - t_N)W_{11}(t_N). \quad (21)$$

A simple physical interpretation of this formula is derived from Fig. 3 by noting that 1 to 0 and 0 to 1 transitions correspond to fluorescence-photon emissions and laser-photon absorptions, respectively. Since photons are absorbed and emitted at random instants, Fig. 3 illustrates a random sequence of events in the system schematized in Fig. 2. The lengths of the lower and upper segments represent the lifetimes of ground and excited states of the atom, respectively. The fluorescence photons emitted by transition from upper to lower segments are shown by vertical segments on the time scale. Figure 3 resembles a random telegraph signal, and the number of photons emitted during any particular time interval equals the number of de-excitations that occurred within this interval. Two intervals of length T containing one and two intermediate photons, respectively, are indicated in Fig. 3. The lower and upper line segments are associated, respectively, with the probabilities

$$\rho_{00}^{(0)}(t), \rho_{00}^{(1)}(t), \rho_{00}^{(2)}(t), \ldots$$

and

$$\rho_{11}^{(0)}(t) = W_{11}(t), \rho_{11}^{(1)}(t), \rho_{11}^{(2)}(t), \ldots.$$

The equation obtained by dividing (20) by $\rho_{11}(t)$,

$$w_0(t) + \sum_{N=1}^{\infty} w_N(t) = 1, \tag{22}$$

relates the probability

$$w_0(t) = \frac{W_{11}(t)}{\rho_{11}(t)} \tag{23}$$

of an interval of length t without intermediate photocounts to the probability

$$\begin{aligned} w_N(t) &= \frac{1}{\rho_{11}(t)} \\ &\times \int_0^t \frac{dt_1}{T_1} \int_0^{t_1} \frac{dt_2}{T_1} \cdots \int_0^{t_{N-1}} \frac{dt_N}{T_1} W_{11}(t-t_1) \\ &\times W_{11}(t_1-t_2)\ldots W_{11}(t_{N-1}-t_N) W_{11}(t_N) \end{aligned} \tag{24}$$

of an interval t containing N intermediate photocounts. Formulas (23) and (24) can be used to calculate the distribution of intermediate photons counted within an interval of length t between photon emissions.

4. DISTRIBUTION OF PHOTONS EMITTED BY A TWO-LEVEL ATOM

According to (23) and (24), the calculation of the desired probabilities requires expressions for $W_{11}(t)$ and $\rho_{11}(t)$ satisfying Eqs. (12) and (14). These are not balance equations, because the coefficient given by (13) depends on the frequency ω. These equations are strictly equivalent to Eqs. (8) with $N = 1$ and Bloch equations (10); i.e., they preserve the phase correlations represented by the off-diagonal elements of the density matrix. Coherence is lost when $\omega = 0$ in Eq. (13), in which case

$$k = 2\chi^2 \frac{\Gamma}{\Gamma^2 + \Delta^2}. \tag{25}$$

Expression (25) describes the absorption coefficient, and Eqs. (11) and (12) with this value of k are balance equations.

One is led to a question: is it correct to change from the Bloch equations to balance equations? The answer is provided by Fig. 4. Figure 4 compares the time dependence of solutions to the Bloch and balance equations for several laser intensities. Rabi oscillation occurs when the intensity is so high that the stimulated transition rate exceeds the spontaneous transition rate. Figure 4b shows that the difference between the solutions to the Bloch and balance equations is insignificant when the field intensity is low. Accordingly, use of the balance equations facilitates the analysis of fluorescence photons statistics presented below.

Solution of Eqs. (12) and (14) with the absorption coefficient given by (25) leads to the following expressions:

$$W_{11}(t) = \frac{k}{A-a}[\exp(-at) - \exp(-At)], \tag{26}$$

$$\rho_{11}(t) = \frac{k}{A+a}\{1 - \exp[-(A+a)t]\}, \tag{27}$$

where

$$\begin{aligned} a &= \gamma - R, \quad A = \gamma + R, \\ \gamma &= k + 1/2T_1, \quad R = \sqrt{k^2 + (1/2T_1)^2}. \end{aligned} \tag{28}$$

If the pump intensity is low ($k \ll 1/T_1$), then $a \approx k$ and $A \approx 1/T_1$. It can be verified by direct substitution that functions (26) and (27) satisfy Eq. (19) with $1/T_1$ replaced by A.

According to (26), the probability $W_{11}(t)$ increases from zero within a time interval on the order of T_1 and then decreases as

$$W_{11}(t) \propto \exp(-kt). \tag{29}$$

Probability (24) calculated by using (23) and approximating $W_{11}(t)$ with (29) is the Poisson distribution

$$w_N(t) = \frac{(kt)^N}{N!}\exp(-kt). \tag{30}$$

Thus, the deviation of fluorescence photon statistics from the Poisson distribution is due to the increase in $W_{11}(t)$ over times on the order of T_1. The vanishing probability of photon emission for $t \longrightarrow 0$ is due to the phenomenon known as *photon antibunching*. Now, let us find the photon-counting distribution taking into account photon antibunching.

The Laplace transform of (24) combined with formula (18) yields

$$(\rho_{11}(t)w_N(t))_\omega = \left(\frac{W_{11}(\omega)}{T_1}\right)^N W_{11}(\omega). \tag{31}$$

The Laplace component of the probability given by (26) is

$$W_{11}(\omega) = \frac{k}{(i\omega - a)(i\omega - A)}. \tag{32}$$

By using (28), it can readily be shown that

$$aA = k/T_1.$$

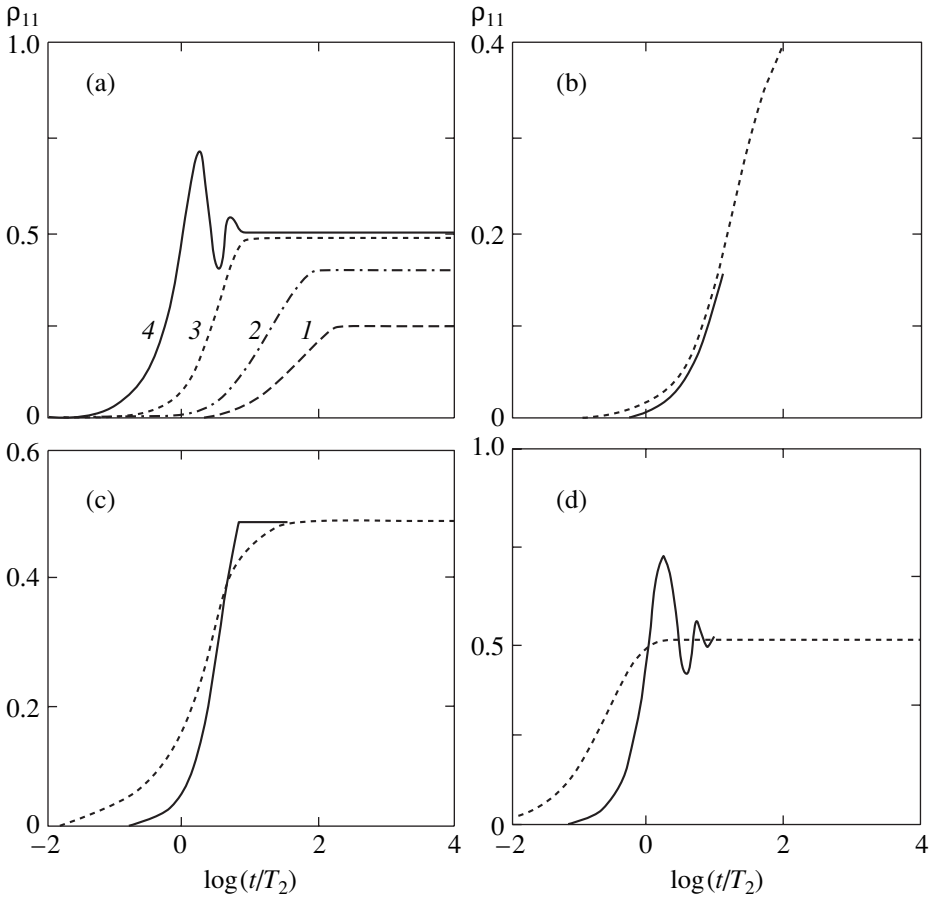

Fig. 4. (a) Time dependence of probability $\rho_{11}(t)$ predicted for several laser intensities: $\chi T_2 = 0.05$ (*1*), 0.1 (*2*), 0.3 (*3*), and 1.0 (*4*). Solutions to the Bloch equations (solid curves) are compared to solutions to balance equations (dashed curves) for $\chi T_2 = 0.1$ (b), 0.3 (c), and 1.0 (d).

Combined with (32), this relation yields

$$\frac{W_{11}(\omega)}{T_1} = \frac{aA}{(i\omega - a)(i\omega - A)}. \quad (33)$$

By virtue of the last two formulas, relation (31) can be rewritten as

$$(\rho_{11}(t)w_N(t))_\omega = k\frac{a^N}{(i\omega - a)^{N+1}}\frac{A^N}{(i\omega - A)^{N+1}}. \quad (34)$$

Here, the first fraction on the right-hand side is the Laplace component of the function

$$\frac{(at)^N}{N!}\exp(-at),$$

i.e., the Poisson function. By taking this into account, using formula (18), and performing the inverse Laplace transform, the desired probability of counting N photons within an interval of length t is found to be proportional to the convolution of Poisson distributions:

$$w_N(t) = \frac{k}{\rho_{11}(t)}\int_0^t dx P_N[(t-x)a]P_N[xA] = \frac{k}{\rho_{11}(t)}\int_0^t dx \frac{[(t-x)a]^N}{N!}e^{-(t-x)a}\frac{[xA]^N}{N!}e^{-xA}. \quad (35)$$

This formula is equivalent to expression (24) for the probability of an interval t containing N intermediate photocounts, but is much more practical, because it contains a single integral.

Figure 5 shows photon-counting distributions predicted by this expression and illustrates the simple fact that many photons can be counted only on time intervals much longer than the mean time $1/k$ between successive emissions of fluorescence photons. The maximum of the distribution corresponds to the number N of photocounts related to the sample time by the approximate formula $kT \approx N$. A similar relation between sam-

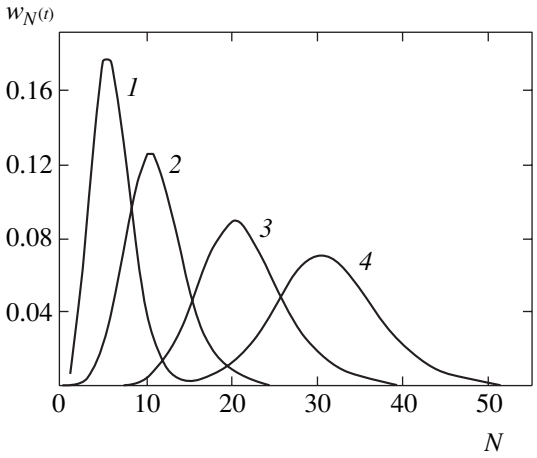

Fig. 5. Photon-counting distribution for sample times $T = 5/k$ (*1*), $10/k$ (*2*), $20/k$ (3), and $30/k$ (*4*).

ple time, number of photocounts, and laser intensity is characteristic of the Poisson distribution. The longer the sample interval, the more photons can be counted. A number of photocounts that is sufficiently large to provide a representative sample corresponds to a sample time much longer than the mean time $1/k$ between successive emissions of fluorescence photons. The resulting distribution resembles, but differs from, the Poisson distribution.

Figure 6 compares these distributions and illustrates the distinction between them. It demonstrates that the width of the Poisson distribution corresponding to the sample time $T = 27/k$ is equal to the width of distribution (35) calculated for the larger sample time $T = 30/k$. When the distributions correspond to equal sample times, distribution (35) is narrower than the Poisson distribution; i.e., a sub-Poissonian distribution is obtained. A sub-Poissonian distribution of fluorescence photons emitted by a single atom interacting with continuous radiation within a short sample interval of length $T < 1/k$ was measured in [18].

Figure 7 shows the distribution predicted by (35) and the Poisson distribution for $T < 1/k$. Such short sample intervals without photocounts occur with a probability close to unity, in agreement with results reported in [18]. The probability of detecting a single photon within such intervals is less than one percent, and the probability of three photocounts within such an interval is less than one-hundred-thousandth of a percent. Even though the deviation from the Poisson distribution reported by Short and Mandel was small, it should be responsible for the negative value of Mandel's Q parameter, which can be measured in experiment.

5. MEAN NUMBER OF PHOTOCOUNTS AND MANDEL'S Q PARAMETER

The distribution of photons emitted by a single atom interacting with a continuous wave laser field has never been calculated by using formula (2). The present theory can be compared with Mandel's theory by using moments of the distribution and Mandel's Q parameter derived from the distribution found here. The mean number of photons is calculated as

$$\langle n(T)\rangle = \sum_{N=0}^{\infty} N w_N(T), \tag{36}$$

with probabilities defined by (24). Substituting the expressions for probability (24), performing the

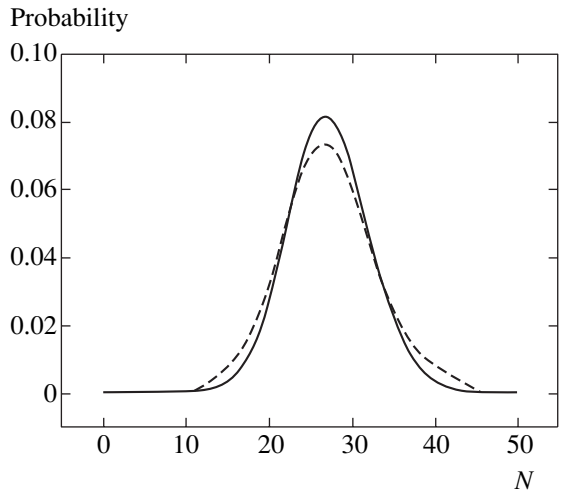

Fig. 6. Photon-counting distribution predicted by (35) for $T = 30/k$ (solid curve) and Poisson distribution of equal width corresponding to $T = 27/k$ (dashed curve).

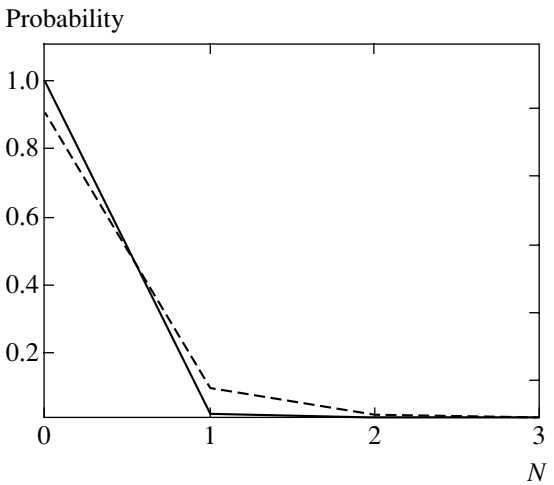

Fig. 7. Sub-Poissonian photon-counting distribution predicted by formula (35) (solid curve) and Poisson distribution of photocounts (dashed curve) for $T = 1/10k$.

Laplace transform, and using formula (17), one obtains

$$[\langle n(T)\rangle\rho_{11}(T)]_\omega = W_{11}(\omega)\sum_{N=1}^{\infty} N\left(\frac{W_{11}(\omega)}{T_1}\right)^N = \frac{W_{11}^2(\omega)/T_1}{(1-W_{11}(\omega)/T_1)^2} = \frac{\rho_{11}^2(\omega)}{T_1}. \quad (37)$$

The inverse Laplace transform yields

$$\langle n(T)\rangle = \frac{1}{\rho_{11}(T)}\int_0^T \frac{dx}{T_1}\rho_{11}(T-x)\rho_{11}(x). \quad (38)$$

This expression differs from formula (3) derived by Mandel. Formula (38) takes into account antibunching of fluorescence photons, whereas formula

$$\langle n\rangle = \langle I\rangle T$$

does not.

The second factorial moment

$$\langle n^{(2)}(T)\rangle = \sum_{N=0}^{\infty} N(N-1)w_N(T) \quad (39)$$

contained in the expression for Mandel's Q parameter can also be calculated by using the Laplace transform, as in the calculation of the mean photon number. The resulting expression

$$\langle n^{(2)}(T)\rangle = \frac{2}{\rho_{11}(T)}\int_0^T \frac{dx}{T_1}\rho_{11}(T-x) \times \int_0^x \frac{dy}{T_1}\rho_{11}(x-y)\rho_{11}(y) \quad (40)$$

also differs from the formula (4) found by Mandel.

The mean number of photocounts per sample time T is obtained by substituting probability (27) into (38):

$$\langle n(T)\rangle = \frac{k}{2\gamma T_1}\left[T\frac{1+e^{-2\gamma T}}{1-e^{-2\gamma T}} - \frac{1}{\gamma}\right]. \quad (41)$$

Since $2\gamma T_1 \approx 1$ for low laser intensities, this result simplifies to

$$\langle n(T)\rangle = \begin{cases} \dfrac{\gamma T}{3}kT, & \gamma T \ll 1, \\ kT - \dfrac{k}{\gamma}, & \gamma T \gg 1. \end{cases} \quad (42)$$

These formulas are different from the expression

$$\langle n\rangle = \langle I\rangle T = kT \quad (43)$$

used by Mandel in (7). Owing to photon antibunching, the number of photons counted within a short sample interval is a quadratic, rather than linear, function of sample time. Indeed, formula (43) is predicted by the present theory when photon antibunching is neglected, which corresponds to the Poisson distribution. Therefore, it is not quite correct to use formula (43), which does not allow for photon antibunching, in expression (7) for calculating the parameter $Q(T)$ as a measure of deviation of a real photon-counting distribution from the Poisson distribution.

The second factorial moment is obtained by substituting (27) into (40):

$$\langle n^{(2)}(T)\rangle = \left(\frac{k}{2\gamma T_1}\right)^2 \times \left[-\left(\frac{3T}{\gamma}\right)\frac{1+e^{-2\gamma T}}{1-e^{-2\gamma T}} + T^2 + \frac{3}{\gamma^2}\right]. \quad (44)$$

Formulas (41) and (44) are combined to obtain the following expression for Mandel's Q parameter:

$$Q(T) = \frac{1}{2\gamma T_1}\frac{k}{\gamma} \times \left\{\frac{-3\gamma T(1+e^{-2\gamma T}) + [(\gamma T)^2+3](1-e^{-2\gamma T})}{\gamma T(1+e^{-2\gamma T}) - (1-e^{-2\gamma T})} - \frac{\gamma T(1+e^{-2\gamma T})}{1-e^{-2\gamma T}} + 1\right\}. \quad (45)$$

Let us compare this result with that predicted by Mandel's formula (7). It was shown in [9, 15, 19] that the autocorrelation function is related to the probability ρ_{11} as follows:

$$g^{(2)}(t) = \frac{\rho_{11}(t)}{\rho_{11}(\infty)}.$$

Combining this formula with (27) and (7), one obtains

$$Q_M(T) = \frac{k}{2\gamma}\left[-1 + \frac{1-e^{-2\gamma T}}{2\gamma T}\right]. \quad (46)$$

According to (45) and (46), Mandel's Q parameter is proportional to $k/2\gamma = kT_1$, i.e., small when the laser intensity is low. However, the time behavior of Mandel's Q parameter is independent of the laser intensity, being determined by the dimensionless time γT. If $\gamma T \gg 1$, then $Q = 4Q_M$. The difference in the theoretical

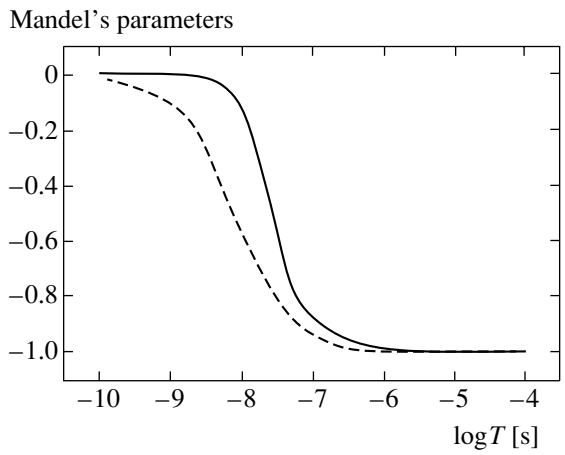

Fig. 8. Mandel's Q parameters calculated by using formulas (45) (solid curve) and (46) (dashed curve) for $T_1 = 10^{-8}$ s.

expressions for Mandel's Q parameter should be explained by difference in the methods actually used to measure the moments of the distribution. Figure 8 compares the functions $Q(T)2\gamma T_1\gamma/k$ and $Q_M(T)2\gamma/k$. Mandel's Q parameter predicted by (45) varies from zero to the limit value $Q(\infty) = -4kT_1$ exponentially. Note that the transition between the domains of short- and long-time behavior corresponds to a change in T within an order of magnitude in the neighborhood of $T \approx T_1$. Thus, expression (45) implies that a fluorescence photon "remembers" the preceding one during a coherence time approximately equal to T_1. Mandel's Q parameter predicted by (46) varies from zero to the limit value $Q_M(\infty) = -kT_1$ according to a hyperbolic law, i.e., at a slower rate. Formula (45) predicts an appreciable deviation from the Poisson distribution only for $T > T_1$, whereas formula (46) predicts that it must be observed for $T \le T_1$. The difference in behavior of Mandel's Q parameter given by (45) and (46) can be explained by the fact that these formulas correspond to different fluorescence photon counting methods.

6. CONCLUSIONS

1. When only intermediate photons are counted in an experiment, formula (24) yields the desired expression for the statistics of fluorescence photons emitted by a single two-level atom interacting with a continuous wave laser field.

2. If the stimulated transition rate k is lower than the spontaneous transition rate $1/T_1$ t by at least an order of magnitude, then formula (24) reduces to the substantially simpler form of (35).

3. The photon-counting distributions found numerically by using formula (35) for several sample times are sub-Poissonian, in agreement with the experimental results obtained in [18]. The deviation from the Poisson distribution is more pronounced for short sample times, when the number of photocounts is small (cf. Figs. 6 and 7).

4. The expressions for the mean photon number and second factorial moment of the photon-counting distribution corresponding to the photon counting method proposed in the present study are different from those obtained by Mandel, and the expressions for Mandel's Q parameter differ accordingly. Whereas both time dependence and limit values of the parameter are quantitatively different, both formulas compared here predict similar dependence on sample time.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 04-02-17024.

REFERENCES

1. R. I. Glauber, in *Quantum Optics and Electronics*, Ed. by C. de Witt, A. Blandin, and C. Cohen-Tannoudji (Gordon and Breach, New York, 1965; Mir, Moscow, 1966).
2. J. R. Klauder and E. C. G. Sudarshan, *Fundamentals of Quantum Optics* (Benjamin, New York, 1968; Mir, Moscow, 1970).
3. M. Lax, *Fluctuations and Coherence in Classical and Quantum Physics* (Gordon and Breach, New York, 1968; Mir, Moscow, 1974).
4. R. Loudon, *The Quantum Theory of Light* (Claredon, Oxford, 1973; Mir, Moscow, 1976).
5. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge Univ. Press, Cambridge, 1995; Fizmatlit, Moscow, 2000).
6. L. Mandel, Opt. Lett. **4**, 205 (1979).
7. L. Mandel, Proc. Phys. Soc. London **72**, 1037 (1958).
8. P. L. Kelley and W. H. Kleiner, Phys. Rev. **136**, 316 (1964).
9. M. S. Kim and P. L. Knight, Phys. Rev. A **36**, 5265 (1987).
10. E. Barkai, Y. J. Jung, and R. Silbey, Phys. Rev. Lett. **87**, 207403 (2001).
11. Y. J. Jung, E. Barkai, and R. Silbey, Adv. Chem. Phys. **123**, 199 (2002).
12. M. S. Kim, P. L. Knight, and K. Wodkiewicz, Opt. Commun. **62**, 385 (1987).
13. B. R. Mollow, Phys. Rev. A **12**, 1919 (1975).
14. I. S. Osad'ko, Zh. Éksp. Teor. Fiz. **113**, 1606 (1998) [JETP **86**, 875 (1998)].
15. I. S. Osad'ko, *Selective Spectroscopy of Single Molecules* (Fizmatlit, Moscow, 2000; Springer, Berlin, 2002).
16. Y. Zheng and F. L. H. Brown, Phys. Rev. Lett. **90**, 238305 (2003).
17. Y. Zheng and F. L. H. Brown, J. Chem. Phys. **119**, 11814 (2003).
18. R. Short and L. Mandel, Phys. Rev. Lett. **51**, 384 (1983).
19. I. S. Osad'ko and L. B. Yershova, J. Lumin. **87–89**, 184 (2000).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 73–79.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 88–94.

ATOMS, MOLECULES, OPTICS

Bremsstrahlung of Fast Charged Particles on Clusters in a Wide Spectral Range

V. A. Astapenko

Moscow Institute of Physics and Technology, Dolgoprudnyĭ, Moscow oblast, 141700 Russia
e-mail: astval@mail.ru

Received February 17, 2005

Abstract—Within the framework of the first Born approximation and a simple model of the structural factor, the bremsstrahlung of fast charged particles on polyatomic clusters is calculated and analyzed with regard to the polarization mechanism in a wide spectral range including a domain of high frequencies. The role of cooperative phenomena in the static and polarization channels of bremsstrahlung is investigated. It is established that these phenomena, being negligible for static bremsstrahlung, substantially influence the polarization bremsstrahlung. It is shown that the constructive interference between the contributions of the atoms of a cluster to the polarization bremsstrahlung substantially increases its intensity and changes its dependence on the basic parameters of the problem compared with the case of bremsstrahlung on an isolated atom. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The scattering of a charged particle by a target that possesses electronic degrees of freedom leads to bremsstrahlung by two channels. One of these channels, the traditional (static) bremsstrahlung (SB), is associated with the acceleration of the incident particle in the field of the target. The second, polarization bremsstrahlung (PB), is associated with the alternating dipole moment of the target that arises during the scattering of the incident particle [1, 2].

In general, the bremsstrahlung by these channels differently depends on the frequency and angle of radiation and on the mass and energy of the incident particle. If the radiation occurs in a medium (plasma, amorphous substance, monocrystal, or polycrystal), the static and polarization mechanisms of bremsstrahlung exhibit different dependence on the parameters of the medium. The last circumstance is largely attributed to the fact that the SB is formed at small distances from the target, whereas the PB, at large distances. As a result, the cooperative phenomena associated with the interference between the contributions of different atoms of the target to bremsstrahlung play different roles in the static and polarization mechanisms. The cooperative phenomena discussed here should manifest themselves in the purest form in the PB on polyatomic clusters, which is the subject of the present paper.

Radiation of a charged particle scattered by a cluster was considered earlier for certain ranges of parameters. For example, the bremsstrahlung of a fast but nonrelativistic electron on a cluster in the low-frequency range ($\omega < 1$ au) with regard to the polarization mechanism was investigated in [3] (see also [1, Ch. 7]). The calculations were performed within the first Born approximation with respect to the interaction between electron and target. The cluster form factor was calculated with the use of a jellium model. The bremsstrahlung cross section was studied as a function of frequency, velocity (of electron), and the radiation angle. It was shown that, in the frequency domain where the main contribution to PB is made by collective excitations of cluster electrons (the domain of a giant resonance), the polarization channel dominates the static one and the interchannel interference leads to asymmetry of the radiation spectrum.

In the recent work [4], Kurkina calculated the characteristics of the SB for scattering of a slow electron by metallic clusters. These calculations were based on the expansion of the electron wavefunction in terms of partial waves. A model of spherical jellium was used for the electrostatic potential of a cluster. The author demonstrated that, as the number of atoms in a cluster increases and the energy of radiating electron decreases, a resonance structure arises in the high-frequency region of the SB spectrum. This structure consists of a series of peaks whose positions are determined by the type of the cluster and the energy of the electron after emitting a photon. The origin of the spectral peaks was associated with quasistationary states of the electron in the field of the cluster–target. The polarization channel of bremsstrahlung was not considered in [4].

As the energy of the incident particle increases, the role of polyatomic interference phenomena in PB on clusters must increase because the magnitude of the minimal momentum transferred to the target decreases and the effective projectile–target distance increases. As the energy of the incident particle increases, the spectral domain of the bremsstrahlung allowed by the

energy conservation law increases. In the case of a relativistic incident particle, the frequency range $\omega \gg 1$ au becomes topical, in which one can neglect the collective excitations of electrons in the cluster and assume that the polarizability of atomic clusters is described in terms of the polarizability of an isolated atom to a reasonable degree of accuracy. It is also essential that, in the relativistic case, the role of the recoil during the emission of a photon by the polarization channel increases because the photon momentum becomes comparable to the minimal variation of the momentum of the impinging particle. As a result, the range of angles of PB of a relativistic electron narrows down in the high-frequency range [5, 6]. For the scattering of a relativistic charged particle by a cluster, the angular–frequency dependence of PB should be additionally modified when the number of atoms in the target is varied.

The present paper is devoted to the theoretical analysis of bremsstrahlung of fast, including relativistic, charged particles on clusters in a wide spectral range of $\omega \gg 1$ au. The main objective of the paper is to determine and analyze the basic principles of bremsstrahlung that are associated with the interference between the contributions of various atoms of the target to the process.

2. BASIC RELATIONS

Let us calculate the intensity of the polarization and static channels of bremsstrahlung of a fast charged particle on a cluster within a simple model. The main assumptions of this approach are the first Born approximation with respect to the interaction between the incident particle and the target and a jellium model for the form factor of the cluster. In addition, we will use a quasiclassical formula for the SB amplitude and an approximate expression for the generalized polarizability of the cluster atoms. Applying a standard quantum-mechanical procedure (see [7] for details) for the differential intensity of bremsstrahlung by each channel, normalized by the number N of atoms in the cluster, we obtain the following expression:

$$\frac{dI}{d\omega\Omega_{\mathbf{n}}} = \frac{1}{N}\int_{q_{\min}}^{q_{\max}} T(q)dq, \tag{1}$$

where $\mathbf{q} = \mathbf{p}_f - \mathbf{p}_i + \mathbf{k}$ is the momentum transferred to the target from the incident particle; $T(q)$ is a partial intensity of bremsstrahlung; $d\Omega_{\mathbf{n}}$ is the solid angle in the direction of radiation; ω and $\mathbf{k}$ are the frequency and the wavevector of the photon, respectively; $\mathbf{p}_i$ and $\mathbf{p}_f$ are the initial and final momenta of the incident particle; $q_{\max} = 2\mu v$, where μ is the reduced mass of the electron and of the incident particle and v is the velocity of the incident particle; and $q_{\min}$ is defined by formula (13) below. Throughout the paper, we use the atomic system of units $\hbar = e = m_e = 1$.

Within the approach used in this paper, the partial intensity of PB can be represented as

$$T_{\text{pol}}(q) = \frac{2Z_p^2}{\pi c^3 v q} S(q, N) \times \left|Z_{\text{pol}}(\omega, q)^2\right| I_\phi(q, v, \omega, \theta), \tag{2}$$

where Z_p is the charge of the incident particle, c is the velocity of light, $S(q, N)$ is the structure factor of a cluster, $Z_{\text{pol}}(\omega, q)$ is the effective polarization charge of the atoms of the cluster, $I_\phi(q, v, \omega, \theta)$ is the kinematic integral that results from the integration with respect to the azimuthal angle $\mathbf{q}$, and θ is the angle between the vectors $\mathbf{p}_i$ and $\mathbf{k}$ (the radiation angle of photons). The kinematic integral $I_\phi(q, v, \omega, \theta)$ represents a rather complicated function of the parameters of the problem; the explicit form of this function is given in [7].

Note that expression (2) is obtained for a sufficiently high-frequency range in which $\omega \gg I_a$, where I_a is the ionization potential of the atoms that constitute a cluster. The opposite case of low frequencies ($\omega < I_a$) is considered in [3].

We will use the following model representation for the structure factor of the cluster:

$$S(q, N) = N^2 F_J^2(q, N) + N[1 - F_J^2(q, N)], \tag{3}$$

where

$$F_J(q, N) = 3\frac{j_1(qr(N))}{qr(N)} \tag{4}$$

is the form factor of a spherical cluster, normalized by the number of atoms, in the jellium model;

$$j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x} \tag{5}$$

is the spherical Bessel function of order 1; and $r(N)$ is the cluster radius, which depends on the number of atoms and can be calculated by the formula

$$r(N) = r_{\text{WS}}\sqrt[3]{N} = \sqrt[3]{\frac{3N}{4\pi n_a}}, \tag{6}$$

where r_{WS} is the Wigner–Seitz radius and n_a is the solid-state concentration of atoms in the cluster.

The first term on the right-hand side of equality (3) represents a coherent part of the structure factor of the cluster, and the second term, a noncoherent part. Note that form factor (4) is the spatial Fourier image, normal-

ized by the number of atoms, of the probability of the distribution of atoms in the cluster in the jellium model:

$$w_J(r, N) = \frac{3N\Theta(r(N)-r)}{4\pi r^3(N)}, \quad (7)$$

where $\Theta(x)$ is the Heaviside step function. In the case of a monoatomic cluster, structure factor (4) is equal to unity.

The polarization charge of the cluster atoms can be represented as

$$Z_{pol}(\omega, q) = \omega^2|\alpha(\omega, q)| \approx \omega^2|\alpha(\omega)|F_a(q), \quad (8)$$

where $\alpha(\omega)$ and $F_a(q)$ are the dipole polarizability and the form factor of an atom, respectively. These quantities are calculated by the method proposed in [7]. The imaginary part of the polarizability is determined by the optical theorem, in terms of the photoabsorption cross section, using the data provided at the site [8]. The real part of the polarizability is recovered by the Kramers–Kronig relation. The atomic form factor is calculated in the Slater approximation by the formula obtained in [9].

An approximate expression for the partial intensity of SB in the quasiclassical approximation $\varepsilon_i \gg \omega$ (ε_i is the initial energy of the incident particle) and the relativistic limit ($v \approx c$) has the form

$$T_{st}(q) \approx \frac{2}{3\pi c^3 v q} S(q, N)\left(\frac{Z_p}{m_p}\right)^2 \times Z^2[1-F_a(q)]^2 \frac{[1-(v/c)^2](1+\cos^2\theta)}{[1-(v/c)\cos\theta]^2}, \quad (9)$$

where m_p is the mass of the incident particle and Z is the charge of the atomic nucleus. Note that the relative error of formula (9) for nonrelativistic velocities of the incident particle is no greater than 30%.

The formulas presented above describe the intensity of two channels of bremsstrahlung for sufficiently high radiation frequencies $\omega \gg I_a$ when a fast charged particle is scattered by a cluster. We neglect the interchannel interference term because the PB and SB amplitudes differently depend on the transferred momentum and, in the relativistic case, on the radiation angle as well.

3. DISCUSSION OF THE RESULTS

Let us apply the formulas obtained in the previous section to calculate the intensity of bremsstrahlung by static and polarization channels when a fast electron is scattered by polyatomic clusters.

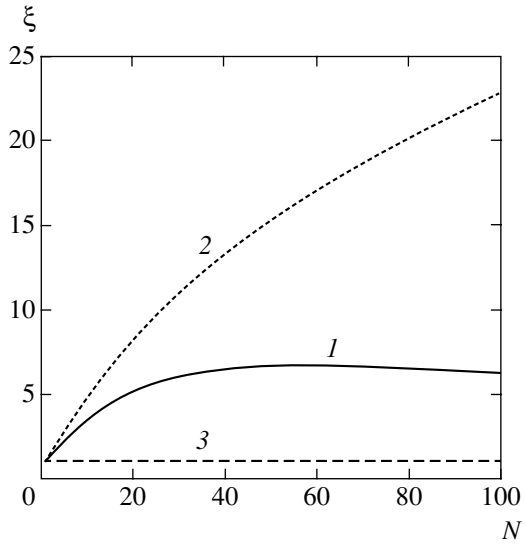

Fig. 1. Cooperative phenomena in PB (*1* and *2*) and SB (*3*) of an electron scattered by a copper cluster ($\gamma = 10$ and $\hbar\omega = 1$ keV); (*1*) $\theta = 1$ rad and (*2*) $\theta = 0.5$ rad.

The cooperative phenomena in bremsstrahlung are characterized by the following relation:

$$\xi = \frac{dI(N)}{dI(N=1)}, \quad (10)$$

where dI is the differential intensity of bremsstrahlung by one of the channels, normalized by the number of atoms in a cluster. It is obvious that $\xi = 1$ in the absence of cooperative phenomena. In the opposite limit case of constructive interference between the contributions of the cluster atoms to the bremsstrahlung intensity, we have

$$\xi = \chi N, \quad \chi < 1.$$

The coefficient χ takes into account the fact that not all the transferred momenta that are essential in the process on an individual atom make a significant contribution to the coherent part of bremsstrahlung on a cluster.

The dependence of the parameter ξ on the number of atoms in a copper cluster for both channels of bremsstrahlung is demonstrated in Fig. 1. The energy of a "bremsstrahlung" photon is equal to 1 keV, and the Lorentz factor is $\gamma = [1-(v/c)^2]^{-1/2} = 10$. In the case of a polarization channel, the function $\xi(N)$is shown for two values of the radiation angle $\theta = 0.5$ and 1 rad. One can see that the cooperative phenomena are negligible for the static channel of bremsstrahlung but are quite significant for the polarization channel. The analysis shows that the effect of the cooperative phenomena in PB appreciably increases as the radiation angle decreases. Moreover, the role of these phenomena increases with the energy of the incident particle and as the frequency of the bremsstrahlung photon decreases.

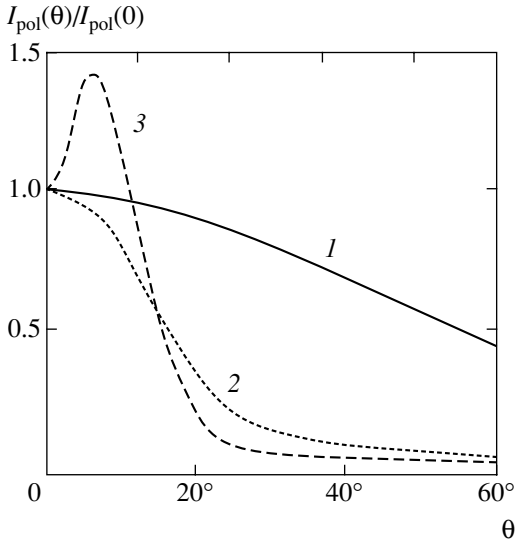

Fig. 2. Angular dependence of the normalized intensity of PB of an electron on an isolated atom (*1*) and on a copper cluster for $N = 100$ (*2*), and $N = 1000$ (*3*); $\hbar\omega = 5$ keV and $\gamma = 10$.

The curves in Fig. 1 and the results of calculations imply that, in the case of the polarization channel, starting from a certain value of N_{sat} that depends on the radiation angle and the energy of the incident particle, the radiation intensity saturates for a certain number of atoms in the cluster. The analysis shows that N_{sat} increases as the radiation angle decreases and the energy of the incident particle increases.

To account for the results listed above, we take into account that, according to formulas (3)–(6) for the structure factor, the constructive interference of the contributions of different atoms of the cluster occurs only for sufficiently small values of the transferred momentum,

$$q < 1/r(N), \tag{11}$$

where $r(N)$ is the cluster radius (see (6)). Otherwise, the structure factor of the cluster (normalized by the number of atoms) is equal to unity, and there are no cooperative phenomena.

It is essential that inequality (11) is inconsistent with the condition

$$q > 1/r_a, \tag{12}$$

which defines the range of transferred momenta in which SB is not small by virtue of the obvious inequality $r(N) > r_a$ (r_a is a characteristic atomic radius). Hence, the cooperative phenomena in SB on a cluster are negligible.

At the same time, a partial amplitude of PB is large when the inequality opposite to (12) holds; thus, there is no similar prohibition on the cooperative phenomena in the polarization channel. Let us write out an expression for the minimal momentum that is transferred to the target from the incident particle and appears in integral (1):

$$q_{min}(\omega, v, \theta) = \frac{\omega}{v}\left(1 - \frac{v}{c}\cos\theta\right). \tag{13}$$

The fact that the cooperative phenomena are essential,

$$q_{min} < 1/r(N), \tag{14}$$

and formula (6) for the cluster radius imply the following expression for the number N_{sat} of atoms in a cluster at which the saturation occurs:

$$N_{sat} = \frac{4\pi n_a v^3}{3\omega^3\left(1 - \frac{v}{c}\cos\theta\right)^3}. \tag{15}$$

Equality (15) shows that N_{sat} strongly depends on the radiation angle and on the energy of the incident particle in the relativistic case. For instance, for the parameters of Fig. 1, we have

$$N_{sat}(\theta = 1 \text{ rad}) = 27 \text{ and } N_{sat}(\theta = 0.5 \text{ rad}) = 1312.$$

The effect of the cooperative phenomena on the angular dependence of the intensity I_{pol} of PB on a cluster is demonstrated by Fig. 2. This figure represents the intensity of PB, normalized by its value at zero angle, as a function of the radiation angle for different numbers of atoms in a copper cluster, including the monoatomic case, for a photon energy of 5 keV and a Lorentz factor of $\gamma = 10$. One can see that the angular distribution of the PB intensity is narrowed down as the number of atoms N increases; moreover, in the limit of large values of this number, this distribution becomes nonmonotonic. This nonmonotonicity disappears in the nonrelativistic case as the radiation frequency decreases and increases with the number of atoms.

Note that these variations of the angular distribution of the intensity of PB on clusters can be observed experimentally only for heavy incident particles. In the case of light particles (an electron or a positron) and small radiation angles, the static channel dominates in which cooperative effects are small.

To describe the relative contribution of PB to the process, we introduce an R factor by the equality

$$R = \frac{dI_{pol}}{dI_{st}}, \tag{16}$$

where I_{st} is the intensity of SB. The angular dependence of the R factor of an electron for $\gamma = 10$, a photon energy of 1 keV, and different numbers of atoms in a copper cluster is represented in Fig. 3. One can see that the role

of PB increases with the number of atoms. For example, in the monoatomic case, the angle at which the intensities of PB and SB become equal amounts to 30°, while, for $N = 100$, this angle is 10°.

The condition under which the cooperative phenomena are essential in the PB spectrum can be obtained from inequality (14) with regard to the explicit expression (13) for the minimal momentum transferred to the target. This condition is expressed as

$$\omega < \omega_{\max} = \sqrt[3]{\frac{4\pi n_a}{3N}} \frac{v}{1 - \frac{v}{c}\cos\theta}. \qquad (17)$$

For frequencies greater than $\omega_{\max}$, cooperative phenomena in bremsstrahlung on a cluster are small.

In the nonrelativistic limit $v \ll c$, the maximal frequency above which the cooperative phenomena disappear does not depend on the radiation angle and is equal to

$$\omega_{\max}^{\text{nrel}} = \frac{v}{r_{WS}\sqrt[3]{N}}. \qquad (18)$$

For metallic clusters, the Wigner–Seitz radius r_{WS} ranges from 2 to 4, so that the characteristic value of frequency (18) for a moderate-size cluster and for the velocity of the incident particle on the order of 10 au amounts to 1 au.

In the frequency range $\omega < 1$ au, the computation model used in this paper becomes inadequate because, in this case, the dynamical polarizability of a cluster is largely determined by the collective excitations of the electrons of the cluster. For a nonrelativistic incident particle, such a situation was considered in [3].

In the relativistic limit $\gamma \gg 1$, formula (17) is conveniently represented as

$$\omega_{\max} = \frac{1}{r_{WS}\sqrt[3]{N}} \frac{2\gamma^2}{4\gamma^2 \sin^2(\theta/2) + 1}, \qquad (19)$$

which explicitly demonstrates the dependence of the maximal frequency $\omega_{\max}$ on the energy of the incident particle. It follows from equality (19) that, in contrast to the nonrelativistic case, the effect of cooperative phenomena on the spectrum of PB in the relativistic case essentially depends on the radiation angle of a photon. For small angles and large values of the Lorentz factor of the incident particle, the maximal frequency above which the cooperative phenomena in PB disappear may reach significant values. However, one should keep in mind that, in the range of angles $\theta < \gamma^{-2}$, the static channel is dominant in the bremsstrahlung of an electron (positron). Therefore, the problem concerning the role of cooperative phenomena in the spectrum of bremsstrahlung of a light charged particle one a cluster should be solved with regard to the specific values of the parameters of the problem. At the same time, for the bremsstrahlung of heavy charged particles, when SB is negligible, a spectral constraint on the role of cooperative phenomena in the relativistic case is given by frequency (19).

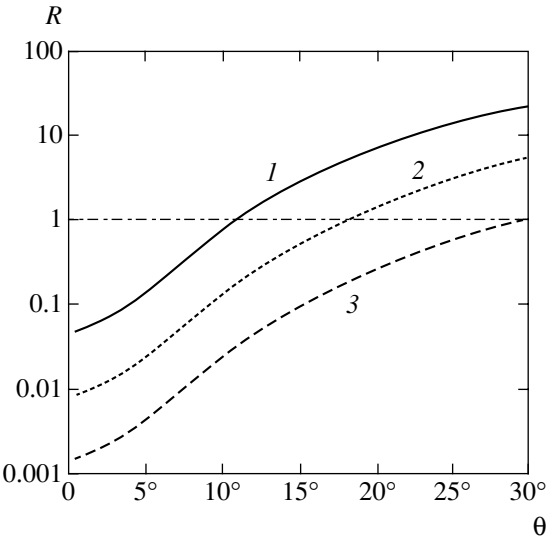

Fig. 3. Angular dependence of the R factor for different numbers of atoms in a copper cluster, including the monoatomic case: $\hbar\omega = 1$ keV and $\gamma = 10$; (*1*) $N = 100$, (*2*) $N = 10$, and (*3*) $N = 1$.

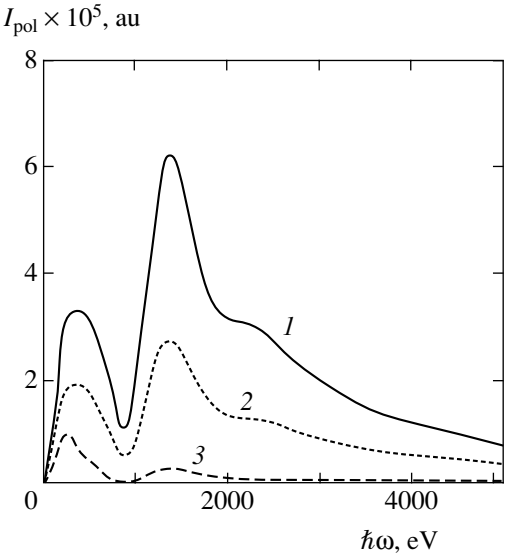

Fig. 4. The spectrum of PB on a copper cluster ($N = 10$) for different values of the Lorentz factor and a radiation angle of 0.5 rad: (*1*) $\gamma = 10^2$, (*2*) $\gamma = 10$, and (*3*) $\gamma = 1.1$.

The spectrum of PB on a copper cluster consisting of ten atoms as a function of the energy of the incident particle is shown in Fig. 4 for a radiation angle of 0.5 rad. One can see that, as the Lorentz factor increases, the radiation intensity increases and the max-

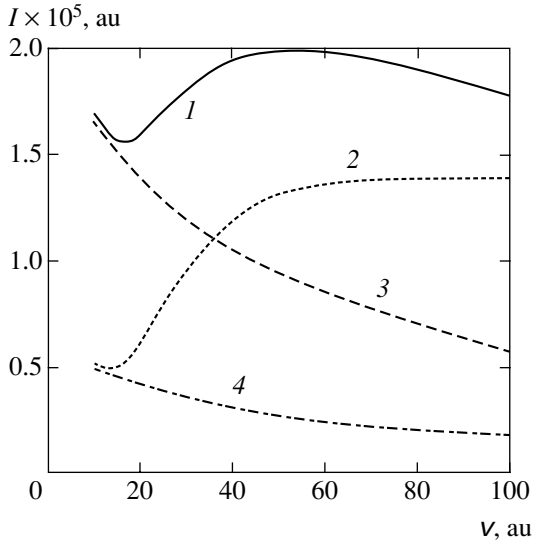

Fig. 5. The intensity of the total bremsstrahlung (*1* and *3*) and PB (*2* and *4*), normalized by the number of atoms in a cluster, as a function of the electron velocity in the nonrelativistic case for a copper cluster ($N = 20$; curves *1* and *2*) and for an isolated atom ($N = 1$; curves *3* and *4*); $\hbar\omega = 200$ eV and $\theta = 1$ rad.

imum of the spectrum shifts toward higher frequencies. These variations are especially manifest when passing from fast but nonrelativistic incident particles to weakly relativistic particles. As the Lorentz factor increases further, the variation of the spectrum of PB on a cluster is not so manifest. As the radiation angle decreases, the PB spectrum of relativistic incident particles moves to higher frequencies according to formula (19), which also implies that the maximal frequency decreases as the number of atoms in the cluster increases.

The intensity of the total bremsstrahlung and of PB, normalized by the number of electrons, for the scattering of nonrelativistic electrons by a copper cluster and by an isolated atom are shown in Fig. 5 as a function of the velocity of the incident particle for a radiation angle of $\theta = 1$ rad and a photon energy of $\hbar\omega = 200$ eV. One can see that, in the nonrelativistic case, there exists an optimal value of the electron velocity v_{opt} for which the intensity of PB on a cluster is maximal. The dependence of this optimal velocity on the parameters of the problem is given by

$$v_{\mathrm{opt}} = \omega r(N). \qquad (20)$$

In the case of scattering by an isolated atom, the cluster radius in (20) must be replaced by the characteristic atomic radius r_a. This formula can be obtained with the use of the expression (13) for the minimal transferred momentum in the case of $v \ll c$.

Equality (20) implies that, as the frequency and the cluster size decrease, the optimal velocity decreases. In particular, for a process on an isolated atom and a photon energy of $\hbar\omega = 200$ eV, the value of v_{opt} lies outside the applicability domain of the Born approximation; therefore, the intensities of bremsstrahlung and PB are monotonically decreasing functions of the velocity of the incident particle (Fig. 5). As the number of atoms in the cluster increases, the optimal velocity increases, as is clear from formulas (20) and (6).

As the radiation angle decreases, the peak in the dependence of the bremsstrahlung intensity on the electron velocity disappears; for PB, this peak becomes less pronounced. On the one hand, this is associated with the increased contribution of the static channel to the process (see formula (9)) and, on the other hand, with the increase in the value of the minimal transferred momentum (13).

Figure 5 illustrates the vanishing of the cooperative phenomena as the velocity of the incident particle decreases: for a given photon energy (200 eV), the total bremsstrahlung and PB on a cluster and on a single atom coincide in the limit of small velocities.

In the high-frequency range, the intensity of PB is a monotonically increasing function of the particle velocity; this is associated with the natural restriction imposed on the optimal velocity of the incident particle: $v_{\mathrm{opt}} < 137$. A monotonic increase in the PB intensity as a function of the energy of the incident particle is also characteristic of the relativistic case (except for the case of low velocities, which is not considered here).

4. CONCLUSIONS

Based on the analysis carried out in this work, one can conclude that the cooperative phenomena play an important role in the bremsstrahlung of a fast (including a relativistic) charged particle scattered by a polyatomic cluster over a wide range of frequencies. These phenomena, associated with the constructive interference of the contributions of atoms to the process by the polarization channel, lead to a nonlinear increase in the intensity of PB as a function of the number of atoms in a cluster. At the same time, for the static mechanism of bremsstrahlung, the contribution of different atoms to the radiation is noncoherent, which is associated with the small values of the impact parameters at which the SB is formed.

In this work, we have investigated the cooperative phenomena as a function of the number of atoms in a cluster. We have determined the number N_{sat} (starting from which the PB intensity ceases to grow nonlinearly) as a function of the basic parameters of the problem. In particular, we have shown that, for relativistic incident particles, the value of N_{sat} strongly increases as the radiation angle decreases. We have also shown that the cooperative phenomena substantially modify the basic characteristics of bremsstrahlung on a cluster compared with that on a single atom. For exam-

ple, in the high-frequency case, the radiation pattern of PB is narrowed down as the number of atoms increases; for sufficiently large clusters, the angular dependence of the intensity of PB of relativistic particles becomes nonmonotonic: a maximum appears for nonzero radiation angles.

As the energy of the incident particle increases, the maximum of the spectral distribution of PB on a cluster moves to higher frequencies. In the relativistic case, the shape of the high-frequency region of the spectrum strongly depends on the radiation angle. For small angles, the intensity of bremsstrahlung decreases, as frequency increases, much slower than for large angles.

The analysis of the bremsstrahlung intensity as a function of the velocity of the incident particle has shown that, in the nonrelativistic case, the character of this function varies from monotonically increasing to monotonically decreasing behavior. In the relativistic limit, the PB intensity monotonically increases with the particle energy. In the limit of small velocities of the incident particle, the role of cooperative phenomena in the bremsstrahlung on a cluster becomes negligible.

The results obtained can be used for interpreting experimental data on the bremsstrahlung of fast charged particles on clusters for sufficiently high energies of the photons.

REFERENCES

1. A. V. Korol', A. G. Lyalin, and A. V. Solov'ev, *Polarization Bremsstrahlung* (S.-Peterb. Gos. Pedagog. Univ., St. Petersburg, 2004) [in Russian].
2. V. A. Astapenko, L. A. Bureeva, and V. S. Lisitsa, Usp. Fiz. Nauk **172**, 155 (2002) [Phys. Usp. **45**, 149 (2002)].
3. J. P. Connerade and A. V. Solov'yov, J. Phys. B **29**, 3529 (1996).
4. L. I. Kurkina, Fiz. Tverd. Tela (St. Petersburg) **46**, 538 (2004) [Phys. Solid State **46**, 557 (2004)].
5. V. A. Astapenko, L. A. Bureeva, and V. S. Lisitsa, Zh. Éksp. Teor. Fiz. **117**, 96 (2000) [JETP **90**, 434 (2000)].
6. V. A. Astapenko, V. M. Buĭmistrov, L. K. Mikhaĭlov, *et al.*, Zh. Éksp. Teor. Fiz. **88**, 1560 (1985) [Sov. Phys. JETP **61**, 930 (1985)].
7. V. A. Astapenko, Zh. Éksp. Teor. Fiz. **126**, 1101 (2004) [JETP **99**, 958 (2004)].
8. http://cindy.lbl.gov/optical_constants.
9. V. P. Shevelko, I. Yu. Tolstikhina, and Th. Stolker, Nucl. Instrum. Methods Phys. Res. B **184**, 295 (2001).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 80–87.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 95–102.

ATOMS, MOLECULES, OPTICS

Vacuum Heating of Large Atomic Clusters by a Superintense Femtosecond Laser Pulse¶

A. V. Getz and V. P. Krainov

Moscow Institute of Physics and Technology, Dolgoprudnyĭ, Moscow oblast, 141700 Russia
e-mail: krainov@online.ru

Received March 11, 2005

Abstract—The analytic approach of vacuum (Brunel) heating mechanism is generalized to the case of large atomic clusters irradiated by a superintense femtosecond laser pulse. The hydrodynamic cluster expansion is taken into account in this approach. Simple universal expressions are obtained for the absorbed laser energy by a cluster and for the radius of an expanding cluster. The absorption of laser energy and the cluster expansion are determined by only one dimensionless field parameter. *© 2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Rapid developments in intense ultrashort laser technology have opened a new regime of laser–cluster interaction, in which intense laser pulses deposit their energy into solid targets faster than the hydrodynamic expansion of the cluster surface occurs. Hot electrons in laser–cluster plasmas can be generated by different absorption or acceleration mechanisms under different experimental conditions. At low laser intensities, inverse bremsstrahlung is the main absorption mechanism, which depends on the electrical conductivity associated with electron mean free path comparable to the interatomic spacing inside the cluster [1]. This absorption strongly decreases as the laser intensity increases.

Brunel [2] proposed that high-intense *p*-polarized laser pulses incident obliquely on an atomically abrupt metal surface could be strongly absorbed by pulling electrons into vacuum during an optical cycle and then returning them to the surface with approximately the quiver velocity. This is the so-called vacuum heating process.

Recently, hot electron generation was studied at high intensities [3]. Hot electron spectra and X-ray spectra from the bremsstrahlung radiation, when hot electrons undergo deceleration in solid targets, showed that, when smooth solid targets were irradiated obliquely by *p*-polarized laser pulses, a group of hot electrons could be heated to a relatively low Maxwellian temperature. Another group of hot electrons with higher energies could be produced by nonlinear resonant absorption if there was a thin layer of preplasma in front of the target surface. Electrons produced by the inverse bremsstrahlung absorption process are known as thermal electrons and have energies less than 1 keV at modest laser intensities. The electrons generated by resonance absorption and the other nonlinear resonant absorption are called hot electrons and have much higher kinetic energies. The energy absorption and the hot electron generation in the interaction of *p*-polarized femtosecond laser pulses with aluminum solid targets have been studied. The laser delivered 150 fs pulses and produced a peak irradiance of 8×10^{15} W/cm^2 at the focus. The measurements suggest that vacuum heating is the main heating mechanism for hot electrons with high energies.

It is clearly seen from other experimental data [4] that, in the case of irradiation of solid targets by a *p*-polarized laser, the outgoing electrons are extracted from the critical surface by the Brunel absorption, once in the laser oscillating period because the electron bunch length is almost equal to the laser wavelength. The laser wavelength and the intensity were 1 μm and 2×10^{18} W/cm^2, respectively, in these experiments.

The third heating mechanism is the elastic reflections of inner electrons from the cluster surface [5, 6]. The absorption of laser energy is equal to $F_{\text{in}}^2/2\omega^2$ at each collision in the presence of laser field (analogously to the induced inverse bremsstrahlung at the collision of an electron with an atomic ion). Here, F_{in} is the electric field strength inside the cluster. The system of units $m_e = e = 1$ is used throughout the paper. Then, the average collision rate ν is determined from the electron motion inside the cluster:

$$\nu \sim V_e/R.$$

Here, V_e is the electron thermal, or quiver, velocity inside the cluster. This mechanism is effective only near the Mie resonance when the Mie frequency

$$\omega_{\text{Mie}} = \omega_p/\sqrt{3}$$

¶ This article was submitted by the authors in English.

1063-7761/05/10101-0080$26.00 © 2005 Pleiades Publishing, Inc.

is equal to the laser frequency ω during the cluster expansion. Then, F_{in} can be larger than the external laser field strength F, resulting in large values of V_e. Far from the Mie resonance, we obtain

$$F_{in} \approx -\left(\frac{\omega}{\omega_{Mie}}\right)^2 F \ll F.$$

We see that the collision rate ν decreases as the cluster size R increases.

Inverse induced bremsstrahlung and vacuum heating are very different heating mechanisms. Electron heating at the inverse induced bremsstrahlung is proportional to the duration of the laser pulse. An electron acquires twice the oscillation field energy at each collision with an atomic ion, but, inside the cluster, the electric laser field is much less compared to the external laser field, because the electron density in the cluster is higher than the critical density. The acquired electron energy at each collision is therefore small, but large absorption in experiments occurs due to large duration of a laser pulse at the leading edge of this pulse. At the trailing part of the laser pulse, the electron plasma becomes subcritical, and the external laser field freely penetrates the entire cluster. The doubled oscillation energy becomes large, but the cluster begins to expand so quickly that the rate of electron–ion collisions becomes very small, and there is no electron heating at the trialing edge. If now the duration of laser pulse is only 50 fs, then there is no time for many electron–ion collisions. Oppositely, the vacuum heating mechanism operates only with the external laser field, because an electron is ejected from the cluster, is heated by the external laser field, and returns to the cluster (where the internal laser field is negligibly small). Therefore, even a small number of electron ejections during the leading part of the femtosecond laser pulse results in large electron heating.

It follows from [7] that, for clusters with radii larger than 10 nm, only a small amount of electrons leave the cluster, i.e., the Coulomb explosion mechanism is not realized, and hydrodynamic pressure of the free electron gas inside the cluster is the dominating mechanism for cluster expansion. The second requirement for fulfillment of the hydrodynamic approximation is that the laser pulse duration be large compared to the time between the neighboring electron–electron collisions. In the opposite case, a particle-in-cell description is required. This means in practice that the laser intensity should be a nonrelativistic quantity, because relativistic electrons practically do not collide with each other (and with atomic ions) during a femtosecond laser pulse.

In this paper, we report a theoretical study on hot electron generation in the interaction of superintense femtosecond laser pulses with large atomic clusters. The vacuum (Brunel) mechanism is generalized to the case of large atomic clusters with the hydrodynamic cluster expansion during the laser pulse taken into account. Femtosecond laser pulses are needed for experiments with clusters in order to prevent fast cluster decay before the peak of the laser pulse. A simplified version of the vacuum heating of electrons for deuterium clusters was considered in [8]. However, this approach did not take the screening effects at the electron ejection into account. Besides, the acquired electron kinetic energy was estimated only qualitatively. Finally, cluster expansion and multiple inner ionization of atomic ions were not considered in [8].

2. VACUUM HEATING

To describe the Brunel energy absorption by a large atomic cluster, we first consider each small part of the cluster surface as a plane irradiated by a superintense femtosecond laser pulse. This is valid if the excursion length F/ω^2 of a free electron ejected by laser field from the cluster is less than the cluster radius R. Here, F and ω are the field strength amplitude and field frequency, respectively. For example, if R = 20 nm (the number of Kr atoms is $N_a = 4.7 \times 10^5$ in such a cluster), $\hbar\omega$ = 1.5 eV (Ti:sapphire laser), then the peak laser intensity should be less than 5×10^{16} W/cm^2 ($F < 1.2$ a.u.). In practice, the inequality

$$F/\omega^2 \geq R$$

can also be permitted, because an electron is ejected from the cluster with zero velocity, and, therefore, it moves and returns to the cluster along the same curved electric field line.

Free electrons are produced at the leading edge of the laser pulse due to a single-ionization process. The local coordinate x is assumed to be directed along the normal to the cluster surface, and $x > 0$ is the region inside the cluster plasma. Inside the liquid cluster matter, the plasma frequency

$$\omega_p = \sqrt{4\pi n_e}$$

is large compared to the laser frequency ω, and, hence, the external electric field practically does not penetrate the cluster. Here, n_e is the electron number density. In addition, we assume that the cluster is sufficiently large, such that the outer ionization can be neglected, unlike in the case of inner ionization [9]. According to the Bethe rule for barrier-suppression outer ionization [9],

the charge of the cluster ion (i.e., the number of ejected electrons) is

$$Q = 4F_{\text{in}}R^2.$$

The condition of a weak outer ionization

$$Q \ll N_e = n_e\frac{4\pi R^3}{3}$$

(N_e is the total number of electrons in the cluster) corresponds to the requirement that

$$F_{\text{in}}/\omega_p^2 \ll R.$$

This inequality is usually fulfilled because $\omega_p \gg \omega$. For example, in the case of a Kr cluster, we have $n_e = 2.0 \times 10^{-3}$ a.u. for single ionization, and $\omega_p = 0.160$ a.u. $\gg \omega = 0.055$ a.u. If $R = 20$ nm, the above inequality is violated only when $F_{\text{in}} > 10$ a.u. and $F > 100$ a.u.

Following Brunel's approach [2], we assume the vacuum region for $x < 0$, where the normal component of the incident linearly polarized laser field

$$F(t) = F\sin\omega t$$

is present on the surface. Here,

$$F = F_0\exp(-t^2/\tau^2)$$

is the Gaussian envelope of the laser pulse and τ is the pulse duration. The reflected field $F(t)$ coincides with the incident field within the laser wavelength λ from the cluster surface. We assume that the excursion length

$$F/\omega^2 < R \ll \lambda,$$

where λ is the laser wavelength. As the field increases for $t > 0$, electrons are pulled out during the first half of the laser period $0 < t < \pi/\omega$. The $(l + 1)$th electron feels the total electric field strength $F(t) + F_l(t)$, where

$$F_l(t) = -4\pi\sum_{i=1}^{l} d\sigma_i = -4\pi\sum_{i=1}^{l} n_i(x_i)dx_i \qquad (1)$$

is the electric field strength produced by the previously ejected electrons and their images inside the cluster, $d\sigma_i$ is the surface number density of the ith electron, and $n_i(x_i)$ is the volume number density of the ejected ith electron ($d\sigma_i = n_i dx_i$). The electrons that are ejected after the $(l + 1)$th electron do not contribute to Eq. (1), because these electrons and their images are on one side of the considered $(l + 1)$th electron.

The quantity F_l is independent of the time t, because the electric field produced by a uniformly charged plane is independent of the distance between this plane and the cluster surface. Hence, putting $t = t_l$ in Eq. (1), we obtain

$$F_l(t) = F_l(t_l) = -F\sin\omega t_l. \qquad (2)$$

The last relation follows from the statement that the total electric field vanishes on the cluster surface, $F(t) + F_l(t) = 0$ when $t = t_l$ and $x_l = 0$. Thus, the total electric field is equal to $F(t) + F(t_l)$ for the $(l + 1)$th electron. The motion of this electron is described by the Newton equation

$$\frac{dv_l}{dt} = -F(\sin\omega t - \sin\omega t_l)$$

($l \gg 1$, and, hence, there is no real difference between the numbers l and $l + 1$). We can integrate this equation to obtain the velocity of the lth electron under the initial condition of a cold ejected electron $v_l(t_l) = 0$:

$$v_l(t) = \frac{F}{\omega}[\cos\omega t - \cos\omega t_l + \omega(t - t_l)\sin\omega t_l]. \qquad (3)$$

We can integrate Eq. (3) to obtain the position of the lth electron under the initial condition $x_l(t_l) = 0$:

$$x_l(t) = \frac{F}{\omega^2}\Big[(\sin\omega t - \sin\omega t_l) - \omega(t - t_l)\cos\omega t_l + \frac{\omega^2}{2}(t - t_l)^2\sin\omega t_l\Big]. \qquad (4)$$

The time instance for the return of the lth electron inside the cluster is determined from the universal implicit equation $x_l(t) = 0$, or

$$\sin\omega t - \sin\omega t_l = \omega(t - t_l)\cos\omega t_l - \frac{\omega^2}{2}(t - t_l)^2\sin\omega t_l. \qquad (5)$$

A nonzero ($t \neq t_l$) solution of this equation is possible only in the interval $0 < t_l < \pi/2\omega$ (the first quarter of the laser period). We restrict ourselves the range $t_l < t < 2\pi/\omega$; $t = t_l$ when $t_l = \pi/2\omega$ (no return), but $t > \pi/2\omega$ when $t_l < \pi/2\omega$. The dependence $t(t_l)$ is shown in Fig. 1. It is seen that small values $\omega t_l \ll 1$ correspond to large values of the returning time $t > 2\pi/\omega$. We neglect their contribution, although they slightly distort the electron ejection during the next laser period, because these electrons have small electron density n_l (see below Eq. (8)). The approximate solution of Eq. (5) is then given by

$$t \approx \frac{2}{\omega^2 t_l} \gg \frac{1}{\omega}.$$

We can repeat all operations during the next laser period analogously to the above consideration.

The absorbed energy over one laser cycle per unit square is given by [2]

$$E_e = \sum \frac{1}{2} v_l^2 n_l dx_l = \frac{1}{2} \int_{\pi/2\omega}^{2\pi/\omega} v_l^3 n_l dt. \quad (6)$$

The electron number density n_l can be found by differentiating Eq. (1) with respect to t_l at a fixed value of t and taking Eq. (2) into account:

$$\frac{dF_l}{dt_l} = -4\pi n_l \frac{dx_l(t)}{dt_l} = -F\omega\cos\omega t_l. \quad (7)$$

Differentiating Eq. (4) with respect to t_l at a fixed value of t and substituting the result in Eq. (7), we obtain the quantity n_l:

$$n_l(t) = \frac{1}{2\pi(t-t_l)^2}. \quad (8)$$

Substituting Eqs. (3) and (8) in Eq. (6), we find

$$E_e = \frac{F^3}{4\pi\omega^3} \times \int_{\pi/2\omega}^{2\pi/\omega} \frac{[\cos\omega t - \cos(\omega t_l + \omega(t-t_l)\sin\omega t_l)]^3}{(t-t_l)^2} dt. \quad (9)$$

The integral is done numerically, and E_e can be written as

$$E_e = \frac{\eta}{8\pi\omega^2}(3F\cos\theta)^3, \quad (10)$$

where $\eta = 0.75$. In Eq. (10), we substituted the external field strength F far from cluster by the field strength at the cluster surface

$$F_s = 3F\cos\theta,$$

which is normal to the cluster surface at each surface point (under the condition that the electric field F_{in} inside the cluster is very small). Here, θ is the angle between the polarization of the linearly polarized laser field and the normal to the cluster surface.

The maximum number of returning electrons during one laser cycle per unit square is

$$N = \frac{F_s}{4\pi} = \frac{3F\cos\theta}{4\pi}.$$

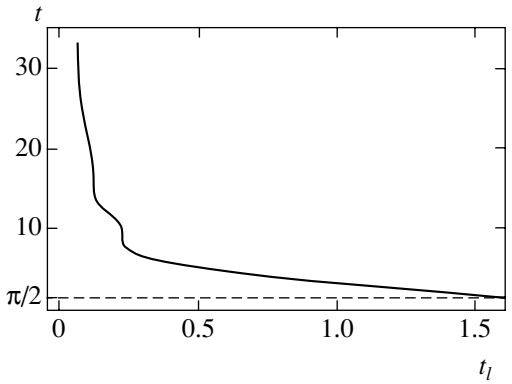

Fig. 1. The universal dependence of the returning time t on the time t_l of the electron ejection according to Eq. (5).

The averaged energy ε of a hot electron is

$$\varepsilon = \frac{E_e}{N} = \frac{\eta\overline{(3F\cos\theta)^2}}{2\omega^2} = \frac{3\eta F^2}{2\omega^2},$$

i.e., it is of the order of the ponderomotive energy (as in a rescattering process for an atom [10], but with a different numerical coefficient). The maximum number of returning electrons during one laser cycle for the whole cluster is (see also Eq. (12) below)

$$N_r = 2\int_0^{\pi/2} N(2\pi R\sin\theta)Rd\theta = \frac{3FR^2}{2}.$$

The number of free electrons inside the cluster is

$$N_e = n_e\frac{4\pi R^3}{3} = \frac{\omega_p^2 R^3}{3}.$$

We then have

$$\frac{N_r}{N_e} = \frac{9F}{2\omega_p^2 R}.$$

For example, when $R = 20$ nm, $F = 0.5$ a.u. (the peak laser intensity is approximately equal to 10^{16} W/cm^2), $\omega = 1.5$ eV, and $\omega_p = 0.16$ a.u. (single ionization of atoms in the Kr cluster), we obtain $N_r/N_e \approx 0.25$. The number of electrons heated during the laser pulse can be estimated by multiplying the quantity N_r/N_e by the number of laser periods $\omega\tau/2\pi = 54$ (at $\tau = 150$ fs). Thus, the conclusion can be made that all electrons are heated in the cluster during the laser pulse.

The mean free time of electron–electron collisions inside the Kr cluster,

$$\tau_{ee} = \frac{3T_e^{3/2}}{4\sqrt{2\pi}n_e \ln\Lambda}$$

($\ln\Lambda \sim 10$ is the Coulomb logarithm), is larger than the pulse duration $\tau = 150$ fs for the electron temperature $T_e > 1.5$ keV during the laser pulse. Therefore, there is no Maxwell distribution among electrons during the femtosecond laser pulse [11]. The kinetic energy of the heated electron in the peak of the laser pulse is $\varepsilon = 3\eta F^2/2\omega^2 \sim 2.5$ keV at $F = 0.5$ a.u.

The absorbed power I per unit square of the cluster surface can be found by dividing E_e over the laser period $2\pi/\omega$. Thus,

$$I = \frac{27\eta}{16\pi^2\omega}F^3\cos^3\theta. \quad (11)$$

We now integrate I over the cluster surface. The power W absorbed by the entire cluster is given by

$$W = 2\int_0^{\pi/2} I\cos^3\theta(2\pi R\sin\theta)Rd\theta = \frac{27\eta R^3}{16\pi\omega}F^3. \quad (12)$$

3. CLUSTER EXPANSION

The energy of laser radiation is first absorbed by the cluster electrons; then, this energy transforms into the kinetic energy of the atomic ions. The cluster expansion is a nonequilibrium statistical process due to high expansion velocities. Thus, the increasing pressure of the electron gas is not compensated by the external medium. Therefore, the well-known expression $P_e dV$ (P_e is the pressure of the electron gas) of the equilibrium statistical physics is not applicable for derivation of the expansion work. Instead, we calculate the variation of the kinetic energy of atomic ions inside the expanding cluster. We assume that the velocity of the radial motion of atomic ions $v(r)$ is a linear function of the radial variable r:

$$v(r) = \frac{dR}{dt}\frac{r}{R}.$$

The kinetic energy of atomic ions in the spherical layer having the width dr is

$$dE_k = 4\pi r^2\frac{n_a M_a v^2(r)}{2}dr,$$

where M_a is the mass of an atomic ion and n_a is the number density of atomic ions. The total kinetic energy of cluster ions can be obtained by integration over r:

$$E_k = \int_0^R 4\pi r^2\frac{n_a M_a v^2(r)}{2}dr = \frac{3}{10}N_a M_a\left(\frac{dR}{dt}\right)^2.$$

We took into account that the number density of atomic ions is

$$n_a = \frac{3N_a}{4\pi R^3},$$

and N_a is the total number of atoms in the cluster. The time derivative of this energy gives the expansion work produced per unit time:

$$\frac{dE_k}{dt} = \frac{3}{5}N_a M_a\frac{dR}{dt}\frac{d^2R}{dt^2}. \quad (13)$$

According to Eqs. (12) and (13), the cluster vacuum heating is determined from the energy balance equation

$$\frac{dE}{dt} = \frac{27\eta R^2}{16\pi\omega}F_0^3\exp\left(-\frac{3t^2}{\tau^2}\right) - \frac{3}{5}N_a M_a\frac{dR}{dt}\frac{d^2R}{dt^2}, \quad (14)$$

where $E(t)$ is the total thermal energy of the heated electrons inside the cluster as a function of time t. The cluster expansion is determined from the Newton equation

$$\frac{d^2R}{dt^2} = \frac{3P_e}{n_a M_a R}. \quad (15)$$

The electron pressure is

$$P_e = \frac{2E}{3V}.$$

This expression is also valid for nonequilibrium process of the cluster expansion. The electron pressure can be produced by both the electron–electron collisions and elastic reflections of electrons from the cluster surface (the latter occurs for high-energy electrons). Hence,

$$\frac{d^2R}{dt^2} = \frac{2E}{M_a N_a R}, \quad (16)$$

where

$$N_a = n_a V = \text{const}$$

is the number of atomic ions in the cluster. Equations (14) and (16) are to be solved under the initial conditions

$$R(-\infty) = R_0, \quad \frac{dR}{dt} = 0 \text{ as } t \longrightarrow -\infty,$$
$$E(-\infty) = 0.$$

We introduce the dimensionless radius $R/R_0 \longrightarrow R$, the dimensionless time $t/\tau \longrightarrow t$, the dimensionless electron energy

$$\frac{2\tau^2}{M_a N_a R_0^2} E \longrightarrow E,$$

and the dimensionless field parameter

$$s = \frac{27\eta(F_0\tau)^3}{16\pi\omega M_a N_a}. \tag{17}$$

Then, the dimensionless universal equations can be written as

$$\frac{dE}{dt} = sR^2\exp(-3t^2) - \frac{6E}{5R}\frac{dR}{dt},$$
$$\frac{d^2R}{dt^2} = \frac{E}{R} \tag{18}$$

and the initial conditions become

$$R(-\infty) = 1, \quad \frac{dR}{dt} = 0 \text{ as } t \longrightarrow -\infty, \quad E(-\infty) = 0.$$

It is seen that the vacuum heating of cluster electrons and the cluster expansion are determined by only one universal dimensionless field parameter s (Eq. (17)). In Figs. 2 and 3, the functions $R(t)$ and $E(t)$ are shown for various values of the parameter s. It should be noted that the multiple inner ionization of atoms inside the cluster does not change the obtained results. It decreases the average energy of a hot electron only. It follows from Eq. (18) that the velocity of cluster expansion

$$V = \frac{dR}{dt} = \text{const as } t \longrightarrow +\infty,$$

and

$$E(t \longrightarrow +\infty) \sim t^{-6/5} \longrightarrow 0$$

due to adiabatic cooling of the cluster electrons after the end of the laser pulse.

In Fig. 4, we present the final dimensionless velocity of the cluster expansion $V(t \longrightarrow +\infty)$ as a function of the field parameter s at $t \longrightarrow +\infty$. It is seen that this dependence can be approximated by the simple linear dependence $V = 0.8s$.

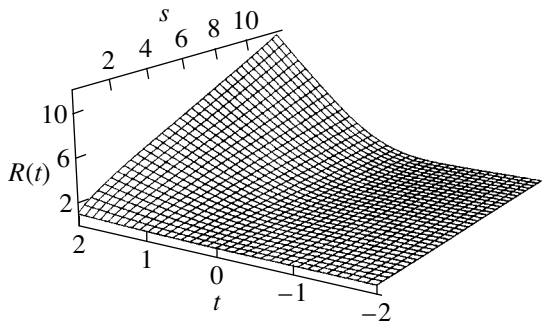

Fig. 2. The universal dependence of the cluster radius $R(t)$ (in units of the initial cluster radius R_0) on time t (in units of the pulse duration τ) for different values of the field parameter s (Eq. (17)).

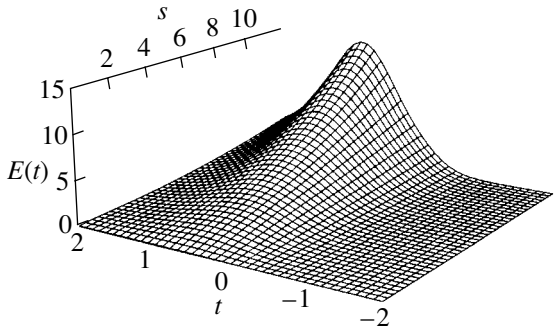

Fig. 3. The universal dependence of the absorption energy $E(t)$ (in units of $M_a N_a R_0^2/2\tau^2$) by a cluster on time t (in units of the pulse duration τ) for different values of the field parameter s.

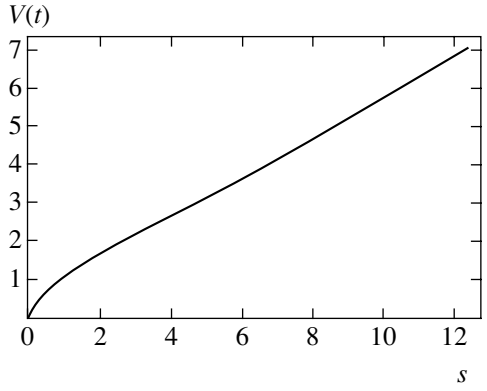

Fig. 4. The final velocity of the cluster expansion $V(t)$ as a function of the field parameter s as $t \longrightarrow +\infty$.

For the laser–Kr cluster interaction at $F_0 = 1$ a.u., $R_0 = 20$ nm, and $\tau = 100$ fs, we find $s = 7.0$. Then, $V = 5.6$ a.u. In the usual units, we find $V = 5.6R_0/\tau$. It then follows that $V \approx 1.1$ nm/fs, which is in agreement

with Fig. 3 in [12] for the same values of the parameters. Hence, in this example, the maximum kinetic energy of Kr ions is approximately equal to 530 keV. Thus, most of the electron heating energy transforms into the energy of cluster ion expansion; the kinetic energy of the heated electron in the peak of the laser pulse is

$$\varepsilon = \frac{3\eta F^2}{2\omega^2} \approx 10 \text{ keV at } F = 1 \text{ a.u.}$$

4. CONCLUSIONS

The hydrodynamic expansion of large Xe clusters (1.5×10^5 atoms per cluster) was observed at the irradiation by a laser pulse with the peak intensity approximately 10^{16} W/cm^2 [13]. The duration of the laser pulse was $\tau = 200$ fs. The authors also took the asymmetric pressure of the laser field into account in addition to the hydrodynamic pressure. It can be effective at plasma resonance when the Mie frequency coincides with the laser frequency. The field pressure produces an asymmetric cluster expansion.

Electron kinetic energy spectra were measured in [14] from the interaction of Kr and Xe clusters with a high-intensity 800 nm femtosecond laser pulse. The sizes of clusters were 10^4 and 2×10^6 atoms per cluster. The peak intensity was 5×10^{16} W/cm^2 at the shortest pulse duration $\tau = 50$ fs. The cutoff in the electron energy spectrum was found to be at $\varepsilon = 6$ keV. According to our approach, the cutoff for electrons inside the cluster is

$$\varepsilon = \frac{3\eta F_0^2}{\omega^2} \approx 14 \text{ keV}.$$

The difference is explained by a decrease of the electron energy due to the adiabatic decrease of the laser field strength when the laser pulse turns off (in accordance with the Lawson–Woodward theorem; see also [15]). This cooling is confirmed by numerical simulation [16] of the interaction of a Xe cluster with a laser pulse ($\lambda = 780$ nm, $\tau = 260$ fs, 4×10^{15} W/cm^2).

Hot electron generation by the vacuum heating process has been studied in the interaction of 150 fs, 5 mJ, 800 nm *p*-polarized laser pulses with solid targets [17]. The measurements have suggested that the "vacuum heating" is the main heating process for hot electrons with high energies. The energy of the vacuum-heated hot electrons has been found to be higher than the prediction from the scaling law of resonance absorption. Particle-in-cell simulations have confirmed that hot electrons are mainly generated by the vacuum heating process under certain experimental conditions.

Spectra of energetic electrons in the 100 keV range were measured from the interaction of intense femtosecond laser pulses with clusters of xenon or argon [18]. The interaction of 28 fs pulses with cluster targets revealed quite different results with respect to the applied laser intensity. At the laser intensity 10^{16} W/cm^2, the absorption by xenon clusters was about 25%; however, at the laser intensity 10^{17} W/cm^2, the laser energy absorption drastically increased to 78%, and the efficient coupling between the laser pulse and cluster target produced electrons with the energy as high as 500 keV. The estimated hot electron temperatures ranged from about 30 to 90 keV depending on atoms.

The Mie resonance is effective only during a very short time (2–3 fs) because of the very fast cluster expansion. Therefore, a short increase of the laser field at the resonance does not result in significant electron heating. The Mie resonance only appears in the self-similar expansion model in [1] and results from the calculation of the dipole moment. In Fig. 7 in [1], the electron energy as a function of time has a sharp and narrow peak because of the Mie resonance. Then, the electron energy strongly decreases, such that the Mie resonance does not actually influence the final electron heating in accordance with the conclusion in the previous section.

In conclusion, the vacuum mechanism suggested by Brunel for irradiation of a planar solid surface by a high-intensity laser pulse is generalized to the case of large spherical clusters with the cluster hydrodynamic expansion taken into account. The universal expressions have been obtained for the absorbed laser energy by a cluster and for the radius of the expanding cluster as functions of time. The absorption of laser energy and the cluster expansion are determined by only one dimensionless field parameter *s* (Eq. (17)).

ACKNOWLEDGMENTS

This work was supported by the International Science and Technology Center (project no. 2155), BRHE (project no. MO-011-0), and the Russian Foundation for Basic Research (project nos. 04-02-16499 and 05-02-16383).

REFERENCES

1. T. Ditmire, T. Donnelly, A. M. Rubenchik, *et al.*, Phys. Rev. A **53**, 3379 (1996).
2. F. Brunel, Phys. Rev. Lett. **59**, 52 (1987).
3. S. Bastiani, P. Audebert, J. P. Geindre, *et al.*, Phys. Rev. E **60**, 3439 (1999).
4. Y. Sentoku, H. Ruhl, K. Mima, *et al.*, Phys. Plasmas **6**, 2855 (1999).
5. F. Megi, M. Belkacem, M. A. Bouchene, *et al.*, J. Phys. B: At. Mol. Opt. Phys. **36**, 273 (2003).

6. V. P. Krainov and A. Roshchupkin, J. Phys. B: At. Mol. Opt. Phys. **34**, L297 (2001).
7. H. M. Milchberg, S. J. McNaught, and E. Parra, Phys. Rev. E **64**, 056402 (2001).
8. V. P. Krainov and M. B. Smirnov, Zh. Éksp. Teor. Fiz. **119**, 719 (2001) [JETP **92**, 626 (2001)].
9. V. P. Krainov and M. B. Smirnov, Phys. Rep. **370**, 237 (2002).
10. P. B. Corkum, Phys. Rev. Lett. **71**, 1994 (1993).
11. M. B. Smirnov, Zh. Éksp. Teor. Fiz. **124**, 48 (2003) [JETP **97**, 42 (2003)].
12. J. Liu, R. Li, P. Zhu, *et al.*, Phys. Rev. A **64**, 033426 (2001).
13. V. Kumarappan, M. Krishnamurthy, and D. Mathur, Phys. Rev. A **66**, 033203 (2002).
14. E. Springate, S. A. Aseyev, S. Zamith, and M. J. J. Vrakking, Phys. Rev. A **68**, 053201 (2003).
15. A. V. Sofronov and V. P. Krainov, J. Phys. B: At. Mol. Opt. Phys. **37**, L329 (2004).
16. E. Springate, N. Hay, J. W. G. Tisch, *et al.*, Phys. Rev. A **61**, 044101 (2000).
17. L. M. Chen, J. Zhang, Q. L. Dong, *et al.*, Phys. Plasmas **8**, 2925 (2001).
18. L. M. Chen, J. J. Park, K. H. Hong, *et al.*, Phys. Plasmas **9**, 3595 (2002).

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 88–97.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 103–114.

NUCLEI, PARTICLES, FIELDS, GRAVITATION, AND ASTROPHYSICS

Energy Spectrum and Anisotropy of Cosmic Rays with $E_0 \geq 10^{17}$ eV from Yakutsk EAS Array Data

A. V. Glushkov and M. I. Pravdin

Shafer Institute of Space Physics and Aeronomy, Siberian Branch, Russian Academy of Sciences, Yakutsk, 677891 Russia

e-mail: a.v.glushkov@ikfia.ysn.ru

Received November 22, 2004

Abstract—Data from the Yakutsk extensive air shower array for the period 1974–2004 are used to analyze the energy spectrum and anisotropy of primary cosmic rays (PCRs) with energy $E_0 \geq 10^{17}$ eV. The spectra from different regions of the sky are shown to differ in shape. Enhanced and reduced particle fluxes come from the disks of the Galaxy and the Supergalaxy (the Local Supercluster of galaxies) at $E_0 \geq 5 \times 10^{18}$ eV and $E_0 \leq (2–3) \times 10^{18}$ eV, respectively. This is interpreted as a manifestation of the possible interaction between extragalactic PCRs and the matter of these spatial structures. © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

Ultrahigh-energy ($E_0 \geq 10^{17}$ eV) cosmic rays have been being recorded on the Yakutsk extensive air shower (EAS) array since 1970 and, in stable operating conditions, since 1974. Over this period, we have repeatedly reported the results of our studies of the primary cosmic ray (PCR) energy spectrum as the experimental data were accumulated [1–5]. The shape and intensity of the spectrum near the cutoff (~10^{20} eV) predicted by Zatsepin and Kuzmin [6] and Greisen [7] are of great importance in revealing the PCR composition and formation sources. The experimental results obtained on different arrays [5, 8, 9] differ in absolute intensity by a factor of 2, but are similar in shape. The intensity variations cannot be described by a single power law. The shape of the spectrum at $E_0 > 10^{19}$ eV is consistent with the assumption that the bulk of the PCR flux originates in extragalactic sources [10, 11]. In this case, the spectrum must exhibit a cutoff near 10^{20} eV. The data from the Yakutsk array [3–5] and the American HiRes array [9] agree with this conclusion, while the data from AGASA (Akeno Giant Air Shower Array) with a large number of events at $E_0 \geq 10^{20}$ eV are in conflict with it [8]. We considered this problem in [12, 13].

Until now, the energy spectrum has been studied without taking into account the locations of the sky areas from which the PCRs arrived. However, as our studies showed [14–23], the primary particle flux deviates appreciably from isotropy in different energy ranges. Below, we consider the influence of this factor on the energy spectrum and the origin of PCRs with $E_0 \geq 10^{17}$ eV as a whole.

2. THE DATA PROCESSING TECHNIQUE

The primary particle energy and flux on the Yakutsk array are determined from the classification parameters, estimated with minimum distortions [2–5]. Before 1992, air showers were selected from the entire array area by the stations that formed equilateral triangles with 1-km sides (a large master). Such master triangles select EASs with $E_0 > 10^{18}$ eV. The charged-particle density measured by ground-based scintillation detectors at the distance $R = 600$ m from the EAS axis with a zenith angle $\rho_{S,\,600}(\theta)$, is a classification parameter for them. The separation between the master stations at the center of the array on an area with a diameter of 1 km is 500 m (a small master). Here, mostly showers with $E_0 \geq 10^{17}$ eV, for which $\rho_{S,\,300}(\theta)$ is better determined, are detected. After 1992, almost the entire area of the array was controlled by the small master. This allowed the spectrum to be analyzed up to its right boundary with homogeneous event selection conditions.

The Greisen–Linsley lateral distribution function (LDF) of charged particles with the parameters estimated on the Yakutsk array [24] is used to determine the axis coordinates and the parameters $\rho_{S,\,600}(\rho_{S,\,300})$:

$$f(R) \propto (R/R_M)^{-1}(1 + R/R_M)^{1-b}, \qquad (1)$$

where R_M is the Moliere radius ($\langle R_M \rangle \approx 70$ m for the Yakutsk array). Subsequently, however, we established [25] that this LDF for air showers with $E_0 \geq 10^{19}$ eV is in poor agreement with the experimental data at dis-

tances $R > 1000$ m from the axis. We suggested a modified fit:

$$f(R) \propto (R/R_M)^{-1.3} \times (1 + R/R_M)^{1.3-b}(1 + R/2000)^{-3.5}. \quad (2)$$

For EASs with $E_0 \geq 2 \times 10^{19}$ eV, the parameter b is almost constant, but depends on the zenith angle. In [13], we determined the axis in such air showers using this refined LDF. As a result, the $\rho_{S,600}$ estimates increased, on average, by about 10% for air showers with the axes located within the array boundaries and up to 20% for air showers with the axes located on the array periphery.

The effective area within which the event recording probability $P \geq 0.9$ with allowance made for the fluctuations of the LDF slope is used in the standard data processing technique to determine the intensity. The total exposure as a function of $\rho_{S,600}$ ($\rho_{S,300}$) and zenith angle is calculated by taking into account the actually operating stations at a given time. The limiting area is bounded by the contour of the corresponding trigger.

To estimate the EAS energy, we use the following relations:

$$E_0 = (4.8 \pm 1.6) \times 10^{17} (\rho_{S,600}(0°))^{1.00 \pm 0.02} \text{ [eV]}, \quad (3)$$

$$\rho_{S,600}(0°) = \rho_{S,600}(\theta) \times \exp((\sec\theta - 1) \times 1020/\lambda_\rho) \text{ [m}^{-2}\text{]}, \quad (4)$$

$$\lambda_\rho = (450 \pm 44) + (32 \pm 15)\log(\rho_{S,600}(0°)) \text{ [g cm}^{-2}\text{]}. \quad (5)$$

The error in formula (3) is determined mainly by the absolute calibration accuracy of the Cherenkov detectors and by the error in the mean atmospheric transparency [2, 3]. It does not affect the shape of the spectrum, but is significant for the intensity estimation.

3. THE ENERGY SPECTRUM

The filled symbols in Fig. 1a indicate the differential energy spectrum obtained in [13] for the small (squares) and large (triangles downward) masters within the array boundaries and for a sample of events with $E_0 \geq 4 \times 10^{19}$ eV from an expanded area with the axes going outside the array boundaries (upright triangles). We took air showers with $\cos\theta \geq 0.5$ ($\theta \leq 60°$) and divided them into $\log E_0 = 0.1$ energy bins.

In general, the shape of the spectrum [13] is identical to that obtained previously [4, 5]. However, the stringent selection of air showers for the energy spectrum leads to the loss of almost all data with $E_0 < 4 \times 10^{17}$ eV. This is because only the most reliably detected

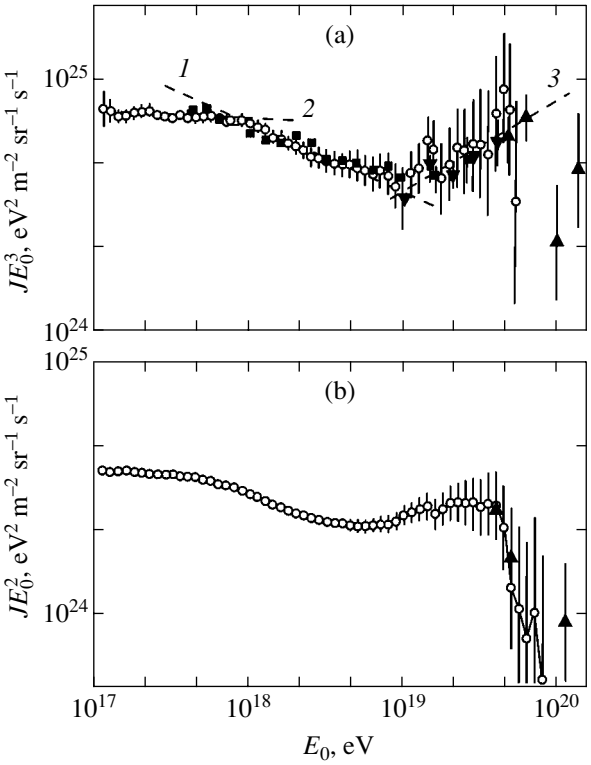

Fig. 1. Differential (a) and integral (b) energy spectra obtained on the Yakutsk EAS array: (○) this paper; our results [13] for the small (■) and large (▼) masters within the array boundaries; (▲) a sample of events with $E_0 \geq 4 \times 10^{19}$ eV from an expanded area with the axes going outside the array boundaries; the dashed lines represent the fits $J(E_0) \approx E_0^{-\gamma}$ with the indices $\gamma_1 = 3.02 \pm 0.02$, $\gamma_2 = 3.27 \pm 0.05$, and $\gamma_3 = 2.62 \pm 0.17$.

air showers (with a probability $P \geq 0.9$) recorded by the master stations when no fewer than eight particles pass through them are selected. In fact, this criterion is needed to correctly estimate the event collection area (at given E_0 and θ), which can be calculated by numerically simulating the entire experiment. Here, the fluctuations of the EAS development, which are not always known completely, play a crucial role.

About 1.2×10^6 air showers have been detected to date on the Yakutsk array. We performed an additional analysis of the events at $10^{17} \leq E_0 < 10^{18}$ eV to see whether they can be used more fully to construct the energy spectrum. It turned out that the effective detection area of such air showers could be found empirically. This is demonstrated by the spectrum indicated in Fig. 1a by the open circles. It was obtained for EASs with $\cos\theta \geq 0.6$ when the entire data set was divided into $\Delta\log E_0 = 0.05$ energy bins. In the above energy range, we restricted our analysis to only six master triangles in the central circle of the array with the radius $R = 500$ m. The spectrum included all of the air showers that were selected by any of these triangles without any

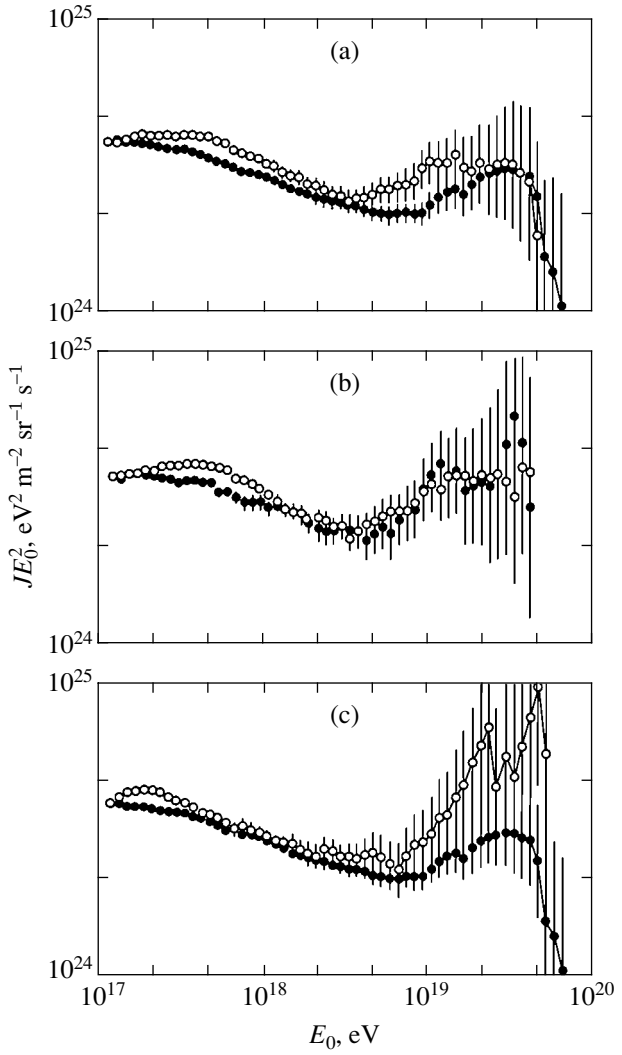

Fig. 2. Integral energy spectra: (a) for 28431 and 7504 air showers with the arrival directions in the Northern (•) and Southern (○) Hemispheres of the Galaxy, respectively; (b) for 3270 and 4234 air showers with the arrival directions at $-10° < b_G < 0°$ (•) and $b_G \le -10°$ (○), respectively; (c) for 28431 and 3276 air showers with the arrival directions at $b_G > 0°$ (•) and $b_G \ge 60°$ (○) in the Northern Hemisphere of the Galaxy, respectively.

selection by the number of particles passed through the stations. It was only required that the EAS axis fall within one of the six circles of radius r inscribed in one of the above six master triangles. The sought-for intensity (at given E_0 and θ) was determined by increasing the sizes of these circles (starting from $r_1 = 20$ m) through a sequential increase in their radius $r_i = 20 + 10(i - 1)$ until the intensity began to decrease. In this case, the $i + 1$ iteration of intensity computation in a given energy bin (with a step of $\Delta \log E_0 = 0.05$) ceased and resumed (from $r_1 = 20$ m) in the next energy bin.

We see from Fig. 1a that the two methods yield identical results. Here, three characteristic portions of the spectrum are clearly distinguished. The first portion (dashed line *1*) corresponds to a power law

$$J(E_0) \propto E_0^{-\gamma} \tag{6}$$

with the index $\gamma_1 = 3.02 \pm 0.02$. The second portion in the energy range $E_0 = (0.7–10) \times 10^{18}$ eV satisfies (6) with $\gamma_2 = 3.27 \pm 0.05$. Finally, the third portion (dashed line *3*) reflects the flattest part of the spectrum with the index $\gamma_3 = 2.62 \pm 0.17$. Below, we will consider and discuss the shape of the spectrum in more detail. For now, we call attention only to the small peak at $E_0 \approx (1–2) \times 10^{19}$ eV. Recently, a wavelet analysis of Yakutsk array data has revealed here a local region of cosmic-ray excess in their arrival directions at a significance level of 0.007 [26]. The pole with the maximum number of events has equatorial coordinates $\alpha_{max} \approx 35° \pm 20°$ and $\delta_{max} \approx 52.5° \pm 7.5°$ and lies in the plane of the Supergalaxy (the Local Supercluster of galaxies).

Figure 1b displays the integral spectrum (35935 events) obtained from the differential spectrum shown in Fig. 1a. The triangles represent the results from [13] for the sample of events with $E_0 \ge 4 \times 10^{19}$ eV from an expanded effective area with the axes going slightly outside the array boundaries. We see that the complex shape of the energy spectrum here is reproduced with a much higher accuracy. Therefore, in our subsequent analysis, we will use the integral spectra.

It was shown in [14–23] that the particle flux at $E_0 \ge 10^{17}$ eV deviates appreciably from isotropy in the PCR arrival directions. The anisotropy has different values and directions in different energy ranges. We believe that this PCR peculiarity can be reflected in the energy spectrum obtained for events from different regions of the sky. Let us first consider the Galaxy. In Fig. 2a, the filled and open circles indicate the spectra for 28431 and 7504 air showers, respectively, from northern ($b_G > 0°$) and southern ($b_G < 0°$) latitudes. Some of the important features that distinguish the two spectra are seen here. First, the north–south asymmetry that we have repeatedly pointed out previously [14, 15, 22, 27] is clearly seen at $5 \times 10^{18} < E_0 < 2 \times 10^{19}$ eV. Second, a statistically significant PCR excess from southern latitudes was also observed at lower energies with its maximum at $E_0 \approx 5 \times 10^{17}$ eV.

To have a more complete picture, let us additionally consider the PCR spectrum from other regions of the Galaxy. Figure 2b shows two approximately identical samples of air showers arriving only from southern latitudes. The spectrum of one of them (3270 air showers) is indicated by the filled circles. It contains events from the latitude band $-10° < b_G < 0°$ immediately adjacent to the Galactic disk. The second sample of 4234 air showers (open circles) includes more southern latitudes $b_G \le -10°$. We see that this sample forms the irregularity of the spectrum at $E_0 \approx 5 \times 10^{17}$ eV noted above.

In Fig. 2c, the open circles indicate the PCR spectrum from the polar ($b_G \geq 60°$) region of the Galaxy for 3276 air showers. The filled circles indicate the spectrum for all events with $b_G > 0°$ (it is shown in Fig. 2a). We see that these spectra also differ significantly in shape. The largest peak compared to the previous spectra is observed in the range of maximum energies ($E_0 \geq 6 \times 10^{18}$ eV). In addition, there is another distinct peak at $E_0 \approx 2 \times 10^{17}$ eV. All these facts are suggestive of a complex formation of the PCR fluxes arriving from different areas of the sky.

In [28], the ASAGA team found a statistically significant anisotropy at $E_0 \approx (5–20) \times 10^{17}$ eV attributable to an enhanced PCR flux from a region near the Galactic center. This result was confirmed and refined by the Australian team on SUGAR (Sydney University Giant Air Shower Recorder) [29]. According to the authors of [28], this flux could be produced by neutrons. It is surmised [30] that such neutrons could originate from several most recent gamma-ray bursts in our Galaxy.

Therefore, the spectrum of ultrahigh-energy particles arriving directly from the Galactic disk is of considerable interest. We considered a sample of (3713) air showers with the arrival directions in the latitude band $|b_G| \leq 5°$. The disk is represented here in the Galactic longitude sector $40° < l_G < 200°$, which is observable on the Yakutsk array. The spectrum from this region of the Galaxy is indicated in Fig. 3a by the open circles. For comparison, the filled circles in this figure indicate the "background" spectrum constructed from 22 465 showers with the arrival directions at Galactic latitudes $|b_G| > 10°$, i.e., lying outside the equatorial region of the Galaxy. We also excluded all of the events with $|b_{SG}| \leq 10°$ that fell into the equatorial region of the Supergalaxy from the background spectrum. The contribution of the Supergalaxy is substantial [14–23], and we will consider it below. For now, we note that the Galactic and Supergalactic planes are almost perpendicular to one another. They intersect in the part of the sky at a Galactic longitude $l_G \approx 137.4°$ surveyed by the Yakutsk array.

In Fig. 3a, we clearly see an excess flux from the Galactic disk at $E_0 > (5–7) \times 10^{18}$ eV that is appreciably higher than the background flux. The region where the Galactic and Supergalactic planes intersect gives the largest contribution [23]. The two spectra are almost indistinguishable at $E_0 < 5 \times 10^{18}$ eV. This is unlikely to agree with the results of [28, 29] where a particle flux from the central region of the Galaxy was detected. Instead of the excess that we expected, we see only a systematic intensity decrease in the spectrum from the Galactic disk in Fig. 3a in the energy range $5 \times 10^{17} < E_0 < 2 \times 10^{18}$ eV with a small peak at $E_0 \approx 10^{18}$ eV. Below, we will return to this question.

Let us now consider the PCR spectrum from the Supergalactic disk. It is indicated in Fig. 3b by the open circles and includes 4295 air showers. We constructed this spectrum for events from the region with Supergalactic latitudes $-8° < b_{SG} < 2°$. No intensity error bars are given in order not to overload the figure (they are approximately the same as those in the spectrum from the Galactic disk in Fig. 3a). The filled circles indicate the background spectrum shown in Fig. 3a by similar circles.

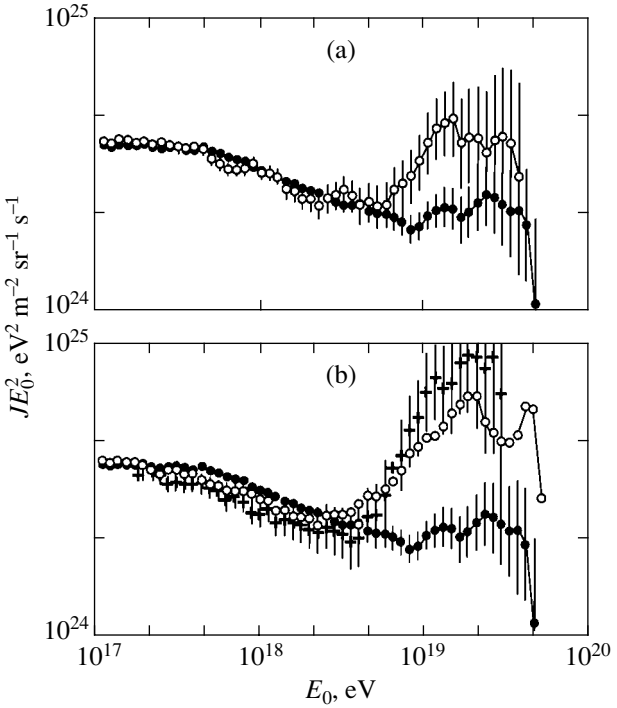

Fig. 3. Integral energy spectra: for 22 465 air showers (•) with the arrival directions outside the equatorial regions of the Galaxy and Supergalaxy ($|b_G| > 10°$ and $|b_{SG}| > 10°$); for 3713 air showers (○) with Galactic latitudes $|b_G| \leq 5°$ (a) and for 4295 air showers in the Supergalactic latitude band $-8° < b_{SG} < 2°$ (b); for air showers (+) from the region of the sky with $-8° < b_{SG} < 2°$ and $100° < l_{SG} < 130°$.

Here, important, in our view, features are immediately apparent. First, an excess particle flux higher than the flux from the Galactic disk is clearly seen in the equatorial region of the Supergalaxy at $E_0 \geq 5 \times 10^{18}$ eV. It is enhanced appreciably (crosses) if a narrower sector ($100° < l_{SG} < 130°$) is taken in the Supergalactic disk, with the particle flux peak in the direction with equatorial coordinates $\alpha \approx 79°$ and $\delta \approx 74°$. Second, 10 to 15% fewer particles than those generally observed for the remaining part of the sky in the background spectrum arrive from the Supergalactic disk at $5 \times 10^{17} < E_0 < 2 \times 10^{18}$ eV.

The results obtained above at $E_0 \geq 5 \times 10^{18}$ eV were not entirely unexpected, since there are numerous experimental data suggesting that the Galaxy plays a role in the origin of the particles at such energies (see, e.g., [14, 15, 27, 31–33]). In addition, Yakutsk data revealed [14–23] that an enhanced PCR flux is also

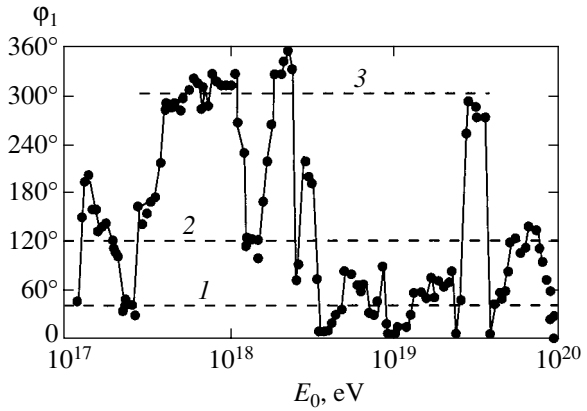

Fig. 4. Variations in the phase of the first harmonic of relation (7) with energy bin position at $\Delta \log E_0 = 0.05$ steps for the sample of showers in the spectrum in Fig. 1.

observed from the Supergalactic disk. Therefore, the spectra in Fig. 3 only confirmed once again these previously established facts.

In contrast, the excess particle fluxes in Fig. 2b at $2 \times 10^{17} < E_0 < 10^{18}$ eV from southern latitudes and in Fig. 2c at $E_0 \approx (1.5–4) \times 10^{17}$ eV from the polar region of the Galaxy require further investigation. These irregularities of the spectrum probably reflect as yet unknown astrophysical processes.

The decrease in particle fluxes at $5 \times 10^{17} < E_0 < 2 \times 10^{18}$ eV in the Galactic disk in Fig. 3, which is particularly clearly seen in the Supergalactic disk, arouses great interest. It is so far hard to tell what caused this decrease. We tested the technique for constructing the energy spectrum, but found no errors. In general, the spectra in Fig. 3 are similar. This suggests that the Galaxy and the Supergalaxy are somehow identically involved in the generation of ultrahigh-energy PCRs.

4. THE PCR ANISOTROPY

Let us additionally consider some of the features of the lateral PCR distribution. Let us analyze the global PCR distribution on the celestial sphere in equatorial coordinates. To this end, we use the method of harmonic analysis [34], which basically consists in choosing the best values of the amplitude A_1 and phase φ_1 of the function

$$f(\alpha) = f_0(1 + A_1 \cos(\alpha - \varphi_1)) \quad (7)$$

by minimizing

$$X^2 = \sum_{i=1}^{n} \frac{(f_i - N_i)^2}{f_i}, \quad (8)$$

where N_i is the number of air showers in sector i $(\Delta\alpha)_i$. In this case,

$$f_0 = \frac{\sum_{i=1}^{n} N_i}{n} = \frac{N}{n}, \qquad A_1 = \frac{f_{\max} - f_{\min}}{f_{\max} + f_{\min}}. \quad (9)$$

We divided the sky into $n = 36$ spherical sectors. The anisotropy was studied in the $h = \Delta \log E_0 = 0.15$ bins by sequentially shifting then in energy by $0.2h$. This was done to analyze in more detail the behavior of the phase of the first harmonic φ_1, which characterizes the direction of the maximum of the global PCR flux.

Figure 4 shows the variations in phase φ_1 with energy bin position for the sample of 35 935 air showers included in the spectrum in Fig. 1. The relative scatter of points in the neighboring bins characterizes the errors in the phase. We see that most of the phases cluster near $\varphi_1 \approx 40°$ (dashed line *1*), $\varphi_1 \approx 120°$ (dashed line *2*), and $\varphi_1 \approx 300°$ (dashed line *3*). The first of them points to the place of intersection of the Galactic and Supergalactic planes (Fig. 5) where the data from [23, 26] reveal an absolute maximum of the PCR flux at $E_0 \geq 8 \times 10^{18}$ eV. The second concentrates near $\varphi_1 \approx 120°$. This direction is in no way associated with the Galactic plane from where one might seemingly expect an anisotropy in the flux of charged PCR particles within the framework of the diffusion model of their propagation. It points to the Supergalactic plane (Fig. 5). Finally, the third direction with $\varphi_1 \approx 300°$ points directly to the Galactic disk. It coincides with the direction of the anisotropy detected in [28, 29] at $E_0 \approx (5–20) \times 10^{17}$ eV.

The results presented in Fig. 4 suggest that the global PCR flux vector in individual energy bins changes significantly. We believe [14–23] that a substantial fraction of the cosmic rays with $E_0 \geq 10^{17}$ eV is extragalactic in origin; therefore, it manifests itself in both the Galaxy and the Supergalaxy. This is probably embodied in the irregularities of the spectra in Figs. 1–3.

4.1. Events with $E_0 \geq 8 \times 10^{18}$ eV

To analyze the results presented in Fig. 4 in more detail, let us first consider the anisotropy in the arrival directions of giant air showers (GASs) with $E_0 \geq 8 \times 10^{18}$ eV. For our analysis, we took all of the events with zenith angles $\theta \leq 60°$ whose arrival directions were determined from at least four stations and whose axes were within the array perimeter. A total of 559 air showers were selected in this way.

In addition, we used 522 events from the catalog [35] recorded by SUGAR. These air showers belong mostly to the Earth's Southern Hemisphere. Together with the

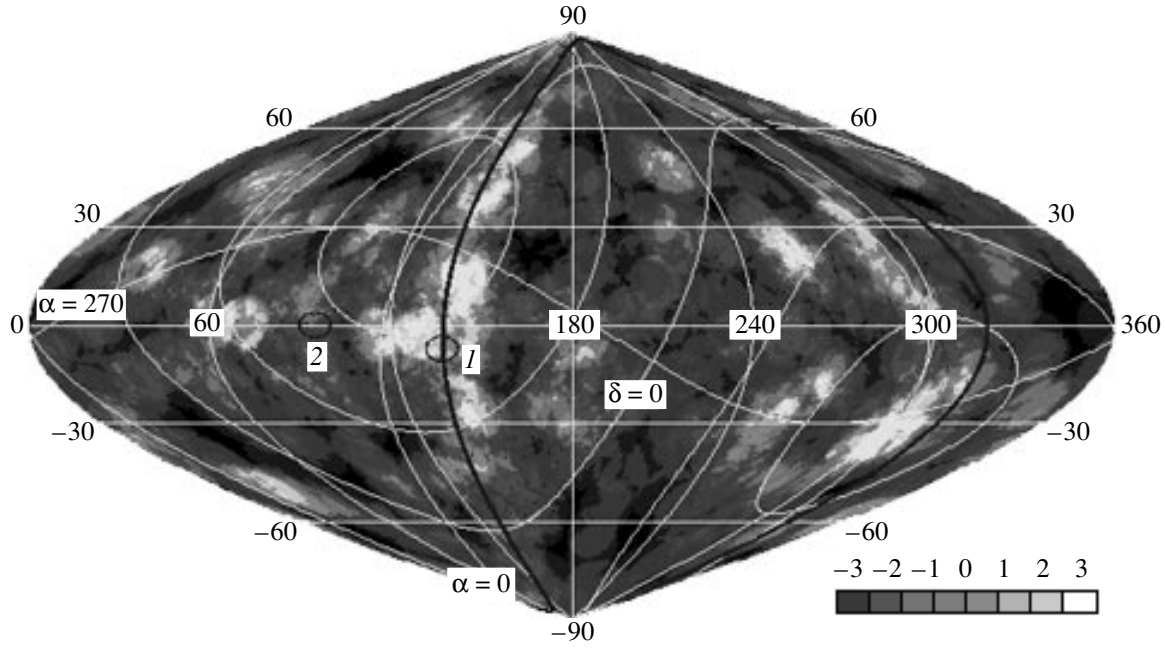

Fig. 5. Deviation of the observed number of air showers N_1 from the expected mean number $\langle N\rangle$ in units of $n_\sigma = (N_1 - \langle N\rangle)/\sqrt{\langle N\rangle}$ on the developed celestial sphere in Galactic coordinates for PCRs with $E_0 \geq 8 \times 10^{18}$ eV and $\theta \leq 60°$, as derived from the Yakutsk EAS and SUGAR data [35]: circles *1* and *2* indicate the pole of the local PCR excess found in [26] and the exit of the Local Galactic Arm, respectively; the dark curve represents the Supergalactic plane; the grayscale indicates the n_σ range.

Yakutsk array data, they give a fairly complete picture of the GAS anisotropy in the surrounding space. The accuracy with which the directions of the air shower axes were determined in [35] is about 5°.

We analyzed the deviations of the observed number of events N_1 from the expected mean number $\langle N\rangle = N_2(\Omega_1/\Omega_2)$ in units of the standard $\sigma = \sqrt{\langle N\rangle}$:

$$n_\sigma = \frac{N_1 - \langle N\rangle}{\sigma}, \qquad (10)$$

where N_1 and N_2 are the numbers of air showers in the solid angles

$$\Omega_1 = 2\pi(1 - \cos\theta_1)$$

and

$$\Omega_2 = 2\pi(1 - \cos\theta_2),$$

respectively ($\theta_1 = 8°$, $\theta_2 = 45°$). The values of n_σ (10) were found when a $1° \times 1°$ area was sequentially displaced over the entire sphere.

Figure 5 shows a map of the distribution of (10) over the developed celestial sphere in Galactic coordinates. For the convenience of perception, the equatorial coordinates are also shown here. Circles *1* and *2* mark the pole of the local PCR excess found in [26] and the exit of the Local Galactic Arm, respectively. The n_σ range is shown at the bottom of the figure in the form of a grayscale. The darkest and lightest regions correspond to deviations of the GAS flux from the mean by $|n_\sigma| \geq 3\sigma$.

Figure 5 shows interesting and important results. First, numerous local regions with relatively high and low GAS fluxes are seen here over the entire sphere. This suggests that the GAS anisotropy is multipolar in pattern. Second, almost no excess flux is observed in the Galactic disk if we do not take into consideration the place of intersection of the Galactic and Supergalactic planes at $l_G \approx 137.4°$. There is no slightest hint at an excess flux even from the Galactic center, where the most active and powerful matter conversion processes take place, although a statistically significant anisotropy is observed in this region at $E_0 \approx (5\text{–}20) \times 10^{17}$ eV [28, 29]. There is no excess GAS flux at the exit of the Local Galactic Arm (circle *2*) either. This probably suggests that the Galaxy plays a minor role in the generation of particles with $E_0 \geq 8 \times 10^{18}$ eV.

However, a completely different picture opens up in the Supergalaxy. A correlation between the GAS arrival directions and the Supergalactic plane in the Earth's Northern Hemisphere is clearly seen in Fig. 5. This correlation is characterized by the mean value

$$\langle n_\sigma\rangle = \frac{\sum_{i=1}^{k} (n_\sigma)_1}{k}. \qquad (11)$$

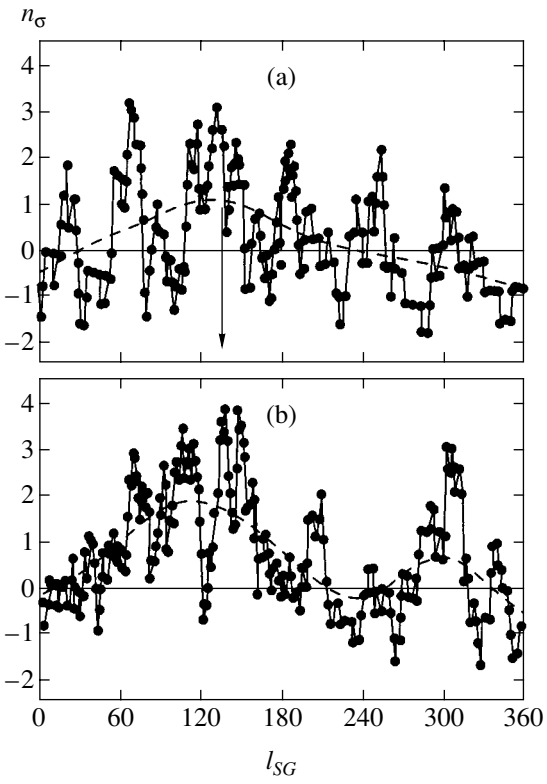

Fig. 6. Variations in n_σ presented in Fig. 5 when scanning along the (a) Galactic and (b) Supergalactic disks; the dashed curves represent the behavior on average when smoothed over 30 neighboring points.

Figure 6b shows the variations in (11) when the scanning solid angles Ω_1 and Ω_2 were sequentially displaced along a latitude of $b_{SG} \approx -3°$. The dashed curve reflects the behavior of this distribution (the dependence on Supergalactic longitude), on average, when smoothed over 30 neighboring points. An intense peak in the sector $\Delta l_{SG} \approx 0°$–$210°$, with its maximum at $l_{SG} \approx 120°$, is also seen here. However, there is no such clear correlation in the Southern Hemisphere, except the GAS excess in the local region of the sky with equatorial coordinates $\alpha \approx 0°$ and $\delta \approx -85°$ (Fig. 5). Curiously enough, it lies near the Supergalactic plane in a direction that is almost opposite to that of the region with the largest GAS excess in the Northern Hemisphere. In [23], we showed that this may not be a coincidence, but could be related to a certain orientation of the Supergalactic disk in space, with its edge directly against the flux of extragalactic particles generated by quasars.

For comparison, Fig. 6a also shows the variations in (11) along the Galactic disk (in the latitude band $|b_G| \le 8°$). We see that the averaged distribution (dashed curve) has a maximum at $l_G \approx 137°$, where the Supergalactic plane passes (see Fig. 5). Its height ($\approx 1\sigma$) is approximately smaller by a factor of 2 than that of the maximum in Fig. 6b at $l_{SG} \approx 120°$, suggesting that the Supergalaxy plays a dominant role in forming the excess GAS flux in the region of intersection of the Galactic and Supergalactic planes.

4.2. Events with $E_0 \approx (5–12) \times 10^{17}$ eV

Let us now consider in more detail the events with $E_0 \approx (5–12) \times 10^{17}$ eV that have attracted increasing attention since the reports [28, 29] appeared, especially since we have observed anomalous changes in the shape of the spectrum in Figs. 2a, 2b, and 3 in this energy range.

Figure 7 shows a map (similar to Fig. 5) of the distribution of (10) in Galactic coordinates for events with the above energies. It characterizes the distribution of local PCR fluxes in the part of the sky surveyed by the Yakutsk array. This sample contains 13 407 air showers with $\cos\theta \ge 0.7$ included in the spectrum in Fig. 1. The darkest and lightest areas reflect the deviations n_σ of the number of events N_1 by 3σ to the higher and lower values, respectively, in the scanning cone with

$$\Omega_1 = 2\pi(1 - \cos 8°)$$

from the expected mean

$$\langle N \rangle = N_2(\Omega_1/\Omega_2)$$

in units of the standard $\sigma = \sqrt{\langle N \rangle}$ for the number of events N_2 in the reference solid angle

$$\Omega_2 = 2\pi(1 - \cos 20°)$$

when a $1° \times 1°$ area is sequentially displaced over the entire portion of the sky under study.

There are numerous local extrema with enhanced and reduced cosmic-ray fluxes in Fig. 7. They form a certain pattern that roughly resembles the multipolar anisotropy in Fig. 5. Light rings with a diameter of about 30° are seen at some locations in Fig. 7.

The global anisotropy in the arrival directions of PCRs with $E_0 \approx (5–12) \times 10^{17}$ eV is characterized by Fig. 8, which shows the distribution of (11) in the intervals of angles $\Delta\alpha = 1.5°$. The dashed curve reflects the behavior on average when smoothed over 70 neighboring points. The maximum of this distribution at $\alpha \approx 300°$ matches the result of our harmonic analysis in Fig. 4.

In Fig. 9, (11) is plotted against latitude in Galactic (filled circles) and Supergalactic (open circles) coordinates (in the intervals of angles $\Delta b = 1.5°$). Here, a latitudinal gradient (dashed lines) is clearly seen in both coordinate systems. The results in Fig. 9 (filled circles) completely confirm the north–south asymmetry in the global particle flux that is present in the spectrum in Figs. 2a and 2b in the form of a peak at $3 \times 10^{17} \le E_0 \le 2 \times 10^{18}$ eV. If we bear in mind that the North Galactic

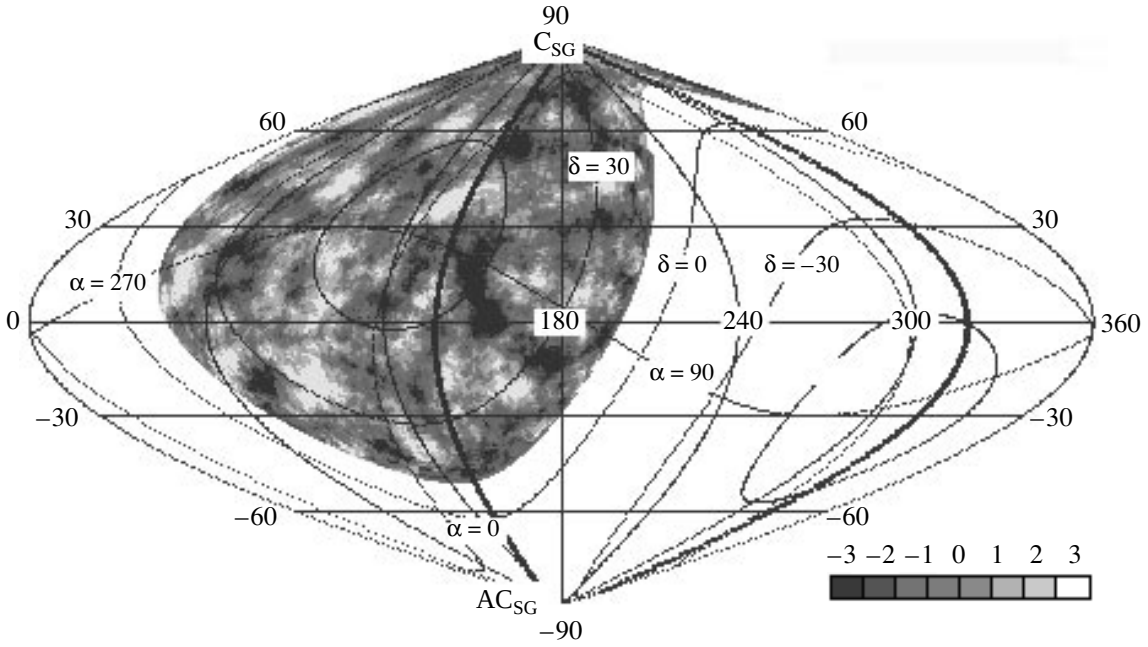

Fig. 7. Deviations of the observed number of air showers N_1 from their expected mean number $\langle N\rangle$ in units of $n_\sigma = (N_1 - \langle N\rangle)/\sqrt{\langle N\rangle}$ on the developed celestial sphere in Galactic coordinates for PCRs with $E_0 \approx (5\text{–}12) \times 10^{17}$ eV and $\theta \le 45°$: the dark curve, C_{SG}, and AC_{SG} are the Supergalactic plane, center, and anticenter, respectively; the grayscale indicates the n_σ range.

Pole points toward the Supergalactic center (C_{SG} in Fig. 7), then, in this case, we cannot exclude the contribution of the latter to the observed gradient. It can manifest itself in stronger absorption of the extragalactic PCR flux passing through the central region of the Supergalaxy than that in other directions.

The distribution in Fig. 9 (open circles) agrees with the spectrum in Fig. 3b. It confirms the existence of a certain deficit of particles in the latitude band $-8° < b_{SG} < 2°$. The north–south asymmetry in these coordinates can be interpreted as evidence for an enhanced PCR flux from the central region of the Galaxy, since the North Pole of the Supergalaxy has Galactic coordinates $l_G \approx 47.4°$ and $b_G \approx 3°$.

The role of the disks of the structures of surrounding space that we considered is reflected in Fig. 10, in which (11) is plotted against Galactic (filled circles) and Supergalactic (open circles) longitudes (in the bands $|b_G| \le 5°$ and $-8° < b_{SG} < 2°$ represented in the spectra in Fig. 3). Here, we also see the gradients

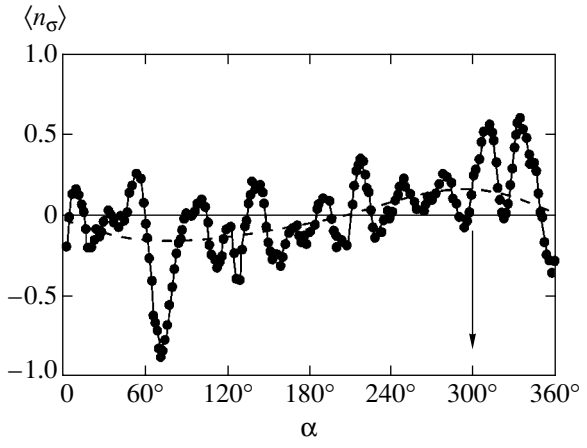

Fig. 8. Distribution of (11) in right ascension (in intervals $\Delta\alpha = 1.5°$) for the sample of events in Fig. 7: the dashed curve represents the behavior on average when smoothed over 70 neighboring points.

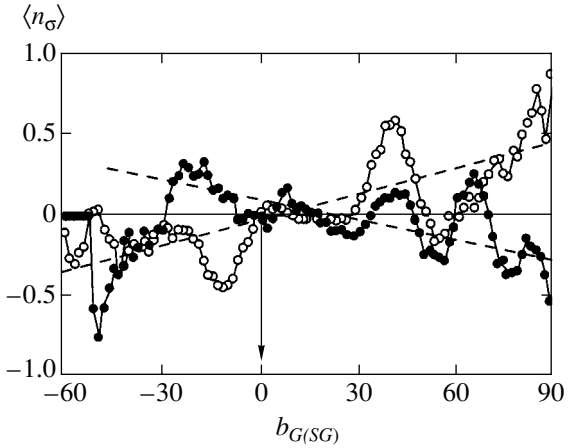

Fig. 9. Distribution of (11) in latitude (in intervals $\Delta b = 1.5°$) for the sample of events in Fig. 7 in Galactic (●) and Supergalactic (○) coordinates: the dashed curves represent the behavior on average.

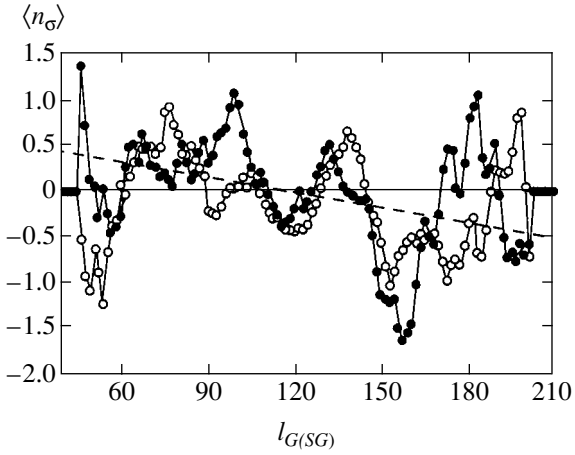

Fig. 10. Dependences of (11) on Galactic (●) and Supergalactic (○) longitudes in the latitude intervals $|b_G| \le 5°$ and $-8° < b_{SG} < 2°$ for the sample of air showers with $E_0 \approx (5–12) \times 10^{17}$ eV and $\theta \le 45°$ presented in Fig. 3: the dashed line represents the behavior on average.

(dashed line) that point to enhanced particle fluxes from the Galactic center and the Supergalactic anticenter.

Note the local peaks and dips in Figs. 8–10 that alternate approximately after 30°. We believe that they did not arise accidentally, but reflect the (somehow) ordered structure of the distribution of PCR sources in Metagalactic space. Quasars could be among the sources of the above energies [18–23]. In [23], we showed that quasars form a certain structure with the same characteristic angular irregularities approximately equal to 30°.

5. CONCLUSIONS

If we attempt to summarize the results obtained above, then the following picture is possible. We have already repeatedly reported previously [14–23] that the PCR flux with $E_0 \ge 10^{17}$ eV probably consists of two components. One of these components is extragalactic in origin and can be generated by quasars [18–23]. Quasars are among the most powerful sources in the Universe. They are located at cosmological distances that significantly exceed the sizes of the Supergalaxy, not to mention the Galaxy. On their way to Earth, the ultrahigh-energy particles pass through these structures. Some of them can be assumed to enter into nuclear reactions with gas. The gas is concentrated most densely in the Galactic and Supergalactic disks, in regions with angular sizes in latitude $|b| \approx 5°–10°$ [36]. The excess PCR fluxes at $E_0 \ge 5 \times 10^{18}$ eV in Figs. 2 and 3 and the anisotropy in Figs. 4–6 are probably attributable to this factor.

In the region of apparent intersection of the Galactic and Supergalactic planes (at $\alpha \approx 40.6°$ and $\delta \approx 59.5°$), the fluxes probably add up to give a local extremum with the absolute maximum found in [26].

As regards energies of $(3–5) \times 10^{17} \le E_0 \le 2 \times 10^{18}$ eV, an appreciable fraction of the extragalactic PCR flux is apparently also present here. This has already been reported previously [20–23]. The intensity decrease in Fig. 3b compared to the background spectrum and the results presented in Figs. 9 and 10 (open circles) can be interpreted as the absorption of the extragalactic particles interacting with the Supergalactic matter. The Galaxy may play a similar role of the target in this energy range. It may well be that these two structures act as targets for extragalactic ultrahigh-energy particles generated by quasars and other galaxies with active nuclei.

Different particles can be assumed to be produced in the nuclear reactions that proceed in the Supergalactic and Galactic gas at $E_0 \le 2 \times 10^{18}$ eV and $E_0 \ge 5 \times 10^{18}$ eV. They reach Earth in the form of cosmic rays producing EASs with distinctly different characteristics, as suggested by our experimental data [37–41]. In these papers, we showed that the experimental data at $E_0 \ge (3–5) \times 10^{18}$ eV completely disagree with the existing theoretical models of EAS development from protons or the nuclei of any other chemical elements.

The results presented above shed additional light on the origin of ultrahigh-energy PCRs as a whole. They point to the importance of taking into account the locations of the sky regions when comparing the experimental data. We plan to continue these studies.

ACKNOWLEDGMENTS

This work was performed under financial support of the Yakutsk EAS array by the Ministry of Science of Russia (reg. no. 01-30), which was included in the "List of Unique Research and Experimental Facilities of National Significance" and was supported by the Russian Foundation for Basic Research, project no. 05-02-17857a.

REFERENCES

1. T. A. Egorov, N. N. Efimov, V. A. Kolosov, *et al.*, in *Proceedings of 12th ICRC* (Hobart, 1971), Vol. 6, p. 2059.
2. A. V. Glushkov, O. S. Diminshteĭn, N. N. Efimov, *et al.*, in *Characteristics of Extensive Atmospheric Showers of Highest Energy Cosmic Rays* (Yakutsk, 1976) [in Russian].
3. A. V. Glushkov, V. M. Grigoryev, M. N. Dyakonov, *et al.*, in *Proceedings of 20th ICRC* (Moscow, 1987), Vol. 5, p. 494.
4. M. I. Pravdin, M. N. Dyakonov, A. V. Glushkov, *et al.*, in *Proceedings of 26th ICRC* (Salt Lake City, 1999), Vol. 3, p. 292.
5. A. V. Glushkov, V. P. Egorova, A. A. Ivanov, *et al.*, in *Proceedings of 28th ICRC* (Tsukuba, 2003), Vol. 1, p. 389.
6. G. T. Zatsepin and V. A. Kuzmin, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 114 (1966) [JETP Lett. **4**, 99 (1966)].
7. K. Gresen, Phys. Rev. Lett. **16**, 748 (1966).

8. N. Sakaki, M. Chikawa, M. Fukushima, *et al.*, in *Proceedings of 27th ICRC* (Hamburg, 2001), Vol. 1, p. 333.
9. T. Abu-Zayyad *et al.*, astro-ph/0208301.
10. V. S. Berezinsky, A. Z. Gazizov, and S. I. Grigorieva, hep-ph/0204357.
11. D. Marco, P. Blasi, and A. V. Olinto, Astropart. Phys. **20**, 53 (2003).
12. A. V. Glushkov, Pis'ma Zh. Éksp. Teor. Fiz. **78**, 1265 (2003) [JETP Lett. **78**, 745 (2003)].
13. M. I. Pravdin, A. V. Glushkov, V. P. Egorova, *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **68**, 1621 (2004).
14. A. V. Glushkov and M. I. Pravdin, Zh. Éksp. Teor. Fiz. **119**, 1029 (2001) [JETP **92**, 887 (2001)].
15. A. V. Glushkov and M. I. Pravdin, Pis'ma Astron. Zh. **27**, 577 (2001) [Astron. Lett. **27**, 493 (2001)].
16. A. V. Glushkov and I. E. Sleptsov, Izv. Ross. Akad. Nauk, Ser. Fiz. **65**, 437 (2001).
17. A. V. Glushkov, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 355 (2001) [JETP Lett. **73**, 313 (2001)].
18. A. V. Glushkov, Izv. Ross. Akad. Nauk, Ser. Fiz. **66**, 1599 (2002).
19. A. V. Glushkov and M. I. Pravdin, Pis'ma Astron. Zh. **28**, 341 (2002) [Astron. Lett. **28**, 296 (2002)].
20. A. V. Glushkov and M. I. Pravdin, Yad. Fiz. **66**, 886 (2003) [Phys. At. Nucl. **66**, 854 (2003)].
21. A. V. Glushkov, Pis'ma Astron. Zh. **29**, 172 (2003) [Astron. Lett. **29**, 142 (2003)].
22. A. V. Glushkov, Yad. Fiz. **67**, 983 (2004) [Phys. At. Nucl. **67**, 961 (2004)].
23. A. V. Glushkov, Yad. Fiz. **68**, 262 (2005) [Phys. At. Nucl. **68**, 237 (2005)].
24. *Catalogue of Highest Energy Cosmic Rays* (World Data Center C2 for Cosmic Rays, Japan, 1988), No. 3.
25. A. V. Glushkov, M. I. Pravdin, I. E. Sleptsov, *et al.*, Yad. Fiz. **63**, 1557 (2000) [Phys. At. Nucl. **63**, 1477 (2000)].
26. A. A. Ivanov, A. D. Krasilnikov, and M. I. Pravdin, in *Proceedings of 28th ICRC* (Tsukuba, 2003), Vol. 1, p. 341.
27. B. N. Afanasiev, M. N. Dyakonow, T. A. Egorov, *et al.*, in *Proceedings of 24th ICRC* (Rome, 1995), Vol. 2, p. 756.
28. N. Hayashida, M. Nagano, D. Nishikawa, *et al.*, Astropart. Phys. **10**, 303 (1999).
29. J. A. Bellido, B. W. Clay, R. B. Dawson, and M. Johnston-Hollitt, astro-ph/0009039.
30. P. L. Biermann, G. M. Tanko, R. Engel, and G. Pugliese, Astrophys. J. **604**, L29 (2004).
31. J. Szabelsky, J. Wdowczyk, and A. W. Wolfendale, Phys. G **12**, 1433 (1986).
32. A. A. Mikhailov, in *Proceedings of 26th ICRC* (Salt Lake City, 1999), Vol. 3, p. 268.
33. A. A. Mikhaĭlov, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 233 (2000) [JETP Lett. **72**, 160 (2000)].
34. J. Linsley, Phys. Rev. Lett. **34**, 1530 (1975).
35. *Catalogue of Highest Energy Cosmic Rays* (World Data Center C2 for Cosmic Rays, Japan, 1986), No. 2.
36. *Physics of Space. Little Encyclopedia*, Ed. by R. A. Syunyaev (Sovetskaya Éntsiklopediya, Moscow, 1986) [in Russian].
37. A. V. Glushkov, I. T. Makarov, E. S. Nikiforova, *et al.*, Astropart. Phys. **4**, 1274 (1995).
38. A. V. Glushkov, I. T. Makarov, M. I. Pravdin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 145 (2000) [JETP Lett. **71**, 97 (2000)].
39. A. V. Glushkov, M. I. Pravdin, I. E. Sleptsov, *et al.*, Yad. Fiz. **63**, 1557 (2000) [Phys. At. Nucl. **63**, 1477 (2000)].
40. A. V. Glushkov, M. I. Pravdin, I. E. Sleptsov, *et al.*, Yad. Fiz. **65**, 1346 (2002) [Phys. At. Nucl. **65**, 1313 (2002)].
41. A. V. Glushkov, L. G. Dedenko, M. I. Pravdin, and I. E. Sleptsov, Zh. Éksp. Teor. Fiz. **126**, 5 (2004) [JETP **99**, 1 (2004)].

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 101, No. 1, 2005, pp. 98–106.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 128, No. 1, 2005, pp. 115–124.

ORDER, DISORDER, AND PHASE TRANSITIONS IN CONDENSED SYSTEMS

Electron Transport, Penetration Depth, and the Upper Critical Magnetic Field in ZrB_{12} and MgB_2¶

V. A. Gasparov, N. S. Sidorov, I. I. Zver'kova, S. S. Khassanov, and M. P. Kulakov

Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia

e-mail: vgasparo@issp.ac.ru

Received November 26, 2004

Abstract—We report on the synthesis and measurements of the temperature dependences of the resistivity ρ, the penetration depth λ, and the upper critical magnetic field H_{c2}, for polycrystalline samples of dodecaboride ZrB_{12} and diboride MgB_2. We conclude that ZrB_{12} behaves as a simple metal in the normal state with the usual Bloch-Grüneisen temperature dependence of ρ(T) and with a rather low resistive Debye temperature T_R = 280 K (to be compared to T_R = 900 K for MgB_2). The ρ(T) and λ(T) dependences for these samples reveal a superconducting transition in ZrB_{12} at T_c = 6.0 K. Although a clear exponential λ(T) dependence in MgB_2 thin films and ceramic pellets was observed at low temperatures, this dependence was almost linear for ZrB_{12} below T_c/2. These features indicate an s-wave pairing state in MgB_2, whereas a *d*-wave pairing state is possible in ZrB_{12}. In disagreement with conventional theories, we found a linear temperature dependence, of $H_{c2}(T)$ for ZrB_{12} ($H_{c2}(0)$ = 0.15 T). © *2005 Pleiades Publishing, Inc.*

1. INTRODUCTION

The recent discovery of superconductivity at 39 K in magnesium diboride [1] has initiated a booming activity in condensed matter physics. This activity has raised considerable interest in the search for superconductivity in other borides [2]. Unfortunately, none of natural candidate MeB_2-type diborides of light metals (Me = Li, Be, Al, Ca) nor any of a large number of the known isostructural transition metal diborides (Me = Ti, Zr, Hf, V, Ta, Cr, Mo, U) have been found to be superconducting [2]. Only in nonstoichiometric compounds ($MoB_{2.5}$, $NbB_{2.5}$, Mo_2B, W_2B, $BeB_{2.75}$) was superconductivity observed [3–6]. We note that the earlier speculation about superconductivity in TaB_2 [7] (in contradiction to other published data [2]) has been disproved by recent resistivity, susceptibility, and specific heat measurements supported by electronic structure calculations [8].

These results do not seem to support the application of the old idea about superconductivity in metallic hydrogen [9] to the explanation of superconductivity in MgB_2 [10]. In spite of this fact, we discuss some aspects of this idea. In particular, it is believed that in MgB_2, the averaged phonon frequencies (in other words, the Debye temperature) must be very high due to the low mass of the boron, which sharply increases the prefactor in the McMillan formula for T_c. Indeed, the band structure calculations have shown that electrons at the Fermi level are predominantly boron-like in MgB_2 and the superconductivity is due to graphite-type "metallic" boron sheets [10]. Furthermore, Eremets *et al.* [11] recently observed that the semiconducting polycrystalline boron (rhombohedral β-B_{105}) transforms to a metal under high pressure and even to a superconductor at about 160 GPa. The critical temperature T_c increases from 6 to 11.2 K at raised pressure up to 250 GPa. This observation supports the old idea that a route for optimizing T_c is in preparation of boron-rich compounds, even though this does not yet work for known borides.

In fact, the search for superconductivity in borides has a long history. Matthias *et al.* [12] discovered several superconducting cubic hexa and dodecaborides (MeB_6 and MeB_{12}) in the 1960's. Many other hexa-and dodecaborides (Me = Ce, Pr, Nd, Eu, Gd, Tb, Dy, Ho, Er, Tm) were found to be ferromagnetic or antiferromagnetic. It was suggested that the superconductivity in YB_6 and ZrB_{12} (having the highest T_c values 6.5–7.1 and 6.03 K, respectively [3]) might be due to the hypothetical cubic metallic boron. However, a much smaller isotope effect on T_c for boron in comparison with the Zr isotopic substitution suggests that the boron in ZrB_{12} serves as an inert background, and Zr is actually crucial for superconductivity [13, 14], even though chemically, ZrB_{12} mainly contains boron.

¶ This article was submitted by the authors in English.

While the superconductivity in ZrB_{12} was discovered a long time ago [12], there has been little effort devoted to the study of basic superconductive properties of this dodecaboride. Only recently were the electron transport of solid solutions $Zr_{1-x}Sc_xB_{12}$ [15] and the band structure calculations for ZrB_{12} [16] reported. Understanding the electron transport properties of the cluster borides and the superconductivity mechanism in these compounds is very important. In this paper, we attempt to address this problem. We report the temperature-dependent resistivity $\rho(T)$, magnetic field penetration depth $\lambda(T)$, and upper critical magnetic field $H_{c2}(T)$ for polycrystalline samples of ZrB_{12}. Comparative data on $\rho(T)$ and $\lambda(T)$ in MgB_2 thin films and pellets are also presented.

The structure of this paper is as follows. In Section 2, we report on synthesis of ZrB_{12} and MgB_2 and the experimental techniques. Section 3 describes the electron transport in these compounds. Section 4 describes the temperature dependence of λ in thin films and polycrystalline samples. The data on $H_{c2}(T)$ are presented in Section 5.

2. EXPERIMENTAL

Under ambient conditions, dodecaboride ZrB_{12} crystallizes in the fcc structure (Fig. 1) of the UB_{12} type (space group $Fm\bar{3}m$), with the lattice parameter a = 0.74075 nm [17]. In this structure, the Zr atoms are located at interstitial openings in close-packed B_{12} clusters [15]. In contrast, the diborides show a phase consisting of two-dimensional graphite-like monolayers of boron atoms with a honeycomb lattice, intercalated with metal monolayers [2]. In our search for the superconducting diboride compounds, we observed superconductivity at 5.5 K in ZrB_2 polycrystalline samples that had a few percent amount of ZrB_{12} impurity [2]. It was recently suggested [18] that this observation could be associated with nonstoichiometry in the zirconium sublattice of ZrB_2. To resolve this issue and to study the electron transport and basic superconducting properties of ZrB_{12}, we successfully synthesized this compound.

Polycrystalline samples of ZrB_{12} were obtained by the conventional solid-state reaction. The starting materials were a zirconium metal powder (99.99% purity) and a submicron amorphous boron powder (99.9% purity). These materials were lightly mixed in appropriate amounts and pressed into pellets 10 mm thick and 20 mm in diameter. The pellets were wrapped in a tungsten foil and baked at 2000°C by electron-beam heating with subsequent slow cooling to room temperature. The process took place for two hours in a high-vacuum chamber at 2×10^{-4} Pa. The resulting polycrystalline pellets had over 90% of the theoretical mass density and were black in color. They demonstrated good metallic conductivity at low temperatures. After regrinding the prepared pellets in an agate mortar, the respective powders were reheated few times for two hours.

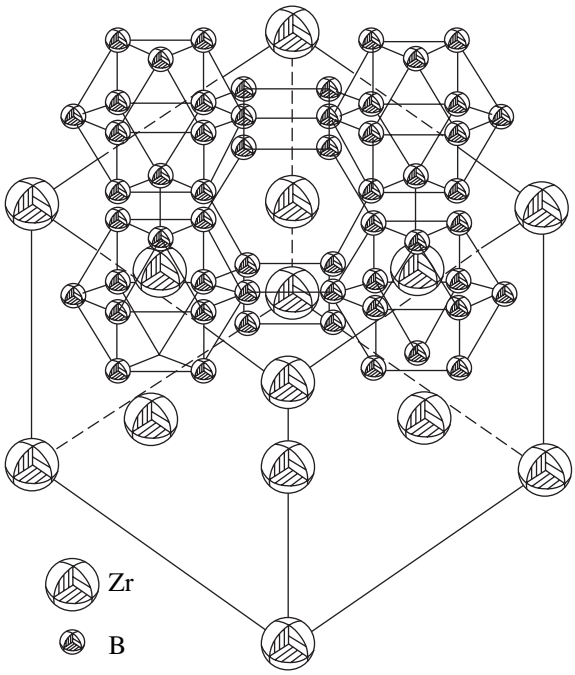

Fig. 1. The lattice structure of dodecaboride ZrB_{12}. For clarity, only B_{12} clusters on the upper face of the lattice are shown.

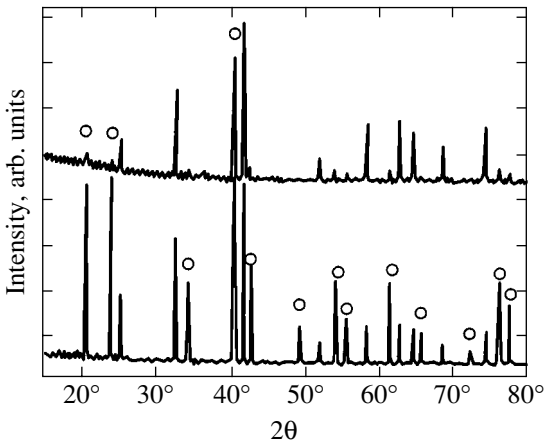

Fig. 2. A typical X-ray $\theta - 2\theta$ scan of ZrB_{12} powders (the lower curve) at room temperature. A similar scan for ZrB_2 pellets studied before [2] is presented by the upper curve. The cycles mark the X-ray reflections from fcc ZrB_{12}.

The powder X-ray diffraction pattern obtained using CuK_α radiation showed that the samples largely consist of the desired ZrB_{12} phase (Fig. 2). Nevertheless, small amounts of ZrB_2 were found to be present and could not be eliminated by subsequent regrinding and annealing. A Rietveld refinement of the ZrB_{12} X-ray pattern, based

on the fcc UB_{12}-type structure presented in Fig. 1, yielded the lattice parameter $a = 0.7388$ nm to be very close to the published values [17]. The polycrystalline MgB_2 pellets have been sintered using a similar technique as outlined in our earlier work [2]. This technique is based on the reactive liquid Mg infiltration of boron powder. X-ray diffraction analysis did not reveal the metallic Mg phase, and we therefore believe that a rather high resistivity ratio ($R_{300\,K}/R_{40\,K} = 12$) is due to the high quality of our samples.

For this study, two highly crystalline, superconducting films of MgB_2 were grown on an r-plane sapphire substrate in a two-step process. Deposition of B precursor films via electron-beam evaporation was followed by *ex-situ* post annealing at 890°C in the presence of bulk MgB_2 and Mg vapor. Scanning electron microscopy showed dense films with surface roughness below 5 nm. For the measurements, we investigate films of 500 and 700 nm thick, with the corresponding T_c values 38 and 39 K. The details of the preparation technique are described elsewhere [19].

For the resistance measurements, we used the spark erosion method to cut the pellets into a rectangular bars with dimensions of about $0.5 \times 0.5 \times 8$ mm^3. The samples were lapped with a diamond paste. To remove any deteriorated surface layers, the samples were etched: ZrB_{12} in hot nitrogen acid and MgB_2 in 2% HCl plus water-free ethanol. A standard four-probe ac (9 Hz) method was used for resistance measurements. Electrical contacts were made with Epotek H20E silver epoxy. The temperature was measured with platinum (PT-103) and carbon glass (CGR-1-500) sensors. A well-defined geometry of the samples provided for the precise resistivity measurements.

The measurements were performed in the liquid-helium variable-temperature cryostat in the temperature range between 1.1 and 350 K. Magnetic measurements of the resistivity and the penetration depth on the polycrystalline samples were carried out using a superconducting coil in applied fields of up to 6 T. The dc magnetic field was applied in the direction of the current flow. The critical temperature measured by the radio-frequency (RF) susceptibility [2] and $\rho(T)$ was found to be $T_{c0} = 6.0$ K for ZrB_{12} samples and 39.0 K for MgB_2 samples.

The $\lambda(T)$ dependence in thin films was investigated using a single-coil mutual inductance technique. This technique, originally proposed in [20] and improved in [21], takes advantage of the well-known two-coil geometry. It was successfully used for the observation of the Berezinskii–Kosterlitz–Thouless vortex–antivortex unbinding transition in ultrathin $YBa_2Cu_3O_{7-x}$, films [22] as well as for the study of the $\lambda(T)$ dependence for MgB_2 films [23]. In particular, this RF technique measures the change of inductance ΔL of a one-layer pancake coil located in the proximity of the sample. The coil is a part of the *LC* circuit driven by a marginal oscillator operating at 2–10 MHz or by the impedance meter (VM-508 TESLA 2–50 MHz). The frequency stability of this oscillator is 10 Hz. The film is placed at a small distance (about 0.1 mm) below the coil and is thermally insulated from the coil by a Teflon foil. Both the sample and the coil are in vacuum, but the coil holder is thermally connected with a helium bath, while the sample holder is isolated and may be heated. During the experiment, the coil was kept at 2.5 K, whereas the sample temperature was varied from 2.5 to 100 K. Such a design allows us to eliminate possible effects in temperature changes in L and C on the measurements. The real part of the complex mutual inductance M between the film and the coil can be obtained as

$$\mathrm{Re}M(T) = L_0\left(\frac{f_0^2}{f^2(T)} - 1\right), \qquad (1)$$

where L_0 and f_0 are the inductance and the resonant frequency of the circuit without the sample. In the London regime, where high-frequency losses are negligible, one can introduce the difference between the temperature-dependent real part of M of the coil with the sample, $\mathrm{Re}M(T)$, and that of the coil at T_0, $\mathrm{Re}M_0$. This difference is a function of $\lambda(T)$,

$$\Delta\mathrm{Re}M(T) = \pi\mu_0\int_0^\infty \frac{M(q)}{1 + 2q\lambda\coth(d/\lambda)}dq, \qquad (2)$$

where μ_0 is the magnetic permeability of vacuum, $M(q)$ plays the role of the mutual inductance at a given wavenumber q in the film plane and depends on the distance between the sample and the coil, d is the sample thickness and μ_0 is the permeability of free space (additional details can be found in [21]). A change in $\Delta\mathrm{Re}M(T)$ is detected as a change of the resonant frequency $f(T)$ of the oscillating signal. When inserted in Eq. (2), this change yields a temperature-dependent London penetration depth $\lambda(T)$.

Measurements of $\lambda(T)$ for polycrystalline ZrB_{12} and MgB_2 samples were performed with a similar *LC* technique but using a rectangular solenoid coil into which the sample was placed. The details of this technique are described elsewhere [24]. For such arrangements, changes in the resonant frequency $f(T) = \omega/2\pi$ of the circuit relative to that above T_c, $f(T_c)$, and at the minimal temperature T_1, $f(T_1)$, are directly related to the inductance of the probe coil and, hence, to $\lambda(T)$ by

$$\lambda(T) - \lambda(T_0) = \delta\frac{f^{-2}(T) - f^{-2}(T_1)}{f^{-2}(T_c) - f^{-2}(T_1)}. \qquad (3)$$

Here, $f(T_c)$ and $f(T_1)$) are the respective resonant frequencies at $T > T_c$ and at the minimal temperature T_1

and $\delta = \sqrt{c^2\rho/2\pi\omega}$ is the skin depth above T_c, which was determined from the resistivity $\rho(T)$ measurements.

3. ELECTRON TRANSPORT

Figure 3 shows the temperature dependence of the resistivity for ZrB_{12} and MgB_2 samples. The inset displays the variation of $\rho(T)$ near the superconducting transition with zero resistance at 6.0 K (the width ΔT = 0.04 K) in ZrB_{12} and at 39 K (ΔT = 0.7 K) in MgB_2 samples. The transition is remarkably narrow for ZrB_{12} samples, which is a clear indication of good-quality samples. The transition temperature is consistent with the previously reported values for ZrB_{12} (6.03 K) [12–14] and is comparably larger than that of ZrB_2 samples (5.5 K) [2]. Although ZrB_{12} mostly contains boron, its room-temperature resistivity is only four times larger than that of MgB_2 and ZrB_2 [2], while the residual resistivity is ten times larger. The resistivity ratio for ZrB_{12} ($\rho(300\text{ K})/\rho(6\text{ K}) \approx 4$) is rather low compared to the single-crystal value 10 [25]. Although an X-ray diffraction analysis revealed small amounts of ZrB_2 (the nonsuperconducting phase [25]), apparently there is no influence of this phase on the $\rho(T)$ dependence, because below T_c the resistivity drops to zero, rather than to the residual value.

One can predict a nearly isotropic resistivity for fcc ZrB_{12}, which can be described by the Bloch–Grüneisen (BG) expression of the electron–phonon (*e–p*) scattering rate [26]:

$$\rho(t) - \rho(0) = 4\rho_1 t^5 \int_0^{1/t} \frac{x^5 e^x dx}{(e^x - 1)^2} = 4\rho_1 t^5 J_5(1/t). \quad (4)$$

Here, $\rho(0)$ is the residual resistivity, $\rho_1 = d\rho(T)/dt$ is the slope of $\rho(T)$ at high temperatures ($T > T_R$), $t = T/T_R$, and T_R is the resistive Debye temperature. As we can see from Fig. 3, the BG equation describes our data reasonably well, indicating the importance of the *e–p* interaction for both metals. The best fit to our data is obtained with T_R = 280 K for ZrB_{12} and T_R = 900 K for MgB_2.

In contrast to ZrB_{12}, the resistivity of MgB_2 samples does deviate from the BG model at low temperatures. This problem has been under consideration by several groups. In particular, Putti *et al.* [27] modified the BG equation by introducing a variable power *n* for the $t^n J_n(1/t)$ term in Eq. (4). The best fit to the data was obtained with $n = 3$, which in fact ignores a small-angle *e–p* scattering. Recently, Sologubenko *et al.* [28] reported a cubic *T*-dependence in the *a*, *b*-plane resistivity below 130 K in single crystals of MgB_2. This was attributed to the interband *e–p* scattering in transition metals.

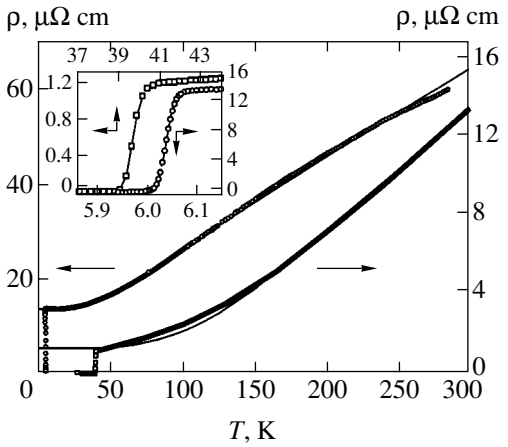

Fig. 3. Temperature dependence of the resistivity ρ for ZrB_{12} (open circles) and MgB_2 (squares) polycrystalline samples. The solid lines represent the Bloch-Grüneisen fits to the experimental data in accordance with Eq. (4).

We stress that there are strong objections to this modified BG model: (i) the cubic $\rho(T)$ dependence is a theoretical model for large-angle *e–p* scattering, and no evidence of it was observed in transition and non-transition metals; (ii) numerous studies of the $\rho(T)$ dependence in transition metals have been successfully described by a sum of the electron-electron (*e–e*), proportional to T^2, and *e–p*, proportional to T^5, contributions to the low-temperature resistivity, which may be easily confused with a T^3 law [26, 29, 30]; (iii) the σ–π inter-band *e–p* scattering plays no role in normal transport in the two-band model for MgB_2 [31].

To investigate whether a combination of *e–e* and *e–p* scattering works for our samples, we decided to add a T^2-term to Eq. (4) [29, 30]. We note that the BG term is proportional to T^5 at low temperatures. Therefore, addition of the T^2-term results in the following expression for the resistivity $\rho(T)$:

$$\frac{\rho(T) - \rho(0)}{T^2} = \alpha + \beta T^3. \quad (5)$$

Here, α and β are parameters of the respective *e–e* and *e–p* scattering terms. When plotted in the $[\rho(T) - \rho(0)]/T^2$ vs T^3 axes, such a dependence yields a straight line with the slope β and the *y*-intercept (T = 0) equal to α. The corresponding plot of our data in Fig. 4 clearly displays the expected linear dependences. The presence of an unusually large T^2-term in MgB_2 data (open squares in Fig. 4) below 150 K is evident (α = 150 pΩ cm/K^2), whereas the *e–p* scattering T^5-term is substantially smaller ($\beta = 2.1 \times 10^{-6}$ pΩ cm/K^5). We note, however, that the α value for MgB_2 is almost 40 times larger than the corresponding values in transition metals such as molybdenum and tungsten (α_{Mo} = 2.5 pΩ cm/K^2 and α_W = 1.5–4 pΩ cm/K^2

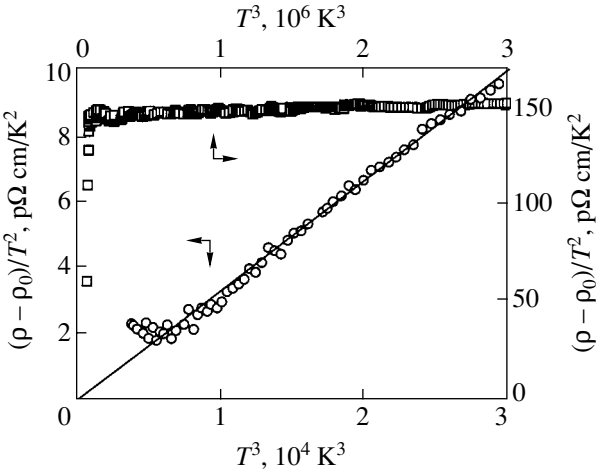

Fig. 4. Temperature dependence of the reduced resistivity $[\rho(T) - \rho(0)]/T^2$ for ZrB_{12} (open circles) and MgB_2 (squares) polycrystalline samples. The solid lines are a guide for the eye.

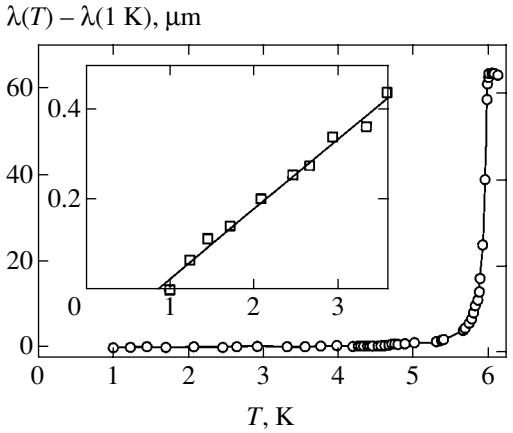

Fig. 5. Temperature dependence of the penetration depth for a ZrB_{12} sample. The solid lines are a guide for the eye. The inset shows the data below 3.5 K in an extended scale.

[29, 30]). In contrast, the ZrB_{12} data display a nearly zero T^2-term.

In general, there are many scattering processes responsible for the T^2-term in the $\rho(T)$ dependence of metals [25]. In particular, umklapp *e–e* scattering strongly contributes to this term. Furthermore, normal collisions are significant in compensated metals and in thermal resistivity [30]. Borides have a rather high T_R value, which depresses the *e–p* scattering, and hence the *e–e* term is easier to observe. Clearly, there is no obvious explanation for such a significant *e–e* scattering contribution in MgB_2. We believe that additional experiments on purer samples are necessary before the final conclusion about the origin of the T^2-term in the $\rho(T)$ dependence for MgB_2 can be drawn. Besides, the T^2-term was recently observed in ZrB_{12} single-crystal samples with the large resistivity ratio equal to 10 [25]. Apparently, the T^2-term depends on the residual resistivity.

4. PENETRATION DEPTH

Our *RF* technique allows us to measure the change in the penetration depth $\lambda(T)$ [24]. We note, however, that there is some uncertainty in determining the absolute values of $\lambda(T)$ for bulk samples because of error in the f_1 determination in Eq. (3). For this reason, we do not attempt to determine the absolute values of $\lambda(0)$ for polycrystalline samples from these data but rather find the temperature-dependent part $\Delta\lambda(T) = \lambda(T) - \lambda(1\ \mathrm{K})$. Figure 5 displays the effect of the superconducting transition in ZrB_{12} on $\lambda(T)$. The striking feature of the curves in Fig. 5 is the linear temperature dependence of $\Delta\lambda$ below $T_c/2 = 3$ K. We emphasize that no frequency dependence of these data was found when oscillator frequency was varied by a factor of two.

In the BCS theory, the London penetration depth $\lambda(T, l)$ is identical to the magnetic penetration depth $\lambda(T)$ in the case of specular and diffuse surface scattering and for negligible nonlocal effects, i.e., for $\delta(T, l) \gg \xi(T, l)$ [24, 32]. Here, l is the mean free path of carriers and ξ is the coherence length. In a BCS-type superconductor (with the conventional s-wave pairing) in the clean limit ($l \gg \xi$), the magnetic penetration depth has an exponentially vanishing temperature dependence below $T_c/2$ (where $\Delta(T)$ is almost constant) [32]:

$$\lambda(T) = \lambda(0)\left[1 + \sqrt{\pi\frac{\Delta(0)}{2k_BT}}\exp\left(-\frac{\Delta(0)}{k_BT}\right)\right]. \qquad (6)$$

Here, $\Delta(0)$ is the value of the energy gap and $\lambda(0)$ is the magnetic penetration depth at zero temperature.

At the same time, the unconventional *d*-wave pairing symmetry causes the energy gap to be suppressed along node lines on the Fermi surface. This results in a linear dependence of $\lambda(T) - \lambda(0) \propto T$ at low temperatures. Such a linear *T*-dependence of λ was recently used as a fingerprint of the *d*-wave symmetry for Cooper pairs in cuprate superconductors [33, 34]. From this standpoint, one could argue that the linear $\lambda(T)$ dependence in ZrB_{12} (Fig. 5) may be considered an indication of the *d*-wave symmetry of the condensate of Cooper pairs.

Recently, however, thermodynamic arguments were suggested [35] that a strictly linear *T*-dependence of λ at low temperatures violates the third law of thermodynamics, because it produces nonvanishing entropy in the zero-temperature limit. Therefore, one should expect a deviation from the linear *T*-dependence of λ at very low temperatures. Indeed, recent experiments in cuprates indicate deviation from the linearity of $\lambda(T)$ from the current-carrying zero-energy surface Andreev bound states [36]. We believe that further experiments

on single crystals of ZrB_{12} are necessary to confirm the actual character of the $\lambda(T)$ behavior below 1.0 K. Such experiments are now in progress and may shed light on the nature of the pairing state in this dodecaboride.

Figure 6 displays the change in $\lambda(T)$ in a MgB_2 polycrystalline sample. These measurements were done on samples freely placed in a rectangular solenoid coil forming an *LC* circuit kept at 2.5 K. In Fig. 7, we show the temperature variation of λ for the best MgB_2 film, determined from the one-coil technique and inversion procedure via Eq. (2). A particular feature of these figures is a very similar exponential *T*-dependence at low temperatures for both film and polycrystalline samples.

We used the conventional *s*-wave approach, Eq. (6), to fit these data. In both cases, we observe satisfactory if not perfect agreement between the fits and low-temperature data for thin films. Our fitting parameters (the superconducting gap value at 0 K) are 2.8 and 2.73 meV for film and polycrystalline samples, respectively. The corresponding reduced gap $2\Delta(0)/k_BT_c$ for these samples was found to be 1.64 and 1.62.

Several recent reports on $\lambda(T)$ measurements [23, 37] in MgB_2 provide strong evidence for a predominately exponential temperature dependence of λ at low temperatures, which is consistent with our observation. The reduced gap obtained from exponential fits to the data was found to be 1.42 [37] and 2.3 [23] for single crystals and thin films, respectively. These values, as well as the value we obtained from our data, are significantly smaller than the BCS weak coupling value $2\Delta(0)/k_BT_c = 3.52$. Several other groups have claimed that $\lambda(T)$ in MgB_2 does follow a power-law or even linear *T*-dependence [38]. The possible reason for this discrepancy is that previous studies were limited to temperatures above 4 K, whereas $\lambda(T)$ shows a clear signature of exponential behavior only below 7 K (see Figs. 6 and 7). Another problem may arise in use of nonetched samples, where the damaged surface layer or the proximity effect associated with the presence of a metallic Mg over layer [19] may significantly complicate the use of the surface-sensitive techniques.

We emphasize that our values of the superconducting gap at low temperatures are in the range of values for 3*D* π-bands obtained by point-contact spectroscopy on MgB_2 single crystals ($\Delta_\sigma(0) = 7.1$ meV and $\Delta_\pi(0) = 2.9$ meV for the σ and π bands, respectively) [39]. Our data also agree with theoretical values predicted by the two-band model [40]. Analysis of the overall temperature dependence of λ within the two-band phenomenological model [41] is now in progress and will be published elsewhere. The essential property of this paper is comparison of the ZrB_{12} and MgB_2 low-temperature data, where the $\lambda(T)$ dependence has a totally different behavior.

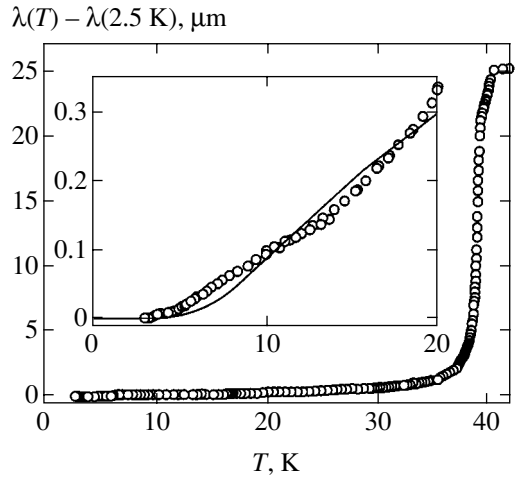

Fig. 6. Temperature variation of the magnetic penetration depth λ for a MgB_2 sample up to T_c. The inset shows the data below $T_c/2$ on an extended scale. The solid line represents the single-gap exponential fit for $\Delta(0) = 2.73$ meV.

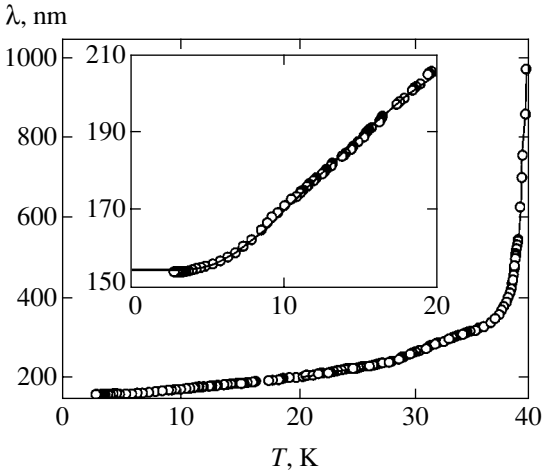

Fig. 7. Temperature variation of λ up to T_c for MgB_2 thin film on Al_2O_3. The inset shows the data below $T_c/2$ on an extended scale. The solid line represents the single-gap exponential fit for $\Delta(0) = 2.8$ meV.

5. UPPER CRITICAL MAGNETIC FIELD

We now turn to the data on electronic transport in a magnetic field. Figure 8 presents the magnetic-field dependent electric resistivity data for ZrB_{12} polycrystalline samples at various temperatures. Two features are clearly seen: (i) the magnetic field shifts the superconducting transition to lower temperatures; (ii) there is a very small longitudinal magnetoresistivity in the normal state. We extracted the completed upper critical magnetic field H_{c2} by extending the maximal-derivative $d\rho/dH$ line (the dashed line in Fig. 8) up to the normal-

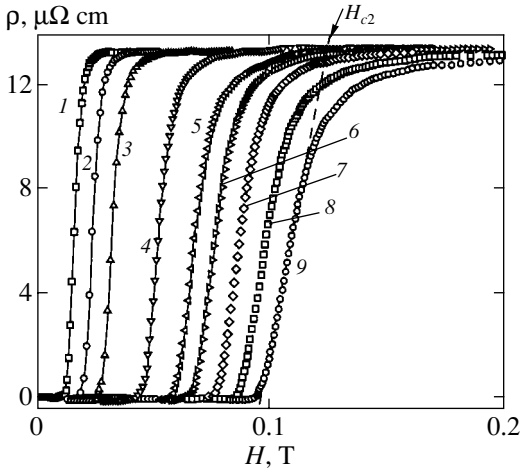

Fig. 8. Magnetic field variation of the resistivity ρ(*T*) in a linear scale for a ZrB_{12} sample at *T* = 5.6 (*1*), 5.2 (*2*), 5.0 (*3*), 4.3 (*4*), 3.6 (*5*), 3.3 (*6*), 2.0 (*7*), 1.8 (*8*), 1.2 (*9*) K. The solid lines are a guide for the eye and the dashed line describes how the resistive transition field H_{c2} has been established.

state level. The crossing point of this line and the normal-state resistivity gave us the value of H_{c2} at various temperatures as indicated by an arrow in Fig. 8. Despite a clear broadening at higher fields, such onset of the resistive transition remains well defined. We note, however, that the resistance may not be an intrinsic property and may be related to the poor grain connection in our polycrystalline samples. Therefore, to obtain a better test for the onset of the superconducting transition, we measured the *RF* susceptibility. Figure 9 shows a plot of the temperature dependence of the resonant frequency *f* of our *LC* circuit as a function of the longitudinal magnetic field. Changes in the resonant frequency are directly proportional to the *RF* susceptibility of the sample. To deduce $H_{c2}(T)$, we used a straight-line fit representing the maximum of the derivative *df*/*dH* (the dashed line in Fig. 9). This straight line was extended up to the normal-state frequency values. We defined H_{c2} as the crossing point of this line with the normal-state frequency $f_n(T)$. As we can see from Fig. 9, this point is very close to the onset point of $f(T)$ in this plot, which makes determination of $H_{c2}(T)$ more reliable.

Figure 10 presents the $H_{c2}(T)$ data that we deduced from these two techniques. A remarkable feature of this plot is a nearly linear increase of H_{c2} with decreasing temperature for both data with no evidence of saturation down to 1.1 K. To obtain the value of $H_{c2}(0)$ from our *RF* data, we assumed that

$$H_{c2}(0) = 0.71T_c\frac{dH_{c2}}{dT}$$

at zero temperature [42]. This assumption yields $H_{c2}(0) = 0.11$ T, which is substantially smaller than the extrapolated value 0.15 T, apparently due to non-BCS or two-gap behavior. Nevertheless, we used this extrapolated number to obtain the coherence length ξ(0) by employing the relation $H_{c2}(0) = \phi_0/2\pi\xi^2(0)$, where ϕ_0 is the magnetic flux quantum. It yields ξ(0) = 60 nm, the value which is substantially larger than a few-angstrom coherence length of high-T_c superconductors. The accuracy of our λ(*T*) measurements in ZrB_{12} did not allow us to determine the absolute values of λ(0). Therefore, the Ginzburg-Landau parameter κ = λ/ξ cannot be determined from these measurements.

Taken as a whole, the temperature dependence of H_{c2} for ZrB_{12} is very similar to that found for MgB_2 [43, 44] and $BaNbO_x$, [45] compounds. Unlike in the conventional BCS theory [42], the $H_{c2}(T)$ dependence is

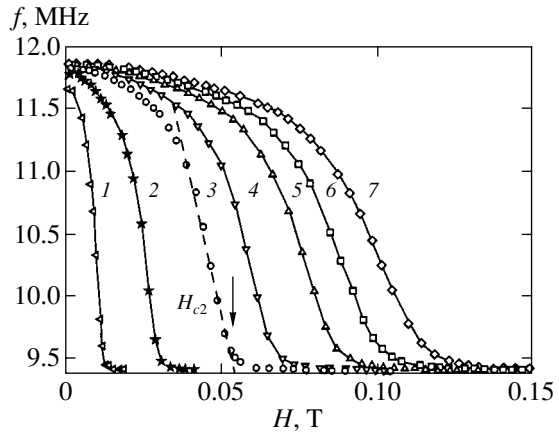

Fig. 9. Magnetic field variation of the resonant frequency of the *LC* circuit for ZrB_{12} sample at the temperatures *T* = 5.8 (*1*), 5.0 (*2*), 4.3 (*3*), 3.3 (*4*), 2.3 (*5*), 1.6 (*6*), 1.1 (*7*) K. The solid lines are a guide to the eye and the dashed line describes a linear extrapolation of the RF data used for the $H_{c2}(T)$ determination.

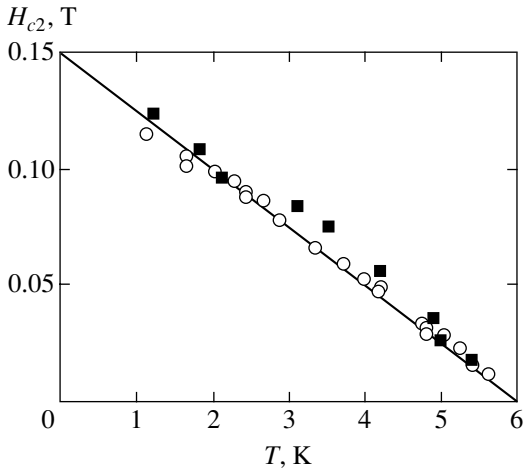

Fig. 10. Temperature variation of the upper critical magnetic field of a ZrB_{12} sample. Different symbols represent the data determined from ρ(*H*) (squares) and *f*(*H*) (circles) data.

linear over an extended region of temperatures with no evidence of saturation at low temperatures. Although the origin of this feature is not completely understood, similar linear $H_{c2}(T)$ dependences have been observed in other high borides and oxide compounds [43–45].

6. CONCLUSIONS

We successfully performed syntheses of polycrystalline samples of dodecaboride ZrB_{12} and diboride MgB_2. We systematically studied the temperature dependences of the resistivity ρ, the magnetic penetration depth λ, and the upper critical magnetic field H_{c2} in these compounds. The electron transport and superconducting properties have been compared with the aim to shed light on the origin of superconductivity in borides. Although the standard Bloch-Grüneisen expression describes the resistivity data in ZrB_{12} fairly well, a better fit was obtained by adding an electron-electron scattering T^2-term in the $\rho(T)$ dependence of MgB_2. This square term dominates the $\rho(T)$ dependence below 150 K in MgB_2, although is almost zero for ZrB_{12}.

The temperature dependence of λ of both polycrystalline and thin-film MgB_2 samples is well described by an *s*-wave behavior of the order-parameter symmetry. Our value of the reduced superconducting gap in MgB_2 samples ($2\Delta(0)/k_BT_c = 1.6$) is significantly smaller than the weak coupling BCS value. However, this value is in a very good agreement with other direct probe measurements of the smaller gap on the π sheets of the Fermi surface. At the same time, we find that λ in ZrB_{12} has a linear temperature dependence over an extended region of temperatures. This feature may be indicative of the *d*-wave pairing, although additional measurements are needed for the final conclusion. We find that the upper critical field $H_{c2}(T)$ deduced from *RF* data is almost the same as that obtained from the resistive data. Both techniques demonstrate an unconventional linear temperature dependence of H_{c2}, with a considerably lower value of $H_{c2}(0) = 0.15$ T. We believe that these observations are clear indicators of the unconventional behavior of electron transport and superconducting properties of dodecaboride ZrB_{12}.

Very useful discussion with V.F. Gantmakher, A.A. Golubov, A. Junod, H. Hilgenkamp, and R. Huguenin, help in sample preparation from V.V. Lomejko, and help in text preparation of L.V. Gasparov are gratefully acknowledged. We are thankful to H.M. Christen, H.-Y. Zhai, M. Paranthaman, and D.H. Lowndes for preparation of excellent MgB_2 films. This work was supported by the Russian Scientific Programs: Superconductivity of Mesoscopic and Highly Correlated Systems (Volna 4G); Synthesis of Fullerenes and Other Atomic Clusters (No. 541-028); Surface Atomic Structures (No. 4.10.99), Russian Ministry of Industry, Science and Technology (MSh-2169.2003.2), the Russian Foundation for Basic Research (No. 02-02- 16874-a), and the INTAS (No. 01-0617).

REFERENCES

1. J. Nagamatsu, N. Nakagawa, T. Muranaka, *et al.*, Nature **410**, 63 (2001).
2. V. A. Gasparov, N. S. Sidorov, I. I. Zver'kova, and M. P. Kulakov, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 601 (2001) [JETP Lett. **73**, 532 (2001)].
3. Z. Fisk, AIP Conf. Proc. **231**, 155 (1991).
4. A. Yamamoto, C. Takao, T. Masui, *et al.*, Physica C (Amsterdam) **383**, 197 (2002).
5. R. Escamilla, O. Lovera, T. Akachi, *et al.*, J. Phys.: Condens. Matter **16**, 5979 (2004).
6. D. P. Young, R. G. Goodrich, P. W. Adams, *et al.*, Phys. Rev. B **65**, 180518(R) (2002).
7. D. Kaczorowski, A. J. Zaleski, O. J. Zogal, and J. Klamut, cond-mat/0103571; D. Kaczorowski, J. Klamut, and A. J. Zaleski, cond-mat/0104479.
8. H. Rosner, W. E. Pickett, S.-L. Drechsler, *et al.*, Phys. Rev. B **64**, 144516 (2001).
9. N. Ashcroft, Phys. Rev. Lett. **21**, 1748 (1968).
10. J. Kortus, I. I. Mazin, K. D. Belashchenko, *et al.*, Phys. Rev. Lett. **86**, 4656 (2001).
11. M. L. Eremets, V. V. Struzhkin, H. K. Mao, and R. J. Hemley, Science **203**, 272 (2001).
12. B. T. Matthias, T. H. Geballe, K. Andres, *et al.*, Science **159**, 530 (1968).
13. C. W. Chu and H. H. Hill, Science **159**, 1227 (1968).
14. Z. Fisk, A. C. Lawson, B. T. Matthias, and E. Corenzwit, Phys. Lett. A **37A**, 251 (1971).
15. K. Hamada, M. Wakata, N. Sugii, *et al.*, Phys. Rev. B **48**, 6892 (1993).
16. I. R. Shein and A. L. Ivanovskiĭ, Fiz. Tverd. Tela (St. Petersburg) **45**, 1363 (2003) [Phys. Solid State **45**, 1429 (2003)].
17. A. Leithe-Jasper, A. Sato, T. Tanaka, *et al.*, Z. Kristallogr.—New Cryst. Struct. **217**, 319 (2002).
18. I. R. Shein, N. I. Medvedeva, and A. L. Ivanovskiĭ, Fiz. Tverd. Tela (St. Petersburg) **45**, 1541 (2003) [Phys. Solid State **45**, 1617 (2003)].
19. M. Paranthaman, C. Cantoni, H. Y. Zhai, *et al.*, Appl. Phys. Lett. **78**, 3669 (2001).
20. V. A. Gasparov and A. P. Oganesyan, Physica C (Amsterdam) **178**, 445 (1991).
21. A. Gauzzi, J. Le Cochec, G. Lamura, *et al.*, Rev. Sci. Instrum. **71**, 2147 (2000).
22. V. A. Gasparov, G. Tsydynzhapov, I. E. Batov, and Qi Li, J. Low Temp. Phys. **139**, 49 (2005); V. A. Gasparov, I. Batov, Qi Li, and C. Kwon, Physica B (Amsterdam) **284–288**, 1021 (2000); Czech. J. Phys. **46** (Suppl. S3), 1401 (1996); Proc. SPIE **2697**, 391 (1996); Phys. Low-Dimens. Struct. **6** (12), 36 (1995).
23. G. Lamura, E. Di Gennaro, M. Salluzzo, *et al.*, Phys. Rev. B **65**, 020506 (2002).

24. V. A. Gasparov, M. R. Mkrtchyan, M. A. Obolensky, and A. V. Bondarenko, Physica C (Amsterdam) **231**, 197 (1994).
25. V. A. Gasparov, M. P. Kulakov, N. S. Sidorov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **80**, 376 (2004) [JETP Lett. **80**, 330 (2004)].
26. J. M. Ziman, *Electrons and Phonons, Theory of Transport Phenomena in Solids* (Oxford Univ. Press, Oxford, 1960; Inostrannaya Literatura, Moscow, 1962).
27. M. Putti, E. G. d'Agliano, D. Marré, *et al.*, Eur. Phys. J. B **25**, 439 (2002).
28. A. V. Sologubenko, J. Jun, S. M. Kazakov, *et al.*, Phys. Rev. B **66**, 014504 (2002).
29. N. V. Vol'kenshtein, V. P. Dyakina, and V. E. Startsev, Phys. Status Solidi **57**, 9 (1973).
30. V. A. Gasparov and R. Huguenin, Adv. Phys. **42**, 393 (1993).
31. I. I. Mazin, O. K. Andersen, O. Jepsen, *et al.*, Phys. Rev. Lett. **89**, 107002 (2002).
32. J. Halbritter, Z. Phys. **243**, 201 (1971).
33. W. N. Hardy, D. A. Bonn, D. C. Morgan, *et al.*, Phys. Rev. Lett. **70**, 3999 (1993).
34. D. A. Bonn, S. Kamal, K. Zhang, *et al.*, Phys. Rev. B **50**, 4051 (1994).
35. N. Schopohl and O. V. Dolgov, Phys. Rev. Lett. **80**, 4761 (1998); **81**, 4025 (1998).
36. A. Carrington, F. Manzano, R. Prozorov, *et al.*, Phys. Rev. Lett. **86**, 1074 (2001).
37. F. Manzano, A. Carrington, N. E. Hussey, *et al.*, Phys. Rev. Lett. **88**, 047002 (2002).
38. Yu. A. Nefyodov, M. R. Trunin, A. F. Shevchun, *et al.*, Europhys. Lett. **58**, 422 (2002).
39. R. S. Gonnelli, D. Daghero, G. A. Ummarino, *et al.*, Phys. Rev. Lett. **89**, 247004 (2002).
40. A. Brinkman, A. A. Golubov, H. Rogalla, *et al.*, Phys. Rev. B **65**, 180517 (R) (2002); A. A. Golubov, A. Brinkman, O. V. Dolgov, *et al.*, Phys. Rev. B **66**, 054524 (2002).
41. F. Bouquet, Y. Wang, R. A. Fisher, *et al.*, Europhys. Lett. **56**, 856 (2001).
42. E. Helfand and N. R. Werthamer, Phys. Rev. Lett. **13**, 686 (1964); Phys. Rev. **147**, 288 (1966).
43. L. Lyard, P. Samuely, P. Szabo, *et al.*, Phys. Rev. B **66**, 180502 (R) (2002).
44. A. V. Sologubenko, J. Jun, S. M. Kazakov, *et al.*, Phys. Rev. B **65**, 180505 (R) (2002).
45. V. A. Gasparov, S. N. Ermolov, S. S. Khasanov, *et al.*, Physica B (Amsterdam) **284**–**288**, 1119 (2000).