Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 435–448.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 499–514.
Original English Text Copyright © 2001 by Kopeliovich.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Characteristic Predictions of Topological Soliton Models[¶]

V. B. Kopeliovich
Institute for Nuclear Research, Russian Academy of Sciences, Moscow, 117312 Russia
e-mail: kopelio@al20.inr.troitsk.ru

Received March 20, 2001

Abstract—Characteristic predictions of the chiral soliton models (the Skyrme model and its extensions) are discussed. The chiral soliton model predictions of low-lying dibaryon states qualitatively agree with recent evidence for the existence of narrow dibaryons in reactions of the inelastic proton scattering on deuterons and the double photon radiation $pp \longrightarrow pp\gamma\gamma$. The connection between magnetic moment operators and inertia tensors valid for arbitrary SU(2) skyrmion configurations allows us to estimate the electromagnetic decay width of some states of interest. Predictions of a different type are multibaryons with a nontrivial flavor (strangeness, charm, or bottom), which can be found, in particular, in high-energy heavy ion collisions. It is shown that the large-B multiskyrmions given by the rational map ansatz can be described within the domain-wall approximation or as a spherical bag with the energy and the baryon number density concentrated at the boundary. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The chiral soliton approach provides a very economical method of describing baryonic systems with different baryon numbers, starting with several basic concepts and ingredients incorporated in the model Lagrangian [1, 2]. The latter is the truncated Lagrangian of effective field theories widely used in describing low-energy meson and baryon interactions [3]. Within this approach, baryons or baryonic systems appear as quantized solitonic solutions of the equations of motion characterized by the so-called winding number or topological charge. If the concept of topological soliton models is accepted and the baryons are indeed skyrmions, it is clear why isospin exists in nature: the number 3 of the SU(2) isospin group generators coincides with the number of space dimensions, thereby allowing for a correlation between SU(2) chiral fields and space coordinates resulting in the appearance of topological solitons.

It has been found numerically that the lowest energy chiral field configurations possess different topological properties—the shape of the mass and B-number distribution—for different values of B. A sphere occurs for the $B = 1$ hedgehog [1], a torus for $B = 2$ [4], a tetrahedron for $B = 3$, a cube for $B = 4$ [5], and higher polyhedrons for greater baryon numbers [5–7]. A paradoxical feature of the approach is that the baryon/nucleon individuality is absent in the lowest energy static configurations (we note that any of the known lowest energy configurations can be made of a number of slightly deformed tori). It is believed that the standard picture of nuclei must emerge when the nonzero modes motion (vibration and breathing) is taken into account. Finding the relative position of states with different quantum numbers (spin, isospin, flavor, SU(3) representation, etc.) requires calculating the zero-mode quantum corrections to the energy of a baryonic system. Corrections of this type were first calculated for configurations of the "hedgehog" type [8] and, later, for axially symmetric configurations [9, 10] and for more general configurations for the SU(2) [11] and SU(3) symmetry groups [12, 13].

The chiral soliton approach provides the concept of nuclear matter that is different from the commonly accepted assumption that the nuclear matter is constructed from separate nucleons only. To find the "smoking gun" for this unusual concept, it is necessary to find some states that cannot be made of separate nucleons, e.g., because of the Pauli exclusion principle. The simplest possibility is to consider the $B = 2$ system, where the Pauli principle strictly and unambiguously forbids definite sets of quantum numbers for the system consisting of separate nucleons.

In this paper, we first discuss the SU(2) case (Section 2), where supernarrow low-lying dibaryons were predicted [14], and estimate their electromagnetic decay width. We next consider the SU(3) extension of the chiral soliton model and extend the previous estimates of the spectra of multibaryons with flavor (strangeness, charm, or bottom quantum number) to higher baryon numbers, where the necessary theoretical information on multiskyrmions is available [7]. A simplified model for large-B multiskyrmions given by rational maps (RM) [15] is presented that allows us to establish the relation to the domain-wall or bag approximation (Section 4). The technical details required for calculations are available in the literature; some of them are given in the Appendices, where several statements valid for any chiral soliton are proved and useful

---

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

1063-7761/01/9303-0435$21.00 © 2001 MAIK "Nauka/Interperiodica"

expressions for the SU(2) skyrmion inertia tensors (still lacking in the literature) are presented.

2. NARROW DIBARYONS BELOW THE $NN\pi$ THRESHOLD

The topological chiral solitons (skyrmions) are classical configurations of chiral fields incorporated in a unitary matrix $U \in$ SU(2) or SU(3) and characterized by the topological or winding number identified with the baryon number B. The classical energy (mass) of these configurations $M_{\rm cl}$ is usually found by minimizing the energy functional that depends on chiral fields. As any extended object, skyrmions also possess other characteristics, e.g., inertia moments Θ (inertia tensors in the general case, see Appendix A), mean square radii of the mass and baryon number distribution, etc. The quantization of the zero modes of chiral solitons allows obtaining the spectrum of states with different values of quantum numbers: spin, isospin, strangeness, etc. [8–13]. Because this approach leads to a reasonable description of various properties of baryons, nucleons, and hyperons, it is interesting to consider predictions of the models of this type for baryonic systems with $B \geq 2$. The energy of SU(2) quantized states with the axial symmetry can be written as [9, 10]

$$E = M_{cl} + \frac{I(I+1)}{2\Theta_I} + \frac{J(J+1)}{2\Theta_J} + \frac{(J_3^{bf})^2}{2B^2\Theta_3}\left(1 - \frac{\Theta_3}{\Theta_I} - B^2\frac{\Theta_3}{\Theta_J}\right), \quad (1)$$

where I and J are the isospin and the spin of the system, J_3^{bf} is the body-fixed third (or z) component of the angular momentum, which can be considered as an additional internal quantum number of the system, and $B = n$ is the azimuthal winding number for the lowest energy axially symmetric configurations. This formula, rigorously obtained from a model Lagrangian [9, 10], has a very transparent physical interpretation. The technical details involving the known Lagrangian of the Skyrme model, expressions for $M_{\rm cl}$, inertia tensors, and some other formulas can be found in Appendix A.

The (generalized) axial symmetry of the configuration with $B = 2$ leads to a certain constraint on the body-fixed third components of the isospin and the angular momentum:

$$J_3^{bf} = -nI_3^{bf} = -nL$$

(see [9, 10]). For the states with $I = 1$ and $J = 0$, or $I = 0$ and $J = 1$, and also $I = J = 1$, it then follows that

$$I_3^{bf} = J_3^{bf} = L = 0.$$

Therefore, the last term in (1), which is proportional to J_3^{bf2}, is absent in these cases. Because the parity of the configuration is equal to $P = (-1)^L$ [10], all the above states have a positive parity. For the state with $I = 0$ and $J = 2$, we can have

$$I_3^{bf} = J_3^{bf} = 0,$$

or

$$I_3^{bf} = L = 1, \quad J_3^{bf} = -2.$$

At large B, it can also be shown (see Appendix A) that only the first two terms in (1), those proportional to $I(I + 1)$ and $J(J + 1)$, are important in the quantum correction to the energy.

It was noted a long time ago [9] that the quantum correction for the deuteron-like state with $I = 0$, $J = 1$ given by $E_d^{\rm rot} = 1/\Theta_J(B = 2)$ is by approximately 30 MeV smaller than the correction for the "quasi-deuteron" state with $I = 1$, $J = 0$ given by $E_{d'}^{\rm rot} = 1/\Theta_I(B = 2)$. This occurs for all known versions of the model, without any tuning of the parameters, and can therefore be considered as an intrinsic property of the chiral soliton models originating from effective field theories. Further investigations of nonzero modes of the two-nucleon system have shown that with many (albeit not all) of them taken into account, the binding energy of the deuteron can be reduced to ~6 MeV [16] if it is considered as a difference between states with the deuteron and the quasi-deuteron quantum numbers. As previously, we consider here the differences of the quantized state energies because they are free of many uncertainties, e.g., those due to unknown loop corrections to the masses of skyrmions (see [17, 18] and discussions below).

In accordance with Eq. (1), some dibaryons are predicted to be decoupled from the 2-nucleon channel as a consequence of the Pauli principle [14]. For example, there is a prediction for the state with the isospin $I = J = 1$, positive parity, and the energy below the threshold for the decay into $NN\pi$ with

$$E_D^{\rm rot} = 1/\Theta_J(B = 2) + 1/\Theta_I(B = 2).$$

This dibaryon cannot be seen in nucleon–nucleon interactions directly, but can be observed in the reaction $NN \longrightarrow NN\gamma\gamma$, where one photon is required to produce D and the second appears from the decay of D, e.g.,

$$pp \longrightarrow D^{++}\gamma \longrightarrow pp\gamma\gamma.$$

The chiral soliton models predict the state D with the isospin $I = J = 1$ at the energy about 50–60 MeV above the NN threshold [14].

In [10], it was shown that the states for which the sum $I + J$ is even (0, 2, etc.) and the parity is positive are forbidden by constraints of the Finkelstein–Rubinstein type arising as a consequence of the requirement that the configuration can be presented as a system of two unit hedgehogs at large relative distances such that these unit skyrmions possess fermionic properties. This implies that the configurations that cannot be considered as consisting of two nucleons were ignored in [10].

In [14], on the contrary, we abandoned this requirement. We also note that the state with $I = 0$, $J = 2$, which was forbidden in [10], can in fact be the ${}^3D_2$ state of two nucleons and should not be forbidden by the FR constraint. This particular case should therefore be analyzed more carefully.

It is possible to estimate the width of the radiative decay $D \longrightarrow NN\gamma$. Electromagnetic nucleon form factors can be described sufficiently well within the Skyrme soliton model in a wide interval of momentum transfers [19]. A reasonable agreement with the data takes place for the deuteron and $2N$ systems [10], and therefore, one can expect reasonable predictions for systems with greater baryon numbers or with unusual properties. The dimensional estimate of the narrow dibaryon decay width was made in [14] providing the lower bound for the decay width given by several eV. To make a more realistic estimate, one can consider a transition of the magnetic type, $D \longrightarrow NN\gamma$ or $d\gamma$. The amplitude of the direct process due to the magnetic dipole transition can be written as

$$M_{D\to NN\gamma} = ie\tilde{\mu}_{D\to NN}\epsilon_{ikl}F_{ik}\Psi_l^D\phi_1^\dagger\phi_2, \tag{2}$$

where $\tilde{\mu}$ is the value of the transition magnetic moment assumed to be of the same order as μ_p; $F_{ik} = e_iq_k - e_kq_i$ is the electromagnetic field strength; and Ψ_l^D, ϕ_1, and ϕ_2 are the respective wave functions of the dibaryon and the nucleons. For the width of this direct decay, we then obtain

$$\Gamma_{D\to NN\gamma} = \alpha\Delta M^2\frac{\tilde{\mu}^2_{D\to NN}}{945\pi^2}(\Delta/M)^{7/2}, \tag{3}$$

which is numerically less than 0.1 eV for

$$\mu \sim \mu_p - \mu_n \approx 4.7/2M_N;$$

here, $\Delta = M_D - 2M$ is the energy release, or the maximum energy of the emitted photon. This estimate agrees with that made previously [14], but the final state interaction could increase it by several orders of magnitude.

To roughly take it into account, one must consider the transition $D \longrightarrow d'$, where d' is the spin-zero quasideuteron, or $D^+ \longrightarrow d$. At this point, an important statement is that the isovector magnetic transition operator for any skyrmion is simply related to its mixed, or interference, inertia tensor $\Theta_{ab}^{\mathrm{int}}$. This statement, known in some particular cases [8, 10], is proved in Appendix B for arbitrary skyrmions and for any type of chiral soliton models: we show that

$$\tilde{\mu}_i^a = -\frac{1}{2}R^{aj}(A)\Theta_{jk}^{\mathrm{int}}O_i^k(A'), \tag{4}$$

where $R^{aj} = D_{aj}^1 = \mathrm{Tr}(A^\dagger\tau^aA\tau^j)/2$, O_i^k are the final rotation matrices, and a is the isotopical (octet in SU(3)) index (for electromagnetic interaction, we must set $a = 3$). $\Theta_{jk}^{\mathrm{int}}$ is given in Appendix A.

For configurations with the generalized axial symmetry and for several known multiskyrmions, only the diagonal elements of Θ^{int} are different from zero, and moreover, only the 33-component remains in the axially symmetric case; we then have

$$\tilde{\mu}_i^3 = -\frac{1}{2}R^{33}(A)\Theta_{33}^{\mathrm{int}}O_i^3(A'), \tag{5}$$

where $\Theta_{33}^{\mathrm{int}} = 2\Theta_{33}^I = 14.8\ \mathrm{GeV}^{-1}$ for $B = 2$ and the accepted values of model parameters (see also Table 1). To obtain numerical values of the transition magnetic moments, we must calculate the rotation matrix elements between the wave functions of the initial and final states. In terms of the final rotation matrices $D_{I_3,L}^I$, these are given by (see, e.g., [20])

$$\Psi_{I,I_3;J,J_3}^D = \sqrt{\frac{2I+1}{8\pi^2}}D_{I_3L}^I\sqrt{\frac{2J+1}{8\pi^2}}D_{J_3,-2L}^J. \tag{6}$$

For the D state, we have $I = J = 1$ and $L = 0$, and for the final d' state, $I = 1$ and $J = 0$. Because $R^{33} = D_{00}^1$, the isotopical part of the matrix element for the $D \longrightarrow d'$ transition is proportional to

$$\langle D_{I_30}^1D_{00}^1D_{I_30}^1\rangle = \int D_{I_30}^1D_{00}^1D_{I_30}^1d\nu = C_{1,0;1,I_3}^{1,I_3}C_{1,0;1,0}^{1,0}/3. \tag{7}$$

One of the Clebsch–Gordan coefficients vanishes, $C_{1,0;1,0}^{1,0} = 0$, and therefore, the $D \longrightarrow d'$ transition magnetic moment is equal to zero for all states including D^{++} and D^0, not only for $D^+ \longrightarrow d'^+$ (which is trivial); this is a consequence of symmetry properties of the rotator wave function with $L = 0$.

For the transition $D \longrightarrow d\gamma$, the isotopical part of the matrix element differs from zero, $\langle D_{0,0}^1D_{00}^1D_{00}^0\rangle = 1/3$, but the angular momentum part proportional to $\langle D_{J_30}^1D_{00}^1D_{J_30}^1\rangle$ is again equal to zero. However, the decay $D^+ \longrightarrow np$ is possible as a result of the second-order isospin violation in the electromagnetic interaction, due to a virtual emission and reabsorption of the photon and due to the isospin violation by the mass difference of the u and d quarks. The order of magnitude estimate of the width of this decay due to the virtual electromagnetic process is

$$\Gamma_{D\to pn} \approx \alpha^2\frac{M}{4\pi}\sqrt{\frac{\Delta}{M}}, \tag{8}$$

which is about ~1 keV. We note that, for the components of D with the charge +2 or 0, the decay into the pp or nn final states is strictly forbidden by the rigorous conservation of the angular momentum and by the Pauli principle.

Table 1. Characteristics of the bound states of skyrmions with the baryon numbers up to $B = 22$

B	$M_{\rm cl}$	$\Theta_F^{(0)}$	Θ_I	$\Theta_{I,3}$	$\bar{\Theta}_J$	Γ	$\tilde{\Gamma}$	$\langle r_0\rangle$	ω_s	ω_c	ω_b
1	1.702	2.05	5.55	5.55	5.55	4.80	15	2.51	0.309	1.542	4.82
2	3.26	4.18	11.5	7.38	23	9.35	22	3.46	0.293	1.511	4.76
3	4.80	6.34	14.4	14.4	49	14.0	27	4.10	0.289	1.504	4.75
4	6.20	8.27	16.8	20.3	78	18.0	31	4.53	0.283	1.493	4.74
5	7.78	10.8	23.5	19.5	126	23.8	35	5.10	0.287	1.505	4.75
6	9.24	13.1	25.4	27.7	178	29.0	38	5.48	0.287	1.504	4.75
7	10.6	14.7	28.9	28.9	220	32.3	43	5.72	0.282	1.497	4.75
8	12.2	17.4	33.4	31.4	298	8.9	46	6.15	0.288	1.510	4.79
9	13.9	20.5	37.7	37.7	375	46	47	6.49	0.291	1.517	4.77
12	18.4	28.0	48.5	48.5	636	64	54	7.31	0.294	1.526	4.79
16	24.5	38.9	63.1	63.1	1107	91	63	8.31	0.301	1.543	4.81
17	25.9	41.2	66.1	66.1	1219	96	65	8.48	0.300	1.542	4.81
22	33.7	56.0	84.2	84.2	2027	135	73	9.55	0.308	1.560	4.84
32*	49.1	86.7	118	118	4154	218	87	11.3	0.319	1.585	4.84

The classical mass of solutions $M_{\rm cl}$ is expressed in GeV; the moments of inertia Θ_F, Θ_I and $\Theta_{I,3}$, Θ_J, $\langle r_0\rangle$, Γ, and $\tilde{\Gamma}$ in GeV^{-1}; and the excitation frequencies for flavor F, $\omega_{s,c,b}$ in GeV; $\langle r_0\rangle = \sqrt{r_B^2}$; Θ_J defines the value of the multiskyrmion isoscalar magnetic moment. For higher baryon numbers, beginning with $B = 9$, calculations are made using the *RM* ansatz. For $B = 32$, it was assumed that the ratio $\mathcal{I}/B^2 = 1.28$ as for the *RM* $B = 22$ skyrmion. The external parameters of the model are $F_\pi = 186$ MeV and $e = 4.12$. The accuracy of calculation is better than 1% for the masses and several percent for other quantities.

For the transitions

$$D^{++} \longrightarrow pp\gamma, \quad D^0 \longrightarrow nn\gamma,$$

and

$$D^+ \longrightarrow (pn)_{I=1}\gamma,$$

the isoscalar magnetic moment operator gives a nonzero contribution. The corresponding matrix element is

$$M_{D\to d'\gamma} = ie\tilde{\mu}^0_{D\to d'}\epsilon_{ikl}F_{ik}\Psi_l^D\Psi^{d'\dagger}. \tag{9}$$

For the rational map parameterization, we have the approximate relation

$$\tilde{\mu}_3^0 \approx J_3\frac{B\langle r_0^2\rangle}{3\Theta^J}, \tag{10}$$

where $\langle r_0^2\rangle$ is the mean square radius of the B-number distribution. Equation (10) coincides with the result in [8] for $B = 1$ and is close to the result in [10] for $B = 2$. The derivation of (10) that is valid for the rational map parametrization of skyrmions will be given elsewhere. The coefficient after J_3 in (10) has a remarkably weak dependence on the baryon number, as can be seen from Table 1. However, numerically, Eq. (10) gives about half the result for $B = 1$ in [8] for the parameters taken here. We thus have

$$\tilde{\mu}^0_{D\to d'} \approx \frac{2\langle r_0^2\rangle}{3\Theta^J}. \tag{11}$$

For the decay width, we then obtain

$$\Gamma_{D\to d'\gamma} = \alpha\frac{4\tilde{\mu}^2_{D\to d'}\Delta^3}{3}. \tag{12}$$

Numerically, $\tilde{\mu}_{D\to d'} \approx 0.35$ GeV^{-1}, and it follows from (12) that $\Gamma_{D\to d'\gamma} \approx 0.3$ keV $(\Delta/60\ {\rm MeV})^3$. The same estimate is valid for the decay rate of $D^+ \longrightarrow np\gamma$ with the np system in the $I = 1$ isospin state.

The experimental evidence for the existence of the narrow dibaryon D in the reaction $pp \longrightarrow pp\gamma\gamma$ has been obtained in Dubna [21], although these data have not been confirmed in the Uppsala bremsstrahlung experiment [22]. Even clearer indications for the existence of low-lying dibaryons were obtained in the experiment at the Moscow meson factory in the reaction $pd \longrightarrow pX$ [23]. As regards its importance, the confirmation of these results is comparable to the discovery of a new elementary particle. The absence of such states would provide definite restrictions on the applicability of the chiral soliton approach and effective field theories.

It should be noted that the model involves a problem with the lowest state with $I = J = 0$, which should be lower than the deuteron-like state. The deuteron must therefore decay into this (0, 0) state and a photon, but a two-nucleon system in the singlet $^1S_0$ state cannot decay because the $0 \longrightarrow 0$ transition is forbidden for the electromagnetic interaction. The loop corrections to the energy of states, or the Casimir energy [16], are different for states that can go over into two nucleons and for states that cannot. Their contribution can change the relative position of these states and shift the (0, 0) state above the deuteron, but a highly nontrivial calculation must be done to verify this.

Some low-lying states with strangeness are also predicted that cannot decay strongly due to the parity and isospin conservation in strong interactions [14]. For example, the dibaryon with the strangeness $S = -2$, $I = 0$, and $J = 1$ and with positive parity has the energy ≈ 0.17 GeV above the $\Lambda\Lambda$ threshold [24], and it cannot decay into two Λ hyperons because of the Pauli principle or into the $\Lambda\Lambda$ final state by the isospin conservation. Therefore, the width of the electromagnetic decay of this state must not exceed several tenths of keV. It is, of course, a special case. Other possible states with the flavor s, c, or b are discussed in the next section.

The masses of neutron-rich light nuclides, such as the tetra-neutron, sexta-neutron, etc., can be estimated using Eq. (1). For the multineutron state with $I = B/2$, the isorotation energy is

$$E^{\mathrm{rot}} = \frac{B(B+2)}{8\Theta_I},$$

and these nuclides are predicted well above the threshold for the strong decay into final nucleons. With increasing the baryon numbers, the energies of neutron-rich states with a fixed difference $N - Z$ decrease, and their widths can therefore be very small. The mass difference of states with the isospin I and the ground states with $I = 0$ (for even B) is equal to

$$\Delta E(B, I) = \frac{I(I+1)}{2\Theta_{I,B}}.$$

For the pairs of nuclei such as ^{8}Li–^{8}Be, ^{12}B–^{12}C, and ^{16}N–^{16}O, it is equal to

$$\Delta E(B, 1) = \frac{1}{\Theta_{I,B}}$$

and decreases with increasing B (i.e., the atomic number), both theoretically (see Table 1 below) and according to data. For $B = 16$, this difference is 10.9 MeV; this is to be compared with the theoretical value of 15.8 MeV, which is not bad for such a crude model.

3. FLAVORED MULTIBARYONS

Another characteristic prediction is that of multibaryons with different values of flavors, such as the strangeness, charm, or bottom quantum numbers. The bound-state approach of multiskyrmions with different flavors is an adequate method to calculate the binding energies of states with quantum numbers s, c, or b. The so-called rigid oscillator model is the most transparent and controllable version of this method [25]. The references to the pioneering papers can also be found in [26]. For the strangeness quantum numbers, the predicted binding energies of flavored states are smaller than the binding energies of ordinary nuclei. For the charm or bottom quantum numbers, the relation is reversed. We now present the main results for flavored multibaryons following [26] and extending them to higher values of the baryon numbers.

To quantize solitons in the SU(3) configuration space in the spirit of the bound-state approach to the description of strangeness, we consider the collective coordinate motion of the meson fields incorporated into a matrix $U \in$ SU(3) (see Appendix A),

$$\begin{aligned} U(r, t) &= R(t)U_0(O(t)\mathbf{r})R^{\dagger}(t), \\ R(t) &= A(t)S(t), \end{aligned} \tag{13}$$

where U_0 is the SU(2) soliton embedded into SU(3) in the standard way (into the upper left corner); $A(t) \in$ SU(2) describes SU(2) rotations; $S(t) \in$ SU(3) describes rotations in the "strange," "charm," or "bottom" directions; and $O(t)$ describes rigid rotations in real space. We have

$$S(t) = \exp(i\mathcal{D}(t)), \quad \mathcal{D}(t) = \sum_{a = 4, \dots, 7} D_a(t)\lambda_a, \tag{14}$$

where λ_a are the Gell–Mann matrices of the (u, d, s), (u, d, c), or (u, d, b) SU(3) groups. The (u, d, c) and (u, d, b) SU(3) groups are totally similar to the (u, d, s) one. For the (u, d, c) group, a simple redefinition of the hypercharge must be made. For the (u, d, s) group,

$$D_4 = \frac{K^+ + K^-}{\sqrt{2}}, \quad D_5 = i\frac{K^+ - K^-}{\sqrt{2}},$$

etc., and for the (u, d, c) group,

$$D_4 = \frac{D^0 + \bar{D}^0}{\sqrt{2}},$$

etc.

The angular velocities of the isospin rotations are defined in the standard way as

$$A^{\dagger}\dot{A} = -i\boldsymbol{\omega} \cdot \boldsymbol{\tau}/2.$$

We do not consider the usual space rotations explicitly because the corresponding inertia moments for baryonic systems are much greater than the isospin inertia moments (see Table 1), and for the lowest possible values of the angular momentum J, the corresponding quantum correction is either exactly zero (for even B) or small.

The magnitude of the D field is small, at least of the order $1/\sqrt{N_c}$, where N_c is the number of QCD colors.

We can therefore safely expand the matrix S in D. To the lowest order in D, the Lagrangian of the model in Eq. (A.1) can be written as

$$\begin{gathered} L = -M_{\mathrm{cl},B} + 4\Theta_{F,B}\dot{D}^{\dagger}\dot{D} \\ -[\Gamma_B\bar{m}_D^2 + \tilde{\Gamma}_B(F_D^2 - F_\pi^2)]D^{\dagger}D - i\frac{N_cB}{2}(D^{\dagger}\dot{D} - \dot{D}^{\dagger}D), \end{gathered} \qquad (15)$$

where

$$\bar{m}_D^2 = (F_D^2/F_\pi^2)m_D^2 - m_\pi^2.$$

Here and below, D is the doublet K^+, K^0 (D^0, D^- or B^+, B^0) and Θ_F is the inertia moment for the rotation into the “flavor” direction (with $F = s$, c, or b and the index c denoting the charm quantum number, except in N_c),

$$\begin{gathered} \Theta_{F,B} = \frac{1}{8}\int(1-c_f) \\ \times\left[F_D^2 + \frac{1}{e^2}((\nabla f)^2 + s_f^2(\nabla n_i)^2)\right]d^3r, \end{gathered} \qquad (16)$$

where f is the profile function of the skyrmion; F_D is the flavor decay constant, i.e., kaon, D meson, or B meson decay constant; and

$$\Gamma_B = \frac{F_\pi^2}{2}\int(1-c_f)d^3r. \qquad (17)$$

The mass term contribution to the static soliton energy is related to Γ by

$$M.t. = m_\pi^2\Gamma/2.$$

The quantity $\tilde{\Gamma}_B$ arises when the flavor symmetry breaking is taken into account in flavor decay constants:

$$\tilde{\Gamma}_B = \frac{1}{4}\int c_f[(\nabla f)^2 + s_f^2(\nabla n_i)^2]d^3r. \qquad (18)$$

It is related to other calculated quantities by

$$\tilde{\Gamma} = 2(M_{\mathrm{cl}}^{(2)}/F_\pi^2 - e^2\Theta_F^{Sk}),$$

where $M_{\mathrm{cl}}^{(2)}$ is the second-order contribution to static mass of the soliton and Θ_F^{Sk} is the Skyrme term contribution to the flavor inertia moment. The contribution proportional to $\tilde{\Gamma}_B$ is suppressed in (15) compared to the term $\sim\Gamma$ by the small factor $\sim F_D^2/m_D^2$ and is more important for strangeness. The term proportional to N_cB arises in (15) from the Wess–Zumino term in the action and is responsible for the difference of the strangeness and antistrangeness (in the general case, flavor and antiflavor) excitation energies [25, 26].

Following the canonical quantization procedure, we write the Hamiltonian of the system including the terms of the order N_c^0 as [25]

$$\begin{gathered} H_B = M_{\mathrm{cl},B} + \frac{1}{4\Theta_{F,B}}\Pi^{\dagger}\Pi \\ + \left(\Gamma_B\bar{m}_D^2 + \tilde{\Gamma}_B(F_D^2 - F_\pi^2) + \frac{N_c^2B^2}{16\Theta_{F,B}}\right)D^{\dagger}D \\ + i\frac{N_cB}{8\Theta_{F,B}}(D^{\dagger}\Pi - \Pi^{\dagger}D), \end{gathered} \qquad (19)$$

where Π is the canonically conjugate momentum to the variable D that describes the oscillator-type motion of the (u, d) SU(2) soliton in the SU(3) configuration space. After the diagonalization that can be done explicitly [25], the normal-ordered Hamiltonian can be written as

$$H_B = M_{\mathrm{cl},B} + \omega_{F,B}a^{\dagger}a + \bar{\omega}_{F,B}b^{\dagger}b + O(1/N_c), \qquad (20)$$

where $a^\dagger$ and $b^\dagger$ are the creation operators of the strangeness (i.e., of antikaons) and antistrangeness (flavor and antiflavor) quantum numbers, and $\omega_{F,B}$ and $\bar{\omega}_{F,B}$ are the frequencies of flavor (antiflavor) excitations. D and Π are related to a and b by [25]

$$D^i = \frac{b^i + a^{\dagger i}}{\sqrt{N_cB\kappa_{F,B}}}, \quad \Pi^i = \frac{\sqrt{N_cB\kappa_{F,B}}(b^i - a^{\dagger i})}{2i} \qquad (21)$$

with

$$\kappa_{F,B} = \sqrt{1 + \frac{16(\bar{m}_D^2\Gamma_B + (F_D^2 - F_\pi^2)\tilde{\Gamma}_B\Theta_{F,B})}{(N_cB)^2}}.$$

For the lowest states, the values of D are small:

$$D \sim [16\Gamma_B\Theta_{F,B}\bar{m}_D^2 + N_c^2B^2]^{-1/4};$$

they increase as $(2|F| + 1)^{1/2}$ with increasing flavor number $|F|$. As noted in [25], deviations of the field D from the vacuum decrease with increasing mass m_D, as well as with increasing number of colors N_c, and the method works for any m_D (and also for charm and bottom quantum numbers). We have

$$\begin{gathered} \omega_{F,B} = \frac{N_cB(\kappa_{F,B} - 1)}{8\Theta_{F,B}}, \\ \bar{\omega}_{F,B} = \frac{N_cB(\kappa_{F,B} + 1)}{8\Theta_{F,B}}. \end{gathered} \qquad (22)$$

It was observed in [26] that, to the leading order in N_c, the difference

$$\bar{\omega}_{F,B} - \omega_{F,B} = \frac{N_cB}{4\Theta_{F,B}}$$

Table 2. The binding energy differences $\Delta\epsilon_{s,\,c,\,b}$ for the states with the isospin $I = T_r + |F|/2$

B	$\Delta\epsilon_{s=-1}$	$\Delta\epsilon_{c=1}$	$\Delta\epsilon_{b=-1}$	$\Delta\epsilon_{s=-2}$	$\Delta\epsilon_{c=2}$	$\Delta\epsilon_{b=-2}$
2	−0.047	−0.03	0.02	−0.053	−0.07	0.02
3	−0.042	−0.01	0.04	−0.036	−0.03	0.06
4	−0.020	0.019	0.06	−0.051	0.022	0.10
5	−0.027	0.006	0.05	−0.063	0.001	0.08
6	−0.019	0.016	0.05	−0.045	0.023	0.10
7	−0.016	0.021	0.06	−0.041	0.033	0.11
8	−0.017	0.014	0.02	−0.040	0.021	0.03
9	−0.023	0.005	0.03	−0.10	−0.003	0.06
12	−0.021	0.003	0.02	−0.09	−0.004	0.04
17	−0.027	−0.013	0.00	−0.11	−0.03	−0.00
22	−0.034	−0.028	−0.03	−0.14	−0.06	−0.03

The binding energy differences $\Delta\epsilon_{s,\,c,\,b}$ are the changes of binding energies of the lowest baryon system with the flavor s, c, or b and the isospin $I = T_r + |F|/2$ compared to the usual u, d nuclei, for the flavor numbers $S = -1, -2$, $c = 1, 2$, $b = -1$ and -2 (see Eq. (24)). The $SU(3)$ multiplets are $(p, q) = (0, 3B/2)$ for even B and $(p, q) = (1, (3B-1)/2)$ for odd B.

coincides with the expression obtained in the collective coordinate approach [24].

The flavor symmetry breaking (FSB) in the flavor decay constants, i.e., the fact that $F_K/F_\pi \approx 1.22$ and $F_D/F_\pi = 1.7 \pm 0.2$ (where we take $F_D/F_\pi = 1.5$ and $F_B/F_\pi = 2$), leads to the increase in the flavor excitation frequencies, in better agreement with the data for charm and bottom. It also leads to some increase in the binding energies of baryon system [26].

The values of $\bar{\Theta}_J$ shown in Table 1 are 1/3 of the trace of the corresponding inertia tensor (see Appendix A). As can be seen from Table 1, the flavor excitation energies increase again for the largest value $B = 22$, and the important property of binding becomes weaker for higher B. However, this can be an artifact of the *RM* approximation discussed in the next section. In particular, for $B \geq 9$, the inertia moments Θ_I and Θ_3 are 1/3 of the trace of the corresponding inertia tensors (see Appendix A).

For large values of $F_D/F_\pi = \rho_D$ and the mass m_D, the following approximate formula for the flavor excitation frequencies can be obtained:

$$\omega_{F,\,B} \approx \tilde{m}_D\left(1 - 2\frac{\Theta^{Sk}_{F,\,B}}{\rho_D^2\Gamma_B}\right) - \frac{N_c B}{2\rho_D^2\Gamma_B} \tag{23}$$

with $\tilde{m}_D^2 = m_D^2 + F_\pi^2\tilde{\Gamma}_B/\Gamma_B$. It is clear from (23) that ω's are smaller than the meson masses m_D, and therefore, the binding always occurs and is to a large degree due to the contribution of the Skyrme term to the flavor inertia Θ^{Sk}_F. As $\rho_D \longrightarrow \infty$, it follows that $\omega_F \longrightarrow m_D$. Because the ratio $\tilde{\Gamma}_B/\Gamma_B$ decreases with increasing B and $\Theta_{F,\,B}/\Gamma_B$ increases as B increases from 1 to 4–7, the energies $\omega_{F,\,B}$ decrease for these B numbers, thereby leading to the increase in the binding of flavored mesons by SU(2) solitons with increasing B up to 4–7. However, for $B = 22$ and 32, the ratio $\Theta_{F,\,B}/\Gamma_B$ is smaller than for $B = 1$, and indeed, ω's are the same and even larger than for $B = 1$.

The binding energy differences between flavored multibaryons and ordinary nuclei in the rigid oscillator approximation are given by

$$\begin{aligned}\Delta\epsilon_{B,\,F} &= |F|\Bigg[\omega_{F,\,1} - \omega_{F,\,B} - \frac{3(\kappa_{F,\,1}-1)}{8\kappa_{F,\,1}^2\Theta_{F,\,1}} \\ &\quad - \frac{T_r(\kappa_{F,\,B}-1)}{4\kappa_{F,\,B}\Theta_{F,\,B}} - \frac{(|F|+2)(\kappa_{F,\,B}-1)^2}{8\kappa_{F,\,B}^2\Theta_{F,\,B}}\Bigg],\end{aligned} \tag{24}$$

and the lowest SU(3) multiplets are considered with the isospin of the flavorless component $T_r = 0$ for even B and $T_r = 1/2$ for odd B. This formula is correct for $|F| = 1$ and for any $|F|$ if the baryon number is sufficiently large to ensure the isospin balance.

The values of $\Delta\epsilon$ shown in Table 2 must be considered as an estimate. They illustrate the restricted possibilities of the *RM* approximation for large-B multiskyrmions.

The isosinglet baryon systems, in particular, those with $|F| = B$, are of special interest. As argued in [26], these states do not belong to the lowest possible SU(3) irreducible representations; they must have $T_r = |F|/2$. It makes sense to calculate the difference between the binding energy of this state and the minimal state

Table 3. The binding energy differences for the states with the isospin $I = 0$

B	$\Delta\epsilon_{s=-1}$	$\Delta\epsilon_{c=1}$	$\Delta\epsilon_{b=-1}$	$\Delta\epsilon_{s=-2}$	$\Delta\epsilon_{c=2}$	$\Delta\epsilon_{b=-2}$	$\Delta\epsilon_{s=-3}$	$\Delta\epsilon_{c=3}$	$\Delta\epsilon_{b=-3}$	$\Delta\epsilon_{b=-B}$
2	–	–	–	–0.075	–0.03	0.02	–	–	–	–0.07
3	0.000	0.034	0.07	–	–	–	–0.08	0.002	0.09	–0.08
4	–	–	–	–0.047	0.030	0.09	–	–	–	–0.13
5	–0.003	0.032	0.06	–	–	–	–0.06	0.035	0.12	–0.15
6	–	–	–	–0.044	0.025	0.09	–	–	–	–0.21
7	0.000	0.040	0.07	–	–	–	–0.04	0.068	0.15	–0.20
8	–	–	–	–0.039	0.023	0.03	–	–	–	–0.28
12	–	–	–	–0.046	0.00	0.03	–	–	–	–0.50
17	–0.020	–0.01	–0.00	–	–	–	–0.08	–0.04	–0.01	–0.82
22	–	–	–	–0.073	–0.06	–0.06	–	–	–	–1.3
32*	–	–	–	–0.088	–0.11	–0.13	–	–	–	–

The binding energy differences between the lowest flavored baryon system with the isospin $I = 0$ and the ground state with the same value of B and $I = 0$ or $I = 1/2$. The first three columns are for $|F| = 1$, the next three columns for $|F| = 2$, and the next three for $|F| = 3$. The state with the flavor value $|F|$ belongs to the $SU(3)$ multiplet with $T_r = |F|/2$. In the last column, the binding energy differences are shown for the isoscalar electrically neutral states with $S = -B$. For $|F| \geq 3$, all estimates are very approximate.

$(p^{\min}, q^{\min})$ with zero flavor, which we identify with the standard nucleus (the ground state). We have

$$\begin{aligned}\Delta\epsilon_{B,F} = {} & |F|\left[\omega_{F,1} - \omega_{F,B} - \frac{3(\kappa_{F,1}-1)}{8\kappa^2_{F,1}\Theta_{F,1}} + \frac{(|F|+2)(\kappa_{F,B}-1)}{8\kappa^2_{F,B}\Theta_{F,B}}\right] \\ & - \frac{1}{2\Theta_{T,B}}\left[\frac{|F|(|F|+2)}{4} - T_r^{\min}(T_r^{\min}+1)\right],\end{aligned} \tag{25}$$

where $T_r^{\min} = 0$, or 1/2 as above.

According to Table 3, the total binding energy, e.g., of the state with $B = 22$ and $S = -2$ is smaller than that of the nucleus with $A = 22$ by 73 MeV, and this state must therefore be well bound. The model used here is too crude for large flavor values, and the results obtained can be used only as an illustration and as a starting point for further investigations. Similar results are also obtained in other versions of the model [27], in particular, in the quark-meson soliton model [28]. For the baryon numbers $B = 3, 4$, estimates of the spectra of baryonic systems with the charm quantum number were made in [29] within the conventional quark model. They are in a relatively good agreement with ours.

In the channel with $B = 2$, the near-threshold state with the strangeness $S = -1$ was observed a long time ago in the reaction $pp \longrightarrow p\Lambda K^+$ [30] and recently confirmed in the COSY experiment [31]. A similar near-threshold $\Lambda\Lambda$ state was observed by the KEK PS E224 collaboration [32]. The Skyrme model explains these near-threshold states with $B = 2$ and predicts similar states for higher values of B. For some values of B beginning with $B \geq 5, 6$, such states with several units of strangeness can be stable with respect to strong interactions. Because of the well-known relation $Q = I_3 + (B + S)/2$ between the charge, the isospin, and the hypercharge of hadrons, the baryon system with several units of strangeness can appear as negatively charged nuclear fragments. For even B and the minimal multiplets $(p, q) = (0, 3B/2)$, the strangeness is $S = -2I$, and the condition for the $Q = -1$ fragment to appear is $-1 = S + B/2$ or $-S = B/2 + 1$. For $B = 6$, this gives $S = -4$; for $B = 8$, $S = -5$; etc. For odd B, the $Q = -1$ state must have the strangeness

$$|S| = (B-1)/2 + 1,$$

i.e., –3, –4, and –5 for $B = 5$, 7, and 9, etc.

The negatively charged long-lived nuclear fragment with the mass about 7.4 GeV observed in the NA52 CERN experiment in a Pb + Pb collision at the energy 158*A* GeV [33] can be, within the chiral soliton models, a fragment with $B = 7$ or 6 and the strangeness $S = -4$ or -5, -6. The confirmation of this result and the search for other negatively charged fragments would be of great importance. For the charm or bottom quantum numbers, the binding energies are greater, but observing these states requires considerably higher incident energies.

4. LARGE-*B* MULTISKYRMIONS FROM RATIONAL MAPS IN THE DOMAIN-WALL APPROXIMATION

The treatment of multiskyrmions was considerably simplified by extensively using the rational map ansatz proposed in [15] (and also adopted in the present paper). At the same time, this ansatz leads to the picture of the multibaryon system at large B that is probably incompatible with the picture for ordinary nuclei. To clarify this point, we here consider large-B multiskyrmions in some kind of a toy model—in the domain-wall approximation; in spite of its simplicity, this model gives relatively good numerical results for the known RM multiskyrmions except those with $B = 1, 2$. Within the rational map ansatz [15], the energy of the skyrmion is given by

$$M = \frac{1}{3\pi}\int\left\{A_N r^2 f'^2 + 2Bs_f^2(f'^2 + 1) + \mathcal{I}\frac{s_f^4}{r^2}\right\}dr \quad (26)$$

in the universal units $3\pi^2 F_\pi/e$.

The coefficient $A_N = 2(N-1)/N$ corresponds to the symmetry group SU(N) [34]. For SU(2), the quantity $\mathcal{I}$ is given in Appendix A. There is the inequality $\mathcal{I} \geq B^2$. Direct numerical calculations have shown and our analytical treatment supports that, at large B and, hence, large $\mathcal{I}$, the multiskyrmion looks like a spherical ball with the profile given by $f = \pi$ inside and $f = 0$ outside the ball. The energy and the B-number density of this configuration is concentrated at its boundary, similarly to the domain wall system considered in [35] in connection with cosmological problems.

We consider such a large-B skyrmion within the "inclined step" approximation. If W is the width of the step and r_0 is the radius of the skyrmion (where the profile is given by $f = \pi/2$), we have

$$f = \pi/2 - (r - r_0)\pi/W \quad \text{for} \quad r_0 - W/2 \leq r \leq r_0 + W/2.$$

We note that this approximation describes the usual domain wall energy [35] with the accuracy ~9%.

We write the energy in terms of W and r_0 and then minimize it with respect to both these parameters and find the minimum energy value. With

$$M(W, r_0) = \frac{1}{3\pi}\left[\frac{\pi^2}{W}(B + A_N r_0^2) + W\left(B + \frac{3\mathcal{I}}{8r_0^2}\right)\right], \quad (27)$$

this gives

$$W_{\min} = \pi\left[\frac{B + A_N r_0^2}{B + 3\mathcal{I}/8r_0^2}\right]^{1/2} \quad (28)$$

and, after the minimization,

$$r_{0\,\min}^2 = \sqrt{\frac{3\mathcal{I}}{8A_N}}.$$

In dimensional units, we then have

$$r_0 = \frac{(6\mathcal{I}/A_N)^{1/4}}{F_\pi e}.$$

Because $\mathcal{I} \geq B^2$, the radius of the minimized configuration grows at least as $\sqrt{B}$. It follows that $W_{\min} = \pi$, which is therefore independent of B for any SU(N). The energy is given by

$$M_{\min} \approx \frac{2B + \sqrt{3A_N\mathcal{I}/2}}{3}. \quad (29)$$

For the SU(2) model, $A_N = 1$ and the energy $M_{\min} = (2B + \sqrt{3\mathcal{I}/2})/3$ should be compared with the lower bound $M_{LB} = (2B + \sqrt{\mathcal{I}})/3$. The formula gives the numbers for $B = 3, \ldots, 22$ in a remarkably good agreement (within 2–3%) with the calculation within the RM approximation [7].

It is not difficult to calculate the corrections to these expressions, of the relative order $1/B$, $1/B^2$, …:

$$\begin{aligned} M(W, r_0) \approx \frac{1}{3\pi}\Bigg\{&\frac{\pi^2}{W}(B + A_N r_0^2) \\ &+ W\left[B(1+\beta) + \frac{3\mathcal{I}}{8r_0^2}(1+\gamma)\right]\Bigg\}, \end{aligned} \quad (30)$$

where

$$\beta = \frac{\pi^2}{12B}, \quad \gamma = \frac{2\pi^2 + 17}{\sqrt{24\mathcal{I}}}.$$

It follows that

$$M_{\min} \approx [2B(1 + \beta/2) + \sqrt{3\mathcal{I}/2}(1 + \gamma/2)]/3. \quad (31)$$

However, the first-order correction in W does not improve the description of masses, and the summation of all terms seems to be required.[1]

We thus see that a very simple approximation confirms the picture emerging from the numerical calculation of the RM skyrmion as a two-phase object, a spherical ball with the profile $f = \pi$ inside and $f = 0$ outside the ball, and a fixed-width envelope with the fixed surface energy density,

$$\rho_M \approx \frac{2B + \sqrt{3\mathcal{I}/2}}{12\pi r_0^2}.$$

We also consider the effect of the mass term. It gives the contribution

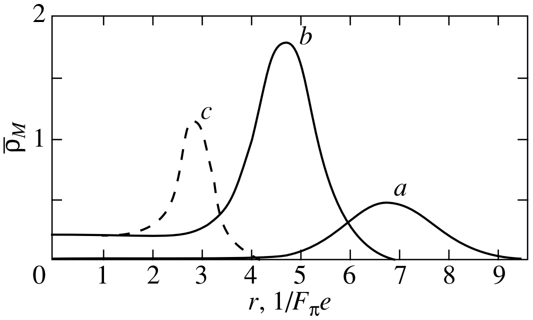
$$M.t. = \tilde{m}\int r^2(1 - \cos f)dr, \quad (32)$$

where

$$\tilde{m} = \frac{8m_\pi^2}{3\pi F_\pi^2 e^2}.$$

---

[1] Detailed analytical treatment of multiskyrmions performed by the author in Pis'ma v ZhETF **73**, 667 (2001) [JETP Lett. **73**, 587 (2001)] confirms the results and conclusions of this section.



The mass density distribution of the rational map multiskyrmion with $B = 22$ as a function of the distance from the center of the skyrmion for different values of mass in the chiral symmetry-breaking term; (*a*) pion mass in the mass term; (*b*) kaon mass, (*c*) *D*-meson mass; the mass density is divided by 10.

For the strangeness, charm, or bottom, the masses m_K, m_D, or m_B must be inserted instead of m_π. In the "inclined step" approximation, we then obtain

$$M.t. \approx \tilde{m}\left[\frac{2}{3}r_0^3 + O(W^2)\right]. \tag{33}$$

In view of this structure of the mass term, it does not affect the width of the step W in the lowest order, but the dimension of the soliton r_0 becomes smaller:

$$r_0 \longrightarrow r_0 - \tilde{m}\frac{r_0^2(B + A_N r_0^2)}{4\pi B}. \tag{34}$$

As was expected from general grounds, dimensions of the soliton decrease with increasing $\tilde{m}$. However, even for large values of $\tilde{m}$, the structure of the multiskyrmion remains the same at large B: it is given by the phase with the broken chiral symmetry inside the spherical wall where the main contribution to the mass and topological charge is concentrated. The behavior of the energy density for $B = 22$ at different values of μ is shown in the figure. The value of the mass density inside the ball is completely determined by the mass term with $1 - c_f = 2$. The baryon number density distribution is quite similar, with the only difference that it is equal to zero inside the bag. It follows from these results that the RM-approximated multiskyrmions cannot model real nuclei at large B, probably for $B > 12$–20, and configurations of the skyrmion crystal type may be more appropriate for this purpose.

In addition to the simple one-shell configurations considered in [7, 15] and here, multishell configurations can also be interesting. Some examples of two-shell configurations with $B = 12, 13, 14$ were considered recently [36]. For these configurations, the profile is given by $f = 2\pi$ at $r = 0$ and decreases to $f = 0$ as $r \longrightarrow \infty$. We can also model this two-shell configuration in the domain-wall or spherical bag approximation with the result

$$M \approx \frac{2B_1 + \sqrt{3\mathcal{I}_1/2}}{3} + \frac{2B_2 + \sqrt{3\mathcal{I}_2/2}}{3}, \tag{35}$$

with the total baryon number $B = B_1 + B_2$. The profile f decreases from 2π to π in the first shell and from π to 0 in the second. The radii of both shells must satisfy the condition

$$r_0^{(2)} \geq r_0^{(1)} + W,$$

and the external shell must therefore be sufficiently large, with the baryon number B_2 given by several tens at least. Because the ratio $\mathcal{I}/B_2$ is larger for smaller B, the energy in Eq. (35) is greater than the energy of the one-shell configuration considered before. Calculations performed in [36] also did not improve the results obtained for the one-shell configuration. However, a more refined analysis would be of interest. The observation concerning the structure of large-B multiskyrmions made above can be useful in view of possible cosmological applications of Skyrme-type models.

5. CONCLUDING REMARKS

We have restricted ourselves to the Skyrme model and its straightforward extensions. However, many of the result are valid in other versions of the model, e.g., in the model with solitons stabilized by the explicit vector (ω) meson or by the baryon number density squared, in the chiral perturbation theory, etc.; see the discussion in the second paper in [14]. The $B = 2$ toruslike configuration has been obtained within these models and in the chiral quark-meson model [28], and it would be interesting to check if there also exist multiskyrmions with $B \geq 3$.

We did not discuss a special class of SU(3) skyrmions, the SO(3) solitons and the problem of their observation. The relevant discussion can be found in [12, 13].

To conclude, the study of some processes, including those at intermediate energies, which to some extent are out of fashion now, can provide a very important check of fundamental principles and concepts of the elementary particle theory including the confinement of quarks and gluons. Confirming the predictions of the chiral soliton approach would give a qualitatively new understanding of the origin of nuclear forces. If the existence of low-energy radiatively decaying dibaryons is reliably established, it will change the long-standing belief that nuclear matter fragments necessarily consist of separate nucleons bound by their interactions. It is therefore extremely important to confirm and check the results of experiments on the dibaryon production and on the production of fragments of flavored matter. This would be possible at accelerators of moderate energies, like COSY (Juelich, Germany), KEK (Japan), Moscow meson factory (Troitsk, Russia), ITEP (Moscow), and several others. The production of multistrange states and the states with charm or bottom quantum numbers is possible in heavy ion collisions and also in accelerators like the Japan Hadron Facility to be built in the near future.

The multiple flavor production realized in the production of flavored multibaryons that is possible, e.g., in heavy ion collisions certainly requires higher energy, but multiple interaction processes and the normal Fermi motion of nucleons inside the nuclei make the effective thresholds much lower [37]. It would allow more complete and reliable verification of the model predictions.

We finally note that low-energy dibaryons were recently obtained in [38] using a quantization procedure different from ours.

ACKNOWLEDGMENTS

This work is supported by the Russian Foundation for Basic Research (project no. 01-02-16615) and UK PPARC (grant PPA/V/S/1999/00004) and was presented in part at the international seminar Quarks-2000 (Pushkin, Russia, May 2000).

APPENDIX A

Inertia Tensors of Multiskyrmions

The Lagrangian density of the SU(2) Skyrme model is given by

$$\begin{aligned}\mathcal{L} &= -\frac{F_\pi^2}{16}\mathrm{Tr}(L_\mu L_\mu) + \frac{1}{32e^2}\mathrm{Tr}G_{\mu\nu}^2 \\ &+ \frac{F_\pi^2 m_\pi^2}{16}\mathrm{Tr}(U + U^\dagger - 2),\end{aligned} \tag{A.1}$$

where $L_\mu = \partial_\mu U U^\dagger$ is the left chiral derivative, $L_\mu = iL_{\mu,k}\tau_k$, τ_k are the Pauli matrices, and $G_{\mu\nu} = \partial_\mu L_\nu - \partial_\nu L_\mu$ is the chiral field strength. The Wess–Zumino term present in the action was discussed in detail in [13], and we omit this discussion here.

We first give the expression for the energy of the SU(2) skyrmion as a function of the profile f and the unit vector $\mathbf{n}$, which is especially useful in some cases. Using the definition $U = c_f + is_f\mathbf{n}\cdot\boldsymbol{\tau}$ and the relation

$$L_{\mu,k}L_{\nu,k} = \partial_\mu f\partial_\nu f + s_f^2\partial_\mu\mathbf{n}\partial_\nu\mathbf{n}, \tag{A.2}$$

we obtain

$$\begin{aligned}M_{\mathrm{stat}} &= \int\left\{\frac{F_\pi^2}{8}[(\nabla f)^2 + s_f^2(\nabla n_i)^2]\right. \\ &\left.+ \frac{s_f^2}{4e^2}\Big[2[\nabla f\times\nabla n_i]^2 + s_f^2[\nabla n_i\times\nabla n_k]^2\Big] + \rho_{M.t.}\right\}d^3r.\end{aligned} \tag{A.3}$$

For the ansatz based on rational maps, the profile f depends on only the variable r, and components of $\mathbf{n}$ depend on only the angular variables θ, ϕ. We have

$$n_x = \frac{2\mathrm{Re}R}{1+|R|^2},\quad n_y = \frac{2\mathrm{Im}R}{1+|R|^2},\quad n_z = \frac{1-|R|^2}{1+|R|^2},$$

where R is a rational function of the variable $z = \tan(\theta/2)\exp(i\phi)$ defining a map $S^2 \longrightarrow S^2$. In this case, the gradients of f and $\mathbf{n}$ are orthogonal (recall that $\nabla_r = \mathbf{n}_r\partial_r + \mathbf{n}_\theta\partial_\theta/r + \mathbf{n}_\phi\partial_\phi/(rs_\theta)$, $\mathbf{n}_r = \mathbf{r}/r = (s_\theta c_\phi, s_\theta s_\phi, c_\theta)$, $\mathbf{n}_\theta = (-c_\theta c_\phi, -c_\theta s_\phi, s_\theta)$, and $\mathbf{n}_\phi = (s_\phi, -c_\phi, 0)$) and $[\nabla f\times\nabla n_1]^2 = f'^2(\nabla n_1)^2$, etc. Using the relations

$$\begin{aligned}n_3^2[\nabla n_2\times\nabla n_3]^2 &= n_1^2[\nabla n_1\times\nabla n_2]^2, \\ n_3^2[\nabla n_1\times\nabla n_3]^2 &= n_2^2[\nabla n_1\times\nabla n_2]^2,\end{aligned} \tag{A.4}$$

we can rewrite (A.3) as

$$\begin{aligned}M_{\mathrm{stat}} &= \int\left\{\frac{F_\pi^2}{8}[f'^2 + s_f^2(\nabla n_i)^2]\right. \\ &\left.+ \frac{s_f^2}{2e^2}\left[f'^2(\nabla n_i)^2 + \frac{s_f^2[\nabla n_1\times\nabla n_2]^2}{n_3^2}\right] + \rho_{M.t.}\right\}d^3r.\end{aligned} \tag{A.5}$$

Introducing the notation

$$\begin{aligned}\mathcal{I} &= \frac{1}{4\pi}\int r^4\frac{[\nabla n_1\times\nabla n_2]^2}{n_3^2}d\Omega \\ &= \frac{1}{4\pi}\int\left(\frac{(1+|z|^2)}{(1+|R|^2)}\frac{|dR|}{|dz|}\right)^4\frac{2idzd\bar{z}}{(1+|z|^2)^2}\end{aligned} \tag{A.6}$$

and using the equation

$$\begin{aligned}\int r^2(\nabla n_k)^2d\Omega &= 2\int r^2\frac{|[\nabla n_1\times\nabla n_2]|}{|n_3|}d\Omega \\ &= 2\int\frac{2idRd\bar{R}}{(1+|R|^2)^2} = 8\pi\mathcal{N},\end{aligned} \tag{A.7}$$

we finally obtain

$$\begin{aligned}M_{\mathrm{stat}} &= 4\pi\int\left\{\frac{F_\pi^2}{8}(f'^2r^2 + 2s_f^2\mathcal{N})\right. \\ &\left.+ \frac{s_f^2}{2e^2}\left[2f'^2\mathcal{N} + \frac{s_f^2\mathcal{I}}{r^2}\right] + r^2\rho_{M.t.}\right\}dr.\end{aligned} \tag{A.8}$$

To find the minimum energy configuration at fixed $\mathcal{N} = B$, one minimizes $\mathcal{I}$ and then finds the profile $f(r)$ by minimizing energy (A.8).

To quantize zero modes, we use the ansatz

$$U(t, r) = A(t)U(O_{ik}(t)r_k)A^\dagger(t)$$

and the obvious relation

$$\partial_t U = \dot{U} = \dot{A}U(\mathbf{r}(t))A^\dagger + AU(\mathbf{r}(t))\dot{A}^\dagger + \dot{r}_t(t)A\partial_i U(\mathbf{r}(t))A^\dagger, \quad (A.9)$$

where $r_i(t) = O_{ik}(t)r_k$ are body-fixed coordinates.

The angular velocities of spatial (or orbital) rotations are introduced as

$$\dot{r}_i = \dot{O}_{ik}r'_k = \dot{O}_{ik}O^{-1}_{kl}r_l(t) = -\epsilon_{ilm}\Omega_m r_l(t)$$

and the integration is performed in the coordinate system bound to the soliton (body-fixed).

The rotation or the zero-mode energy of SU(2) skyrmions as a function of the angular velocities is

$$E_{\rm rot} = \frac{1}{2}\Theta^I_{ab}\omega_a\omega_b + \Theta^{\rm int}_{ab}\omega_a\Omega_b + \frac{1}{2}\Theta^J_{ab}\Omega_a\Omega_b. \quad (A.10)$$

The isotopic inertia tensor for an arbitrary SU(2) skyrmion is given by

$$\Theta^I_{ab} = \int s_f^2\left\{(\delta_{ab} - n_a n_b) \times\left(\frac{F_\pi^2}{4} + \frac{(\nabla f)^2}{e^2}\right) + \frac{s_f^2}{e^2}\partial_l n_a \partial_l n_b\right\}d^3r. \quad (A.11)$$

For the RM ansatz, the trace of this inertia tensor is

$$\Theta^I_{aa}(RM) = 4\pi\int s_f^2\left\{\frac{F_\pi^2}{2} + \frac{2}{e^2}\left(f'^2 + \mathcal{N}\frac{s_f^2}{r^2}\right)\right\}r^2dr. \quad (A.12)$$

The orbital inertia tensor gives the contribution to the energy $\Theta^J_{ab}\Omega_a\Omega_n/2$; using the same notation for an arbitrary configuration, we have

$$\begin{aligned}\Theta^J_{ab} = \int\Big\{&\frac{F_\pi^2}{4}(\partial_i f\partial_k f + s_f^2\partial_i\mathbf{n}\partial_k\mathbf{n}) \\ &+ \frac{s_f^2}{e^2}[\partial_i f\partial_k f(\nabla n_l)^2 + (\nabla f)^2\partial_i\mathbf{n}\partial_k\mathbf{n} \\ &- \partial_i f\partial_l f\partial_l\mathbf{n}\partial_k\mathbf{n} - \partial_k f\partial_l f\partial_l\mathbf{n}\partial_i\mathbf{n} \\ &+ s_f^2[(\nabla n_l)^2\partial_i\mathbf{n}\partial_k\mathbf{n} + (\partial_i\mathbf{n}\partial_l\mathbf{n})(\partial_k\mathbf{n}\partial_l\mathbf{n})]\Big\} \\ &\times\epsilon_{i\alpha a}\epsilon_{k\beta b}r_\alpha r_\beta d^3r.\end{aligned} \quad (A.13)$$

For the RM ansatz, this expression can be simplified as

$$\begin{aligned}\Theta^J_{ab} = \int s_f^2\Big\{&\left[\frac{F_\pi^2}{4} + \frac{f'^2}{e^2} + \frac{s_f^2}{e^2}(\nabla n_l)^2\right] \\ &\times[(\nabla n_l)^2(r^2\delta_{ab} - r_a r_b) - \partial_a\mathbf{n}\partial_b\mathbf{n}r^2] \\ &- \frac{s_f^2}{e^2}[(\partial_i\mathbf{n}\partial_k\mathbf{n})(\partial_i\mathbf{n}\partial_k\mathbf{n})(r^2\delta_{ab} - r_a r_b) \\ &- r^2(\partial_a\mathbf{n}\partial_l\mathbf{n})(\partial_b\mathbf{n}\partial_l\mathbf{n})]\Big\}d^3r.\end{aligned} \quad (A.14)$$

This allows us to obtain the trace of the inertia tensor

$$\Theta^J_{aa}(RM) = 4\pi \times\int s_f^2\left\{\frac{F_\pi^2}{2}\mathcal{N} + \frac{2}{e^2}\left(f'^2\mathcal{N} + \mathcal{I}\frac{s_f^2}{r^2}\right)\right\}r^2dr. \quad (A.15)$$

It is easy to establish the inequality for traces of isotopic and orbital inertia tensors

$$\Theta^J_{aa} - B\Theta^I_{aa} = \frac{8\pi}{e^2}(\mathcal{I} - B^2)\int s_f^4 dr \geq 0, \quad (A.16)$$

because $\mathcal{I} \geq B^2$. The interference (mixed) inertia tensor, which also defines the isovector part of the magnetic transition operator, is equal to

$$\begin{aligned}\Theta^{\rm int}_{ab} = \int s_f^2\Big\{&\left[\frac{F_\pi^2}{4} + \frac{1}{e^2}[(\partial_\nu f)^2 + s_f^2(\partial_\nu\mathbf{n})^2]\right]\partial_i n_l \\ &- \frac{1}{e^2}(\partial_i f\partial_\nu f + s_f^2\partial_i\mathbf{n}\partial_\nu\mathbf{n})\partial_\nu n_l\Big\}n_k\epsilon_{kla}\epsilon_{i\alpha b}r_\alpha d^3r.\end{aligned} \quad (A.17)$$

The components of the spatial angular velocities interfere only with the components ω_1, ω_2, ω_3 of the angular rotation velocities in configuration space.

Numerically, the components of the mixed inertia tensor are much smaller than those of the isotopic or orbital inertia tensor, except in special cases of "hedgehogs," where

$$\Theta^{\rm int} = \Theta^I = \Theta^J,$$

and the axially symmetric configurations where the third components of inertia tensors satisfy the relations

$$\Theta^{\rm int}_{33} = -n\Theta^I_{33} = -\Theta^J_{33}/n.$$

We finally note that the most general formulas for inertia tensors are presented here for the first time. For the RM configurations, they differ in some details from those given in the literature.

APPENDIX B

Electromagnetic Transition Operators

For completeness, we here prove some statements concerning the isovector (octet in the SU(3) case) vector charge and the isovector magnetic moment operator in the general form.

The isovector current and the isospin generator are related by

$$V_{0,a} = \frac{1}{2}\mathrm{Tr}(A^\dagger\lambda_a A\lambda_b)I_b^{bf} = R_{ab}(A)I_b^{bf}, \qquad \text{(B.1)}$$

where in the body-fixed coordinate system (connected with the soliton), the isospin generator is

$$I_b^{bf} = \frac{\partial L^{\mathrm{rot}}(\omega, \Omega)}{\partial\omega_b}. \qquad \text{(B.2)}$$

We have a, $b = 1, 2, 3$ for the SU(2) model and a, $b = 1, \ldots, 8$ for the SU(3) model. To prove this, we consider the ansatz

$$U = \exp(-i\alpha_a\lambda_a/2)A(t)U_0A^\dagger(t)\exp(i\alpha_a\lambda_a/2). \qquad \text{(B.3)}$$

The Noether vector current is the coefficient before the derivative of the probe function, $\partial_\mu\alpha$. In the lowest order in α, we obtain the chiral derivative

$$\begin{aligned} U^\dagger\partial_0 U &= A \\ &\times\left[U_0^\dagger A^\dagger\left(\dot{A} - \frac{i}{2}\dot{\alpha}A\right)U_0 - A^\dagger\left(\dot{A} - \frac{i}{2}\dot{\alpha}A\right)\right]A^\dagger. \end{aligned} \qquad \text{(B.4)}$$

Using the definition of the rotation angular velocities ω_a in configuration space, we obtain

$$A^\dagger\dot{A} - \frac{i}{2}A^\dagger\dot{\alpha}A = -\frac{i}{2}\lambda_b(\omega_b + R_{ab}(A)\dot{\alpha}_a), \qquad \text{(B.5)}$$

where

$$R_{ab}(A) = \frac{1}{2}\mathrm{Tr}(A^\dagger\lambda_a A\lambda_b) \qquad \text{(B.6)}$$

is a real orthogonal matrix. Because the dependence on $\dot{\alpha}$ reduces to a simple addition to angular velocity in accordance with (B.5), Eq. (B.1) follows immediately.

Because of the well known relation

$$Q = B + \frac{I_3}{2} = B + \frac{V_{0,3}}{2}, \qquad \text{(B.7)}$$

the baryonic (topological) charge and the third component of the isospin generator contribute to the charge of the quantized skyrmion.

We also prove that there is a simple relation between the isovector (octet for the SU(3) model) magnetic momentum operator of the skyrmion and the mixed (interference) inertia tensor. We first note that because of the Lorentz invariance, the Lagrangian of an arbitrary chiral model, not only the Skyrme model, can be presented as a linear combination of contributions of the form

$$\mathcal{L}_{M,N} = \mathrm{Tr}(U^\dagger\dot{U}MU^\dagger\dot{U}N - U^\dagger\partial_k UMU^\dagger\partial_k UN), \qquad \text{(B.8)}$$

where M and N are some matrices. For example, $M = N = 1$ for the second-order term. The contribution of the first term in (B.8) to the rotational energy that is proportional to Ω and ω and therefore defines the mixed (interference) inertia tensor is (see (A.9))

$$\begin{aligned} \Theta_{ab}^{\mathrm{int}}\omega_a\Omega_b &= \int\mathrm{Tr}(U_0^\dagger A^\dagger\dot{A}U_0 - A^\dagger\dot{A}) \\ &\times\tilde{M}U_0^\dagger\partial_k U_0\tilde{N}\dot{r}_k d^3r + (M \longleftrightarrow N), \end{aligned} \qquad \text{(B.9)}$$

where $\tilde{M} = A^\dagger MA$ and $\tilde{N} = A^\dagger NA$. Thus,

$$\begin{aligned} \Theta_{ab}^{\mathrm{int}} &= -\frac{i}{2}\epsilon_{bjk}\int r_j(t)\mathrm{Tr}(U_0^\dagger\lambda_a U_0 - \lambda_a) \\ &\times\tilde{M}U_0^\dagger\partial_k U_0\tilde{N}d^3r + (M \longleftrightarrow N), \end{aligned} \qquad \text{(B.10)}$$

where $r_j(t)$ and ∂_k are body-fixed. From the second term in (B.8), we obtain the spatial components of the vector current,

$$\begin{aligned} V_k^a &= \frac{i}{2}\mathrm{Tr}(U_0^\dagger A^\dagger\lambda_a AU_0 - A^\dagger\lambda_a A) \\ &\times\tilde{M}U_0^\dagger\partial_k U_0\tilde{N} + (M \longleftrightarrow N). \end{aligned} \qquad \text{(B.11)}$$

Recalling that

$$A^\dagger\lambda_a A = R_{ab}(A)\lambda_b, \quad R_{ab} = \frac{1}{2}\mathrm{Tr}A^\dagger\lambda_a A\lambda_b,$$

$$\partial_k = O_{lk}\partial_l^{bf},$$

we obtain

$$\begin{aligned} V_k^a &= \frac{i}{2}R_{ab}O_{lk}\mathrm{Tr}(U_0^\dagger\lambda_b U_0 - \lambda_b) \\ &\times\tilde{M}U_0^\dagger\partial_l U_0\tilde{N} + (M \longleftrightarrow N). \end{aligned} \qquad \text{(B.12)}$$

By definition,

$$\mu_i^a = \frac{1}{2}\epsilon_{ijk}\int r_i V_k^a d^3r, \qquad \text{(B.13)}$$

or

$$\begin{aligned} \mu_i^a &= \frac{i}{4}\epsilon_{ijk}R_{ab}(A)O_{qk}O_{pj}\int r_p(t)\mathrm{Tr}(U_0^\dagger\lambda_b U_0 - \lambda_b) \\ &\times\tilde{M}U_0^\dagger\partial_q U_0\tilde{N} + (M \longleftrightarrow N). \end{aligned} \qquad \text{(B.14)}$$

Because

$$\epsilon_{ijk}O_{pj}O_{qk} = \epsilon_{pql}O_{li},$$

we obtain the sought relation between components of the magnetic moment operator and the mixed inertia tensor in the body-fixed coordinate system:

$$\mu_i^a = -\frac{1}{2}R_{ab}(A)\Theta_{bl}^{\text{int}}O_{li}. \tag{B.15}$$

In some particular cases, this relation was used previously [8, 10].

To calculate the transition matrix elements, it is necessary to average this expression over wave functions of some initial and final states (see Section 2).

REFERENCES

1. T. H. R. Skyrme, Proc. R. Soc. London, Ser. A **260**, 127 (1961); Nucl. Phys. **31**, 556 (1962).
2. E. Witten, Nucl. Phys. B **223**, 433 (1983).
3. J. Gasser and H. Leutwyler, Nucl. Phys. B **250**, 465 (1985); U.-G. Meissner, COSY News, No. 7, 6 (1999).
4. V. B. Kopeliovich and B. E. Stern, Pis'ma Zh. Éksp. Teor. Fiz. **45**, 165 (1987) [JETP Lett. **45**, 203 (1987)]; NORDITA Preprint 89/34 (1989); J. J. M. Verbaarschot, Phys. Lett. B **195**, 235 (1987).
5. E. Braaten, S. Townsend, and L. Carson, Phys. Lett. B **235**, 147 (1990).
6. R. A. Battye and P. M. Sutcliffe, Phys. Rev. Lett. **79**, 363 (1997); **86**, 3989 (2001).
7. P. M. Sutcliffe, *Talk presented at the International Seminar Quarks-2000, Pushkin, May 2000.*
8. G. S. Adkins, C. R. Nappi, and E. Witten, Nucl. Phys. B **228**, 552 (1983).
9. V. B. Kopeliovich, Yad. Fiz. **47**, 1495 (1988) [Sov. J. Nucl. Phys. **47**, 949 (1988)].
10. E. Braaten and L. Carson, Phys. Rev. D **38**, 3525 (1988).
11. L. Carson, Nucl. Phys. A **535**, 479 (1991); T. S. Walhout, Nucl. Phys. A **547**, 423 (1992); P. Irwin, Phys. Rev. D **61**, 114024 (2000).
12. G. Guadagnini, Nucl. Phys. B **236**, 35 (1984); A. P. Balachandran *et al.*, Nucl. Phys. B **256**, 525 (1985); B. Schwesinger and H. Weigel, Phys. Lett. B **267**, 438 (1991); V. B. Kopeliovich, Phys. Lett. B **259**, 234 (1991); Yad. Fiz. **51**, 241 (1990) [Sov. J. Nucl. Phys. **51**, 151 (1990)]; Nucl. Phys. A **547**, 315C (1992); C. L. Schat and N. N. Scoccola, Phys. Rev. D **62**, 074010 (2000).
13. V. B. Kopeliovich, Zh. Éksp. Teor. Fiz. **112**, 1941 (1997) [JETP **85**, 1060 (1997)]; Nucl. Phys. A **639**, 75C (1998); V. B. Kopeliovich, B. E. Stern, and W. J. Zakrzewski, Phys. Lett. B **492**, 39 (2000).
14. V. B. Kopeliovich, Yad. Fiz. **58**, 1317 (1995) [Phys. At. Nucl. **58**, 1237 (1995)]; Yad. Fiz. **56** (8), 160 (1993) [Phys. At. Nucl. **56**, 1084 (1993)].
15. C. Houghton, N. Manton, and P. Suttcliffe, Nucl. Phys. B **510**, 507 (1998).
16. R. A. Leese, N. S. Manton, and B. J. Schroers, Nucl. Phys. B **442**, 228 (1995).
17. B. Moussalam, Ann. Phys. (N.Y.) **225**, 264 (1993).
18. H. Walliser, Phys. Lett. B **432**, 15 (1998); N. Scoccola and H. Walliser, hep-ph/9805340; Phys. Rev. D **58**, 094037 (1998).
19. G. Holzwarth, Z. Phys. A **356**, 339 (1996); Nucl. Phys. A **666**, 24 (2000).
20. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1974; Pergamon, New York, 1977) §§58, 110.
21. The DIB-2 Collab., JINR Preprint E1-96-104 (1996); S. B. Gerasimov, nucl-th/9712064; nucl-th/9812077; A. S. Khrykin *et al.*, nucl-ex/0012011; hep-ph/0101184.
22. H. Calen *et al.*, Phys. Lett. B **427**, 248 (1998).
23. E. S. Konobeevsky, M. V. Mordovskoi, S. L. Potashev, *et al.*, Izv. Akad. Nauk, Ser. Fiz. **62**, 2171 (1998); L. V. Filkov, V. L. Kashevarov, E. S. Konobeevski, *et al.*, nucl-ex/9902002; hep-ex/9904003; Phys. Rev. C **61**, 044004 (2000).
24. V. B. Kopeliovich, B. Schwesinger, and B. E. Stern, Nucl. Phys. A **549**, 485 (1992).
25. I. R. Klebanov and K. M. Westerberg, Phys. Rev. D **53**, 2804 (1996); **50**, 5834 (1994); V. B. Kopeliovich, Pis'ma Zh. Éksp. Teor. Fiz. **67**, 854 (1998) [JETP Lett. **67**, 896 (1998)]; hep-ph/9805296.
26. V. B. Kopeliovich and W. J. Zakrzewski, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 675 (1999) [JETP Lett. **69**, 721 (1999)]; Eur. Phys. J. C **18**, 369 (2000).
27. J. P. Garrahan, M. Schvellinger, and N. N. Scoccola, Phys. Rev. D **61**, 014001 (2000); C. L. Schat and N. N. Scoccola, Phys. Rev. D **61**, 034008 (2000).
28. J. Segar, M. Sripriya, and M. Sriram, Phys. Lett. B **342**, 201 (1995); V. B. Kopeliovich and M. S. Sriram, Yad. Fiz. **63**, 552 (2000) [Phys. At. Nucl. **63**, 480 (2000)].
29. B. F. Gibson, C. B. Dover, G. Bhamathi, and D. R. Lehman, Phys. Rev. C **27**, 2085 (1983).
30. J. T. Reed *et al.***,** Phys. Rev. **168**, 1495 (1968).
31. D. Grzonka and K. Kilian, Nucl. Phys. A **639**, 569C (1998); COSY-TOF Collab., COSY News, No. 4, 1 (1999).
32. J. K. Ahn *et al.*, AIP Conf. Proc. **412**, 923 (1997); Nucl. Phys. A **639**, 379C (1998).
33. S. Kabana *et al.*, J. Phys. G **23**, 2135 (1997).
34. T. Ioannidou, B. Piette, and W. J. Zakrzewski, J. Math. Phys. **40**, 6353 (1999); **40**, 6223 (1999).
35. Ya. B. Zeldovich, I. Yu. Kobzarev, and L. B. Okun, Zh. Éksp. Teor. Fiz. **67**, 3 (1974) [Sov. Phys. JETP **40**, 1 (1975)].
36. N. S. Manton and B. M. A. G. Piette, hep-th/0008110.
37. V. B. Kopeliovich, Phys. Rep. **139**, 51 (1986).
38. T. Krupovnickas, E. Norvaisas, and D. O. Riska, nucl-th/0011063.

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 449–461.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 515–528.
Original English Text Copyright © 2001 by Afanasev, Akushevich, Gakh, Merenkov.

NUCLEI, PARTICLES,
AND THEIR INTERACTION

Radiative Corrections to Polarized Inelastic Scattering in the Coincidence Setup[¶]

A. V. Afanasev[a, b], I. Akushevich[a, b, *], G. I. Gakh[c], and N. P. Merenkov[c, **]
[a]*North Carolina Central University, Durham, NC 27707, USA*
[b]*TJNAF, Newport News, VA 23606, USA*
**On leave of absence from the National Center of Particle and High Energy Physics, 220040, Minsk, Belarus*
[c]*NSC Kharkov Institute of Physics and Technology, 61108, Kharkov, Ukraine*
***e-mail: merenkov@kipt.kharkov.ua*

Received April 2, 2001

Abstract—We completely analyze the model-independent leading radiative corrections to the cross section and polarization observables in the semi-inclusive deep-inelastic electron–nucleus scattering with the detection of a proton and the scattered electron in coincidence. The calculations are based on representing the spin-independent and spin-dependent parts of the cross section in terms of the electron structure functions similarly to the Drell–Yan-like representation in electrodynamics. As the applications, we consider the polarization transfer effect from a longitudinally polarized electron beam to the detected proton and the scattering by a polarized target. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Current experiments at the new-generation electron accelerators have reached a new level of precision. This precision requires a new approach to the data analysis and the inclusion of all possible systematic uncertainties. One of the important sources of systematic uncertainties is the electromagnetic radiative effects caused by physics processes in the next orders of perturbation theory.

The purpose of this paper is to develop a unified approach to the computation of radiative effects for the inelastic scattering of polarized electrons in the coincidence setup, namely, in the case where one produced hadron is detected in coincidence with the scattered electron. A broad range of measurements fall into the category of the coincidence electron scattering experiments. This includes deep-inelastic semi-inclusive leptoproduction of hadrons, $(e, e'h)$, and quasielastic nucleon knock-out processes, $(e, e'N)$. Experiments of the former class give access to the flavor structure of quark–parton distributions and fragmentation functions. They are in the focus of experimental programs at CERN, DESY, SLAC, and Jefferson Lab. Some experiments have already been completed, and some are in preparation. The detailed modern review of the activities can be found in [1]. The quasielastic nucleon knock-out process allows studying single-nucleon properties in nuclear medium and probing the nuclear wave function [2, 3].

The different theoretical aspects of strong interactions in the semi-inclusive deep-inelastic scattering (DIS) were studied in a number of papers [4, 5]. The most direct experimental probe of the momentum distribution in nuclei that is presently available is provided by the reaction $A(e, e'N)B$ (see reviews [6]). Specific polarization effects in reactions of this type have been investigated in [7] at the level of the Born approximation with respect to the electromagnetic interaction.

There are several papers dealing with radiative effects for coincidence experiments. The lowest order correction was treated in [8] using the approach of the covariant cancellation of infrared divergences. The leading logarithmic correction was studied in [9] for the charm production. Finally, the radiative correction in quasielastic scattering was recently studied in [10]. Different approaches were used in calculations, and different approximations have been applied. These calculations adopted some specific models for the structure functions. Because the current experimental data do not cover sufficiently wide kinematical ranges, the extrapolation and interpolation procedures must be used in calculating radiative effects. Therefore, the model dependence of the results reduces their generality and hence their applicability. Furthermore, higher order effects, which are important at the current level of the experimental accuracy, were not systematically considered.

The method of the electron structure functions [11] allows the same treatment to be applied to the observed cross section in the lowest order and in higher orders. This results in clear and physically transparent formulas for radiative effects. In this paper, we restrict our consideration to the leading accuracy. This allows us to avoid choosing a preferred model for the hadron struc-

---

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

ture functions and thus to obtain some general formulas for a wide class of physical processes. Whenever needed, the next-to-leading order correction to some specific process can be obtained by the standard procedure. Good examples are the recent calculations of the leading order and the next-to-leading order corrections to polarization observables in DIS [12] and elastic [13] processes.

In this paper, we consider the model-independent radiative corrections to the cross section and polarization observables in the semi-inclusive deep-inelastic scattering of the longitudinally polarized electron off nucleus targets in the case where the target and the detected hadron can be polarized. In Section 2, we use the electron structure function approach to calculate the radiative corrections and to derive the master formulas for the radiatively corrected spin-independent and spin-dependent parts of the corresponding cross sections in the form of the Drell–Yan-type representation in electrodynamics [14]. The result of this section is applicable to leptonic variables if the scattered electron is detected. In Section 3, we apply our master formulas to the case where the polarization of the final nucleon is measured. The radiative corrections to the semi-inclusive DIS on the nucleus target with a vector polarization are calculated in Section 4. In Section 5, we apply our approach to describe the effects of the polarization transfer from the target to the detected nucleon. These effects include the double-spin (hadron–hadron) and triple-spin (electron–hadron–hadron) correlations. In Section 6, we derive the modification of the master formulas for hadronic variables (when the total 4-momentum of all the hadrons is measured instead of the scattered electron) and consider some applications. In Conclusion, we briefly discuss the extension of our results for the radiatively corrected polarization observables beyond the leading-log accuracy.

2. THE MASTER FORMULA

In the recent experiment [15], the polarization transfer to the detected proton was measured in the process with the longitudinally polarized electron beam $^{16}\mathrm{O}(\vec{e}, e, \vec{p})^{15}\mathrm{N}$. This reaction is a particular case of the more general semi-inclusive deep-inelastic polarized process

$$\vec{e}^{\,-}(k_1) + A(p_1) \longrightarrow e^-(k_2) + \vec{p}(p_2) + X. \tag{1}$$

In this paper, we clarify the problem of calculating the electromagnetic radiative corrections to the cross section and polarization observables in a process of this type within the framework of the electron structure function approach.

For process (1) with a definite spin orientation of the proton detected in the final state, we define the cross section in terms of the leptonic and hadronic tensors as

$$d\sigma = \frac{\alpha^2}{(2S_A+1)V(2\pi)^3}\frac{L_{\mu\nu}H_{\mu\nu}}{2\hat{q}^4}\frac{d^3k_2}{\varepsilon_2}\frac{d^3p_2}{E_2}, \tag{2}$$

where S_A is the target spin, $\varepsilon_2(E_2)$ is the energy of the scattered electron (the detected proton), and $\hat{q}$ is the 4-momentum of the virtual photon that probes the hadron block. The hadronic tensor can be expressed through the hadron electromagnetic current J_μ as

$$H_{\mu\nu} = \sum_X \langle p_1|J_\mu(\hat{q})|p_2, X\rangle$$

$$\times \langle X, p_2|J_\nu(-\hat{q})|p_1\rangle \delta(P_x^2 - M_x^2),$$

$$P_x = \hat{q} + p_1 - p_2,$$

where P_x is the total 4-momentum of the undetected hadron system and M_x is its invariant mass.

The electron structure function approach leads to the summation of the leading-log contributions to the leptonic tensor in all orders of the perturbation theory. These contributions arise because of the radiation of the hard collinear as well as the soft and virtual photons and the electron–positron pairs by electrons in the initial and final states. In the leading approximation, the electron tensor on the right-hand side of Eq. (2) can be written as [16]

$$L_{\mu\nu}(k_1, k_2) = \iint \frac{dx_1 dx_2}{x_1 x_2^2} D(x_2, Q^2)$$

$$\times [D(x_1, Q^2)\hat{Q}^B_{\mu\nu}(\hat{k}_1, \hat{k}_2) + i\lambda D_\lambda(x_1, Q^2)\hat{E}^B_{\mu\nu}(\hat{k}_1, \hat{k}_2)], \tag{3}$$

$$Q^2 = -(k_1 - k_2)^2, \quad \hat{k}_1 = x_1 k_1, \quad \hat{k}_2 = \frac{k_2}{x_2},$$

where the respective structure functions $D(x, Q^2)$ and $D_\lambda(x, Q^2)$ describe the radiation of the unpolarized and the longitudinally polarized electrons. At the level of the next-to-leading accuracy, these functions already differ in the first order of the perturbation theory, but in the framework of the leading approximation used here, they only differ in the second order. This difference is caused by the leading contribution of the e^+e^--pair production in the singlet channel to the D function (the effect of the final electron identity). For the unpolarized and longitudinally polarized electron, these contributions are different and are given by [16] (KMS), [17]

$$D^S = \left(\frac{\alpha L}{2\pi}\right)^2\left[\frac{2(1-x^3)}{3x} + \frac{1-x}{2} + (1+x)\ln x\right],$$

$$L = \frac{Q^2}{m_e^2},$$

$$D^S_\lambda = \left(\frac{\alpha L}{2\pi}\right)^2\left[\frac{5(1-x)}{2} + (1+x)\ln x\right],$$

where m_e is the electron mass.

Taking the singlet channel contribution into account usually leads to very small effects (of the order 10^{-4})

because, as one can see, the terms inside the brackets tend to compensate each other (see, e.g., [18]). In what follows, we do not distinguish between D and D_λ, which corresponds to taking only the nonsinglet channel contribution into account (for the corresponding D functions, see [17, 18]). This approximation allows us to write compact formulas for the radiatively corrected cross sections. We also omit the quantity Q^2 from the arguments of the D functions.

The quantity λ entering the right-hand side of Eq. (3) is the degree of the longitudinal polarization of the electron beam. The integration limits are defined below. Representation (3) follows from the quasireal electron approximation [19]. The physical interpretation of the variables x_1 and x_2 is as follows: $1 - x_1 = \omega/\varepsilon_1$ is the ratio of the energy of all the collinear photons and the e^+e^- pairs radiated by the initial electron to the energy of that electron, and $(1 - x_2)/x_2$ is a similar ratio for the scattered electron.

In the Born approximation, we have

$$\begin{aligned} Q^B_{\mu\nu}(k_1, k_2) &= q^2 g_{\mu\nu} + 2(k_1k_2)_{\mu\nu}, \\ E^B_{\mu\nu}(k_1, k_2) &= 2(\mu\nu k_1 k_2), \\ (\mu\nu k_1 k_2) &= \epsilon_{\mu\nu\rho\sigma} k_{1\rho} k_{2\sigma}, \\ (k_1k_2)_{\mu\nu} &= k_{1\mu}k_{2\nu} + k_{1\nu}k_{2\mu}, \quad q = k_1 - k_2. \end{aligned} \tag{4}$$

In the general case, the hadronic tensor on the right-hand side of Eq. (2) depends on the 4-momenta p_1, p_2, the virtual photon 4-momentum $\hat{q} = \hat{k}_1 - \hat{k}_2$, and the 4-vector of the hadron spin S that satisfies the conditions $S^2 = -1$ and $(Sp_2) = 0$. For example, in the case under consideration,

$$H_{\mu\nu} = H^{(u)}_{\mu\nu} + H^{(p)}_{\mu\nu},$$

$$\begin{aligned} H^{(u)}_{\mu\nu} = {} & h_1 \tilde{g}_{\mu\nu} + h_2 \tilde{p}_{1\mu}\tilde{p}_{1\nu} + h_3 \tilde{p}_{2\mu}\tilde{p}_{2\nu} \\ & + h_4(\tilde{p}_1\tilde{p}_2)_{\mu\nu} + ih_5[\tilde{p}_1\tilde{p}_2]_{\mu\nu}, \end{aligned} \tag{5}$$

$$\begin{aligned} H^{(p)}_{\mu\nu} = {} & (Sp_1)[h_6(\tilde{p}_1N)_{\mu\nu} + ih_7[\tilde{p}_1N]_{\mu\nu} \\ & + h_8(\tilde{p}_2N)_{\mu\nu} + ih_9[\tilde{p}_2N]_{\mu\nu}] + (S\hat{q})[h_{10}(\tilde{p}_1N)_{\mu\nu} \\ & + ih_{11}[\tilde{p}_1N]_{\mu\nu} + h_{12}(\tilde{p}_2N)_{\mu\nu} + ih_{13}[\tilde{p}_2N]_{\mu\nu}] \\ & + (SN)[h_{14}\tilde{g}_{\mu\nu} + h_{15}\tilde{p}_{1\mu}\tilde{p}_{1\nu} + h_{16}\tilde{p}_{2\mu}\tilde{p}_{2\nu} \\ & + h_{17}(\tilde{p}_1\tilde{p}_2)_{\mu\nu} + ih_{18}[\tilde{p}_1\tilde{p}_2]_{\mu\nu}], \\ N_\mu = {} & \epsilon_{\mu\nu\rho\sigma} p_{1\nu} p_{2\rho} \hat{q}_\sigma = (\mu p_1 p_2 \hat{q}), \\ [ab]_{\mu\nu} = {} & a_\mu b_\nu - a_\nu b_\mu, \\ \tilde{g}_{\mu\nu} = {} & g_{\mu\nu} - \frac{\hat{q}_\mu \hat{q}_\nu}{\hat{q}^2}, \quad \tilde{p}_{i\mu} = p_{i\mu} - \frac{(\hat{q}p_i)\hat{q}_\mu}{\hat{q}^2}, \quad i = 1, 2, \end{aligned} \tag{6}$$

where h_i (i = 1–18) are the hadron semi-inclusive structure functions that depend on four invariants in general. These invariants can be taken as $\hat{q}^2$, $(\hat{q}p_1)$, $(\hat{q}p_2)$, and (p_1p_2).

The j component of the proton polarization P^j that could be measured experimentally is defined as the ratio of the spin-dependent part of cross section (2) (which is caused by the contraction of the leptonic tensor with the spin-dependent part of the hadronic one $H^{(p)}_{\mu\nu}$, with the given j component of the proton spin) to the spin-independent one (which is caused by the contraction of $L_{\mu\nu}$ with $H^{(u)}_{\mu\nu}$),

$$P^j = \frac{d\sigma^{(p)}(\lambda, S^j, k_1, k_2, p_1, p_2)}{d\sigma^{(u)}(\lambda, k_1, k_2, p_1, p_2)}. \tag{7}$$

We note that P^j is nonzero even if $\lambda = 0$ (the case of the unpolarized electron beam) because of nonzero single-spin correlations in semi-inclusive processes.

In principle, three independent components can be measured in process (1): P^l (longitudinal), P^t (transverse), and P^n (normal), which could be taken relative to the definite physical directions and planes created by 3-momenta of the particles participating in the process. If no additional particle (photons and e^+e^- pairs) radiated by electrons with the 4-momenta k_1 and k_2 is detected, there are three independent directions along $\mathbf{p}_2$, $\mathbf{k}_1$, and $\mathbf{k}_2$. In this case, any component of the proton polarization and the corresponding proton spin components S^j are defined for the Born kinematics and their directions are not affected by the radiation.

Combining formulas (2) for the cross section and definitions (3) and (4) of the lepton and (5) and (6) of the hadron tensors and taking the above discussion into account, we can write the cross section of process (1) as

$$\begin{aligned} \varepsilon_2 E_2 \frac{d\sigma(\lambda, S^j, k_1, k_2, p_1, p_2)}{d^3k_2 d^3p_2} = {} & \iint \frac{dx_1 dx_2}{x_2^2} D(x_1) D(x_2) \\ & \times \hat{\varepsilon}_2 E_2 \frac{d\sigma^B(\lambda, S^j, \hat{k}_1, \hat{k}_2, p_1, p_2)}{d^3\hat{k}_2 d^3p_2}, \end{aligned} \tag{8}$$

where $j = l, t, n$. The factor $1/x_1$ that enters the definition of $L_{\mu\nu}$ is absorbed into the flow in the reduced Born cross section that is by definition given by (see Eq. (2))

$$\begin{aligned} \hat{\varepsilon}_2 E_2 \frac{d\sigma^B(\lambda, S^j, \hat{k}_1, \hat{k}_2, p_1, p_2)}{d^3\hat{k}_2 d^3p_2} = {} & \frac{\alpha^2}{(2S_A + 1)\hat{V}(2\pi)^3} \\ & \times \frac{L^B_{\mu\nu}(\hat{k}_1, \hat{k}_2, \lambda) H_{\mu\nu}(S^J, \hat{q}, p_1, p_2)}{2\hat{q}^4}, \end{aligned}$$

where $\hat{V} = x_1 V$. Within the chosen accuracy, representation (8) is valid for both the spin-dependent ($d\sigma^{(p)}$) and spin-independent (($d\sigma^{(u)}$) parts of the cross section.

In theoretical calculations, it is often useful to parameterize the proton spin 4-vector, which enters the definition of the hadron tensor, in terms of the particle 4-momenta [20]. In our case, we have four 4-momenta to express any component of the proton spin S^j such that

$$S^j = S^j(k_1, k_2, p_1, p_2). \tag{9}$$

We temporarily imagine that the chosen parametrization on the right-hand side of Eq. (9) is stabilized by the relative substitution

$$k_1 \longrightarrow \hat{k}_1, \quad k_2 \longrightarrow \hat{k}_2,$$

$$S^{j_s}(k_1, k_2, p_1, p_2) = S^{j_s}(\hat{k}_1, \hat{k}_2, p_1, p_2).$$

(In what follows, we label such stabilized parametrizations by an index with a lowercase letter.) In this case, we can write the Born cross section in the integrand on the right-hand side of Eq. (8) as

$$\begin{aligned}
&\hat{\varepsilon}_2 E_2 \frac{d\sigma^B(\lambda, S^j, \hat{k}_1, \hat{k}_2, p_1, p_2)}{d^3\hat{k}_2 d^3 p_2} \\
&= \hat{\varepsilon}_2 E_2 \frac{d\sigma_j^B(\lambda, \hat{k}_1, \hat{k}_2, p_1, p_2)}{d^3\hat{k}_2 d^3 p_2}.
\end{aligned} \tag{10}$$

If the proton spin S^J is unstable under the above substitution (in this case, we use an uppercase index), it can always be expressed in terms of the stabilized one by means of an orthogonal matrix,

$$\begin{gathered}
S^J(k_1, k_2, p_1, p_2) = A_{J_j}(k_1, k_2, p_1, p_2) S^j(\hat{k}_1, \hat{k}_2, p_1, p_2), \\
A_{J_j} = -S^J S^j.
\end{gathered} \tag{11}$$

Using this formula and recalling that, in the class of processes considered here, the hadron tensor depends on the proton spin linearly, we can write the master representation for the spin-dependent part ($d\sigma^{(p)}$) of the cross section of process (1) for an arbitrary orientation of the proton spin as

$$\begin{aligned}
&\varepsilon_2 E_2 \frac{d\sigma(\lambda, S^J, k_1, k_2, p_1, p_2)}{d^3 k_2 d^3 p_2} = A_{J_j} \iint \frac{dx_1 dx_2}{x_2^2} \\
&\times D(x_1) D(x_2) \hat{\varepsilon}_2 E_2 \frac{d\sigma_j^B(\lambda, \hat{k}_1, \hat{k}_2, p_1, p_2)}{d^3\hat{k}_2 d^3 p_2},
\end{aligned} \tag{12}$$

where the summation over the index $j = l, t, n$ is implied.

This representation is the electrodynamical analogue of the Drell–Yan formula well known in QCD [14], which has previously been applied to calculate the electromagnetic radiative corrections to the total cross section of the electron–positron annihilation into hadrons [17], to the small-angle Bhabha scattering cross section at LEP1 [18], to unpolarized [21] and polarized deep-inelastic cross sections [12], and to the polarized elastic electron–proton scattering [13]. In the next section, we show how this representation can be used to describe the leading radiative corrections in polarized semi-inclusive deep-inelastic events. Within the leading accuracy, we must find adequate parametrizations of the proton spin 4-vector, calculate the elements of the orthogonal matrix A_{J_j}, derive the spin-independent and spin-dependent parts of the Born cross section for a given parametrization S^j, and determine the x_1 and x_2 integration limits in master formula (12).

3. THE ANALYSIS OF SEMI-INCLUSIVE DEEP-INELASTIC EVENTS WITH THE POLARIZATION TRANSFER

We begin with the parametrizations of the proton spin 4-vector in process (1). To describe this process, we use the set of invariant variables

$$\begin{gathered}
z = \frac{2p_1p_2}{V}, \quad z_{1,2} = \frac{2k_{1,2}p_2}{V}, \quad y = \frac{2p_1(k_1 - k_2)}{V}, \\
x = \frac{-q^2}{2p_1q}, \quad V = 2p_1k_1, \quad q = k_1 - k_2.
\end{gathered} \tag{13}$$

It is physically justified to determine the longitudinal component of the proton spin along the direction of $-\mathbf{p}_1$ as seen from the rest frame of the detected proton. This direction is not affected by the lepton collinear radiation, and the corresponding parametrization is given by

$$S_\mu^l = \frac{zp_{2\mu} - 2\tau_2 p_{1\mu}}{m\sqrt{z^2 - 4\tau_1\tau_2}}, \quad \tau_1 = \frac{M^2}{V}, \quad \tau_2 = \frac{m^2}{V}, \tag{14}$$

where $M(m)$ is the mass of the target nucleus (detected proton). It is easy to verify that, in the rest frame of the proton ($p_2 = (m, 0)$, this longitudinal component is equal to $(0, -\mathbf{n}_1)$, where $\mathbf{n}_1 = \mathbf{p}_1/|\mathbf{p}_1|$, and in the laboratory system ($p_1 = (M, 0)$), it is equal to $(|\mathbf{p}_2|, E_2\mathbf{n}_2)/m$, where $\mathbf{n}_2$ is the unit vector in the direction of the detected proton 3-momentum.

For the fixed longitudinal component, we have several possibilities to determine the transverse and normal components. We first take the transverse component in the plane ($\mathbf{k}_1$, $\mathbf{p}_2$) and the normal component in the plane that is perpendicular to it. The orientations of these planes do not change under the substitution $\mathbf{k}_1 \longrightarrow \hat{\mathbf{k}}_1$, and we therefore have

$$\begin{gathered}
S_\mu^t \\
= \frac{(z^2 - 4\tau_1\tau_2)k_{1\mu} + (2z_1\tau_1 - z)p_{2\mu} + (2\tau_1 - zz_1)p_{1\mu}}{\sqrt{V(z^2 - 4\tau_1\tau_2)[1]}}, \\
S_\mu^n = \frac{2(\mu k_1 p_1 p_2)}{\sqrt{V^3[1]}}, \\
[1] = zz_1 - \tau_2 - z_1^2\tau_1, \quad (S^jS^i) = -\delta_{ji}.
\end{gathered} \tag{15}$$

Totally similarly to the above procedure, we can determine another stabilized set of transverse and normal components relative to the plane ($\mathbf{k}_2$, $\mathbf{p}_2$),

$$\begin{gathered}\tilde{S}^t_\mu = \frac{(z^2-4\tau_1\tau_2)k_{2\mu} + (2z_2\tau_1 - z(1-y))p_{2\mu} + (2\tau_2(1-y) - zz_2)p_{1\mu}}{\sqrt{V[z^2-4\tau_1\tau_2][2]}},\\ \tilde{S}^n_\mu = \frac{2(\mu k_2 p_1 p_2)}{\sqrt{V^3[2]}}, \quad [2] = zz_2(1-y) - \tau_2(1-y)^2 - z_2^2\tau_1.\end{gathered} \tag{16}$$

The sets in Eqs. (15) and (16) represent the complete list of stabilized parametrizations of the proton spin components under the condition that the longitudinal component is chosen in accordance with Eq. (14). There are many unstable parametrizations that can be taken relative to an arbitrary plane ($a\mathbf{k}_1 + b\mathbf{k}_2$, $\mathbf{p}_2$) with arbitrary numbers a and b. In what follows, we consider only the physically favorable set with $a = -b = 1$. The corresponding transverse and normal components are given by

$$\begin{gathered}S^T_\mu = \frac{(z^2-4\tau_1\tau_2)q_\mu + (2(z_1-z_2)\tau_1 - zy)p_{2\mu} + (2y\tau_2 - z(z_1-z_2))p_{1\mu}}{\sqrt{V(z^2-4\tau_1\tau_2)[q]}},\\ S^N_\mu = \frac{2(\mu q p_1 p_2)}{\sqrt{V^3[q]}}, \quad [q] = zy(z_1-z_2) + xy(z^2-4\tau_1\tau_2) - (z_1-z_2)^2\tau_1 - y^2\tau_2.\end{gathered} \tag{17}$$

We now consider the relation between the stabilized set (for definiteness, we work with set (15)) and an unstable one. It is obvious that this relation can be written as

$$S^N = S^n\cos\theta + S^t\sin\theta, \quad S^T = -S^n\sin\theta + S^t\cos\theta, \tag{18}$$

where

$$\cos\theta = -(S^N S^n) = -(S^T S^t) = \frac{z(z_1(1+y)-z_2) + xy(z^2-4\tau_1\tau_2) - 2z_1(z_1-z_2)\tau_1 - 2y\tau_2}{2\sqrt{[1][q]}},$$

$$\begin{gathered}\sin\theta = -(S^N S^t) = (S^T S^n) = \frac{\eta}{2}\sqrt{\frac{z^2-4\tau_1\tau_2}{[1][q]}},\\ \eta = \operatorname{sgn}[(p_1p_2k_1k_2)]\sqrt{\frac{16}{V^4}(p_1p_2k_1k_2)^2},\\ (p_1p_2k_1k_2) = \epsilon_{\mu\nu\rho\sigma}p_{1\mu}p_{2\nu}k_{1\rho}k_{2\sigma},\\ \frac{16(p_1p_2k_1k_2)^2}{V^4} = x^2y^2(4\tau_1\tau_2 - z^2)\\ + 2xy[z(z_2 - z_1(1-y)) - 2z_1z_2\tau_1 - 2(1-y)\tau_2]\\ -(z_2 - z_1(1-y))^2.\end{gathered}$$

One can verify that the necessary condition $\cos^2\theta + \sin^2\theta = 1$ is satisfied.

We can now write the spin-independent part (which is actually independent of the proton spin only) and the spin-dependent part of the cross section of process (1) as

$$\begin{gathered}\varepsilon_2 E_2\frac{d\sigma_{(u),L}}{d^3k_2d^3p_2}\\ = \iint\frac{dx_1dx_2}{x_2^2}D(x_1)D(x_2)\hat{\varepsilon}_2E_2\frac{d\hat{\sigma}^B_{(u),l}}{d^3\hat{k}_2d^3p_2},\end{gathered} \tag{19}$$

$$\begin{gathered}\varepsilon_2E_2\frac{d\sigma_N}{d^3k_2d^3p_2} = \iint\frac{dx_1dx_2}{x_2^2}D(x_1)D(x_2)\\ \times\hat{\varepsilon}_2E_2\left[\cos\theta\frac{d\hat{\sigma}^B_n}{d^3\hat{k}_2d^3p_2} + \sin\theta\frac{d\hat{\sigma}^B_t}{d^3\hat{k}_2d^3p_2}\right],\end{gathered} \tag{20}$$

$$\begin{gathered}\varepsilon_2E_2\frac{d\sigma_T}{d^3k_2d^3p_2} = \iint\frac{dx_1dx_2}{x_2^2}D(x_1)D(x_2)\\ \times\hat{\varepsilon}_2E_2\left[-\sin\theta\frac{d\hat{\sigma}^B_n}{d^3\hat{k}_2d^3p_2} + \cos\theta\frac{d\hat{\sigma}^B_t}{d^3\hat{k}_2d^3p_2}\right],\end{gathered} \tag{21}$$

where $d\hat{\sigma}^B$ with any lowercase index denotes the corresponding Born cross section given at the shifted values of $k_{1,2} \longrightarrow \hat{k}_{1,2}$. The corresponding shifted dimensionless variables introduced by relation (13) are given by

$$\begin{gathered}\hat{x} = \frac{x_1xy}{x_1x_2+y-1}, \quad \hat{y} = \frac{x_1x_2+y-1}{x_1x_2}, \quad \hat{V} = x_1V,\\ \hat{z} = \frac{z}{x_1}, \quad \hat{z}_1 = z_1, \quad \hat{z}_2 = \frac{z_2}{x_1x_2}.\end{gathered} \tag{22}$$

Equations (19)–(21) are the straightforward consequences of master representation (12). Obviously, in order to obtain $d\sigma_n$ and $d\sigma_t$ in the left-hand sides of Eqs. (20) and (21), we must set $\cos\theta = 1$, $\sin\theta = 0$.

Next, we must derive the Born cross sections that enter the right-hand sides of Eqs. (19)–(21). The spin-independent part of the cross section for the longitudinally polarized electron beam (with the degree λ) is expressed in terms of the hadron structure functions $h_1, \ldots, h_5$ as

$$\varepsilon_2 E_2 \frac{d\sigma^B_{(u)}}{d^3k_2 d^3p_2} = \frac{\alpha^2 V}{2(2S_A+1)(2\pi)^3 q^4} H_1, \tag{23}$$

$$\begin{aligned} H_1 = {} & -\frac{2xy}{V}h_1 + (1 - y - xy\tau_1)h_2 + (z_1 z_2 - xy\tau_2)h_3 \\ & + (z_2 + z_1(1-y) - xyz)h_4 - \lambda\eta h_5. \end{aligned}$$

We note that the phase space of the detected proton can also be expressed in terms of invariant variables (13) as

$$\frac{d^3p_2}{E_2} = \frac{V}{2|\eta|}dz_1 dz_2 dz. \tag{24}$$

If the proton spin is directed along S^l, the spin-dependent part of the Born cross section is given by

$$\begin{aligned} \varepsilon_2 E_2 \frac{d\sigma^B_l}{d^3k_2 d^3p_2} = {} & -\frac{\alpha^2 V^3 \eta\sqrt{z^2 - 4\tau_1\tau_2}}{8(2S_A+1)m(2\pi)^3 q^4} \\ & \times\left[H_2 + \frac{z(z_1 - z_2) - 2y\tau_2}{z^2 - 4\tau_1\tau_2}H_3\right], \end{aligned} \tag{25}$$

$$H_2 = (2-y)h_6 + (z_1+z_2)h_8 + \frac{\lambda}{\eta}(\eta_1 h_7 + \eta_2 h_9),$$

$$H_3 = (2-y)h_{10} + (z_1+z_2)h_{12} + \frac{\lambda}{\eta}(\eta_1 h_{11} + \eta_2 h_{13}),$$

$$\eta_1 = y[z_2 - z_1(1-y) - xz(2-y) + 2x(z_1+z_2)\tau_1],$$

$$\begin{aligned} \eta_2 = {} & (z_1 - z_2)(z_2 - z_1(1-y)) \\ & + xyz(z_1+z_2) - 2xy(2-y)\tau_2. \end{aligned}$$

For the transverse orientation of the spin (along S^t), we have

$$\begin{aligned} \varepsilon_2 E_2 \frac{d\sigma^B_t}{d^3k_2 d^3p_2} = {} & \frac{\alpha^2 V^2 \eta}{8(2S_A+1)(2\pi)^3 q^4} \\ & \times\sqrt{\frac{V}{z^2 - 4\tau_1\tau_2}}\left[\psi H_3 - \frac{z^2 - 4\tau_1\tau_2}{\sqrt{[1]}}H_4\right], \end{aligned} \tag{26}$$

$$\begin{aligned} \psi & = \frac{xy(z^2 - 4\tau_1\tau_2) + (z - 2z_1\tau_1)(z_1 - z_2) + (zz_1 - 2\tau_2)y}{\sqrt{[1]}} \\ & = 2\sqrt{[q]}\cos\theta, \end{aligned}$$

where H_4 can be obtained from H_1 by the simple replacement $h_i \longrightarrow h_{i+13}$.

Finally, for the normal orientation of the proton spin (along S^n), the spin-dependent part of the cross section of process (1) is given by

$$\begin{aligned} & \varepsilon_2 E_2 \frac{d\sigma^B_n}{d^3k_2 d^3p_2} \\ & = \frac{\alpha^2 V^2 \sqrt{V}}{8(2S_A+1)(2\pi)^3 q^4}\left[-\frac{\eta^2}{\sqrt{[q]}}H_3 - \psi H_4\right]. \end{aligned} \tag{27}$$

We must also determine the integration limits for x_1 and x_2 in master representation (12). They can be obtained from the condition that the semi-inclusive deep-inelastic process occurs. For the electron–proton scattering, this is possible under the condition that the hadron state involves at least a proton and a pion. This leads to the inequality

$$x_1 x_2 + y - 1 - x_1 xy \geq x_2\delta, \quad \delta = \frac{(m + m_\pi)^2 - m^2}{V}, \tag{28}$$

where m_π is the pion mass. For the integration limits, we then have

$$1 > x_2 > \frac{1 - y + xyx_1}{x_1 - \delta}, \quad 1 > x_1 > \frac{1 + \delta - y}{1 - xy}. \tag{29}$$

For the electron–nucleus scattering process (1) considered here, we must change the pion mass entering the definition of δ by the bound energy of the ejected proton in a given nucleus.

It is interesting to note that, in the case where the final proton polarizations are measured relative to stabilized orientations, the corresponding Born values and the leading radiative corrections to them are expressed in terms of the same hadron structure functions. The situation changes radically if the polarizations are measured relative to unstable orientations. In this case, the contributions to the polarizations caused by the radiative corrections due to the hard collinear radiation are expressed in terms of different sets of hadron structure functions compared to those used in the Born polarizations. To make this more transparent, we write the spin-dependent part of the Born cross section for the orientations of the proton spin along S^N and S^T,

$$\varepsilon_2 E_2 \frac{d\sigma_T^B}{d^3k_2 d^3p_2} = \frac{\alpha^2 V^2 \eta}{4(2S_A+1)(2\pi)^3 q^4} \times \sqrt{\frac{V[q]}{z^2 - 4\tau_1\tau_2}} H_3, \tag{30}$$

$$\varepsilon_2 E_2 \frac{d\sigma_N^B}{d^3k_2 d^3p_2} = -\frac{\alpha^2 V^2 \sqrt{V[q]}}{4(2S_A+1)(2\pi)^3 q^4} H_4. \tag{31}$$

These formulas can be derived from Eqs. (20) and (21) if the $D(x_i)$ functions are taken to be the δ function, which corresponds to the radiationless process (or to the Born approximation).

4. SEMI-INCLUSIVE DEEP-INELASTIC SCATTERING ON A POLARIZED TARGET

In this section, we apply the master representation to the analysis of polarized phenomena in the semi-inclusive deep-inelastic scattering of the polarized nucleus,

$$\vec{e}^{\,-}(k_1) + \vec{A}(p_1) \longrightarrow e^-(k_2) + H(p_2) + X, \tag{32}$$

where H is an arbitrary hadron and the nucleus A has a definite vector polarization P. In this case, the leptonic tensor is the same as above (see Eqs. (3) and (4)) and the hadronic tensor has the same structure as defined by Eqs. (5) and (6), where the nucleus polarization P must be used instead of the proton spin S and (Sp_1) must be replaced with (Pp_2). We also use the notation $g_1, \ldots, g_{18}$ for the corresponding hadron structure functions.

To find the various asymmetries measured in studying the polarization phenomena, it is necessary to know the polarization-independent and polarization-dependent parts of the cross section at different orientations of the target polarization. The corresponding analysis can therefore be performed in the same way as in Section 3.

We first define parametrizations of the nucleus polarization 4-vector in terms of the 4-momenta. As a stabilized set, we can choose the longitudinal and transverse components given in [12],

$$P_\mu^l = \frac{2\tau_1 k_{1\mu} - p_{1\mu}}{M}, \qquad P_\mu^t = \frac{k_{2\mu} - (1 - y - 2xy\tau_1)k_{1\mu} - xyp_{1\mu}}{\sqrt{Vxy(1 - y - xy\tau_1)}}, \tag{33}$$

and for the normal component, we use

$$P_\mu^n = \frac{2(\mu k_1 k_2 p_1)}{\sqrt{V^3 xy(1 - y - xy\tau_1)}}. \tag{34}$$

It is easy to verify that parametrizations (33) and (34) are not changed after the substitution $k_{1,2} \longrightarrow \hat{k}_{1,2}$. In the laboratory system, this set corresponds to the longitudinal polarization directed along $\mathbf{k}_1$, the transverse polarization in the plane $(\mathbf{k}_1, \mathbf{k}_2)$, and the normal one in the plane that is perpendicular to the $(\mathbf{k}_1, \mathbf{k}_2)$ plane.

Another set of polarizations can be chosen such that the longitudinal component is along the $\mathbf{q}$ direction in the laboratory system and the transverse one is in the plane $(\mathbf{q}, \mathbf{k}_1)$. In this case, the normal component coincides with (34) and

$$P_\mu^L = \frac{2\tau_1(k_{1\mu} - k_{2\mu}) - yp_{1\mu}}{M\sqrt{y^2 + 4xy\tau_1}}, \qquad P_\mu^T = \frac{(1 + 2x\tau_1)k_{2\mu} - (1 - y - 2x\tau_1)k_{1\mu} - x(2 - y)p_{1\mu}}{\sqrt{Vx(1 - y - xy\tau_1)(y + 4x\tau_1)}}. \tag{35}$$

The transformation between sets (35) and (33) is implemented by the orthogonal matrix

$$\begin{gathered} P^L = \cos\theta_1 P^l + \sin\theta_1 P^t, \\ P^T = -\sin\theta_1 P^l + \cos\theta_1 P^t, \\ \cos\theta_1 = \frac{y(1 + 2x\tau_1)}{\sqrt{y(y + 4x\tau_1)}}, \\ \sin\theta_1 = -2\sqrt{\frac{x\tau_1(1 - y - xy\tau_1)}{y + 4x\tau_1}}. \end{gathered} \tag{36}$$

Master equation (12) can be applied to the polarization-independent part of cross section (32) and to the polarization-dependent part. Therefore, we must derive the Born cross section for the stabilized set. A simple calculation gives

$$\varepsilon_2 E_2 \frac{d\sigma_{(u)}^B}{d^3k_2 d^3p_2} = \frac{\alpha^2 V}{(2S_A + 1)(2\pi)^3 q^4} G_1. \tag{37}$$

We note that the numerical coefficient in front of G_1 is twice the coefficient in front of H_1 on the right-hand side of Eq. (23). The reason is that we do not fix the spin state of the final hadron H in this case.

The polarization-dependent part of the cross section for the longitudinal stabilized polarization is given by

$$\varepsilon_2 E_2 \frac{d\sigma_l^B}{d^3k_2 d^3p_2} = -\frac{\alpha^2 V^3 \eta}{4(2S_A + 1)M(2\pi)^3 q^4} \times [(2\tau_1 z_1 - z)G_2 - y(1 + 2x\tau_1)G_3 + 2\tau_1 G_4], \tag{38}$$

where the functions G_i, $i = 1, \ldots, 4$, can be derived from H_i by replacing the hadron structure functions h_j with g_j.

For the transverse polarization, the corresponding part of the cross section can be written as

$$\varepsilon_2 E_2 \frac{d\sigma_t^B}{d^3k_2 d^3p_2} = -\frac{\alpha^2 V^2 \eta \sqrt{Vxy(1-y-xy\tau_1)}}{4(2S_A+1)(2\pi)^3 q^4} \times \left[\frac{z_2 - xyz - z_1(1-y-2xy\tau_1)}{xy(1-y-xy\tau_1)} G_2 + 2G_3 + \frac{1+2x\tau_1}{x(1-y-xy\tau_1)} G_4\right]. \tag{39}$$

For the normal polarization, the spin-dependent part of the cross section is

$$\varepsilon_2 E_2 \frac{d\sigma_n^B}{d^3k_2 d^3p_2} = \frac{\alpha^2 V^2}{4(2S_A+1)(2\pi)^3 q^4} \times \sqrt{\frac{V}{xy(1-y-xy\tau_1)}} [\eta^2 G_2 - y(z_2(1+2x\tau_1) - z_1(1-y-2x\tau_1) - xz(2-y))G_4]. \tag{40}$$

Using master representation (12) leads to the radiatively corrected contributions (within the leading accuracy) to the cross section of process (32),

$$\varepsilon_2 E_2 \frac{d\sigma_{(u), N}}{d^3k_2 d^3p_2} = \iint \frac{dx_1 dx_2}{x_2^2} D(x_1) D(x_2) \hat{\varepsilon}_2 E_2 \frac{d\hat{\sigma}^B_{(u), n}}{d^3\hat{k}_2 d^3p_2}, \tag{41}$$

$$\varepsilon_2 E_2 \frac{d\sigma_L}{d^3k_2 d^3p_2} = \iint \frac{dx_1 dx_2}{x_2^2} D(x_1) D(x_2) \times \hat{\varepsilon}_2 E_2 \left[\cos\theta_1 \frac{d\hat{\sigma}_l^B}{d^3\hat{k}_2 d^3p_2} + \sin\theta_1 \frac{d\hat{\sigma}_t^B}{d^3\hat{k}_2 d^3p_2}\right], \tag{42}$$

$$\varepsilon_2 E_2 \frac{d\sigma_T}{d^3k_2 d^3p_2} = \iint \frac{dx_1 dx_2}{x_2^2} D(x_1) D(x_2) \times \hat{\varepsilon}_2 E_2 \left[-\sin\theta_1 \frac{d\hat{\sigma}_l^B}{d^3\hat{k}_2 d^3p_2} + \cos\theta_1 \frac{d\hat{\sigma}_t^B}{d^3\hat{k}_2 d^3p_2}\right]. \tag{43}$$

We also write the cross sections on the left-hand sides of Eqs. (42) and (43) in the Born approximation,

$$\varepsilon_2 E_2 \frac{d\sigma_L^B}{d^3k_2 d^3p_2} = \frac{\alpha^2 V^3 \eta}{4(2S_A+1)(2\pi)^3 M q^4} \times \left[\frac{yz - 2(z_1 - z_2)\tau_1}{\sqrt{y(y+4x\tau_1)}} G_2 + \sqrt{y(y+4x\tau_1)} G_3\right], \tag{44}$$

$$\varepsilon_2 E_2 \frac{d\sigma_T^B}{d^3k_2 d^3p_2} = \frac{\alpha^2 V^3 \eta \sqrt{V}}{4(2S_A+1)(2\pi)^3 q^4} \times \left[-\sqrt{\frac{y+4x\tau_1}{x(1-y-xy\tau_1)}} G_4 + \frac{xz(2-y) - z_2 + z_1(1-y) - 2x\tau_1(z_1+z_2)}{\sqrt{x(y+4x\tau_1)(1-y-xy\tau_1)}} G_2\right]. \tag{45}$$

Thus, the polarization-dependent parts of the Born cross section involve fewer hadron structure functions than the radiatively corrected cross sections.

We can also use the 4-vector p_2 to parametrize the nucleus polarization 4-vector. With the longitudinal polarization chosen along $\mathbf{p}_2$ in the laboratory system, the stabilized set can be defined with respect to the plane $(\mathbf{k}_1, \mathbf{p}_2)$ and the unstable set with respect to the plane $(\mathbf{q}, \mathbf{p}_2)$ as in Section 3; the corresponding calculations are very similar to those in Section 3. But the parametrizations used in this section look more physical and can also be used to describe the polarization phenomena in inclusive deep-inelastic events.

5. POLARIZATION TRANSFER FROM THE TARGET TO THE DETECTED PROTON

We now consider the effects of the polarization transfer from the vector polarized target to the detected proton in the process

$$\vec{e}^-(k_1) + \vec{A}(p_1) \longrightarrow e^-(k_2) + \vec{p}(p_2) + X \tag{46}$$

for the longitudinally polarized electron beam and the vector polarization of the target. In this case, the general form of the hadronic tensor is given by

$$H_{\mu\nu} = H_{\mu\nu}^{(u)} + H_{\mu\nu}^{(S)} + H_{\mu\nu}^{(W)} + H_{\mu\nu}^{(SW)}, \tag{47}$$

where $S(W)$ labels the vector polarization of the target (the spin of the detected proton). All the effects caused by the first three terms on the right-hand side of Eq. (47) were considered in previous sections, and we now investigate the radiative corrections to the hadron double-spin correlations that precisely arise due to the last term,

$$\begin{aligned} H_{\mu\nu}^{(SW)} = {} & (Sp_2)(Wp_1)[f_1\tilde{g}_{\mu\nu\nu} + f_2\tilde{p}_{1\mu}\tilde{p}_{1\nu} \\ & + f_3\tilde{p}_{2\mu}\tilde{p}_{2\nu} + f_4(\tilde{p}_1\tilde{p}_2)_{\mu\nu} + if_5[\tilde{p}_1\tilde{p}_2]_{\mu\nu}] \\ & + (Sp_2)(Wq)[f_6\tilde{g}_{\mu\nu} + f_7\tilde{p}_{1\mu}\tilde{p}_{1\nu} + f_8\tilde{p}_{2\mu}\tilde{p}_{2\nu} \\ & + f_9(\tilde{p}_1 p_2)_{\mu\nu} + if_{10}[\tilde{p}_1\tilde{p}_2]_{\mu\nu}] \\ & + (Sp_2)(W_N)[f_{11}(\tilde{p}_1 N)_{\mu\nu} + if_{12}[\tilde{p}_1 N]_{\mu\nu} \\ & + f_{13}(\tilde{p}_2 N)_{\mu\nu} + if_{14}[\tilde{p}_2 N]_{\mu\nu}] \\ & + (Sq)(Wp_1)[f_{15}\tilde{g}_{\mu\nu} + f_{16}\tilde{p}_{1\mu}\tilde{p}_{1\nu} + f_{17}\tilde{p}_{2\mu}\tilde{p}_{2\nu} \end{aligned}$$

$$\begin{aligned}
&+ f_{18}(\tilde{p}_1 p_2)_{\mu\nu} + i f_{19}[\tilde{p}_1\tilde{p}_2]_{\mu\nu}] \\
&+ (Sq)(Wq)[f_{20}\tilde{g}_{\mu\nu} + f_{21}\tilde{p}_{1\mu}\tilde{p}_{1\nu} + f_{22}\tilde{p}_{2\mu}\tilde{p}_{2\nu} \\
&\quad + f_{23}(\tilde{p}_1\tilde{p}_2)_{\mu\nu} + i f_{24}[\tilde{p}_1\tilde{p}_2]_{\mu\nu}] \\
&+ (Sq)(WN)[f_{25}(\tilde{p}_1 N)_{\mu\nu} + i f_{26}[\tilde{p}_1 N]_{\mu\nu} \\
&\quad + f_{27}(\tilde{p}_2 N)_{\mu\nu} + i f_{28}[\tilde{p}_2 N]_{\mu\nu}] \\
&+ (SN)(Wp_1)[f_{29}(\tilde{p}_1 N)_{\mu\nu} + i f_{30}[\tilde{p}_1 N]_{\mu\nu} \\
&\quad + f_{31}(\tilde{p}_2 N)_{\mu\nu} + i f_{32}[\tilde{p}_2 N]_{\mu\nu}] \\
&+ (SN)(Wq)[f_{33}(\tilde{p}_1 N)_{\mu\nu} + i f_{34}[\tilde{p}_1 N]_{\mu\nu} \\
&\quad + f_{35}(\tilde{p}_2 N)_{\mu\nu} + i f_{36}[\tilde{p}_2 N]_{\mu\nu}] \\
&+ (SN)(WN)[f_{37}\tilde{g}_{\mu\nu} + f_{38}\tilde{p}_{1\mu}\tilde{p}_{1\nu} + f_{39}\tilde{p}_{2\mu}\tilde{p}_{2\nu} \\
&\quad + f_{40}(\tilde{p}_1\tilde{p}_2)_{\mu\nu} + i f_{41}[\tilde{p}_1\tilde{p}_2]_{\mu\nu}].
\end{aligned} \tag{48}$$

Thus, the coefficients of the polarization transfer from the target to the detected proton are described, in general, by 41 structure functions. If the electron beam is unpolarized, only the symmetric part of the hadronic tensor contributes, which corresponds to double-spin (hadron–hadron) correlations in the cross section of process (46). The antisymmetric part of the hadron tensor contributes in the case of the longitudinally polarized electron beam due to triple-spin (electron–hadron–hadron) correlations.

The corresponding radiatively corrected parts of the cross section for the unstable orientations of the target nucleus polarization S^J (given by Eq. (35)) and the detected proton spin W^I (given by Eq. (17)) can be written as

$$\begin{aligned}
\varepsilon_2 E_2 \frac{d\sigma_{JI}}{d^3k_2 d^3p_2} &= \sum_{j,i} A_{Jj} B_{Ii} \\
&\times \iint \frac{dx_1 dx_2}{x_2^2} D(x_1) D(x_2) \hat{\varepsilon}_2 E_2 \frac{d\hat{\sigma}^B_{ji}}{d^3\hat{k}_2 d^3 p_2},
\end{aligned} \tag{49}$$

where the Born cross section in the integrand is defined for the stable orientations of S^j (given by Eqs. (33) and (34)) and W^i (given by Eqs. (14) and (15)) and depends on the shifted variables

$$\hat{\varepsilon}_2 E_2 \frac{d\hat{\sigma}^B_{ji}}{d^3\hat{k}_2 d^3 p_2} = \hat{\varepsilon}_2 E_2 \frac{d\sigma^B(\lambda, S^j, W^i, \hat{k}_1, \hat{k}_2, p_1, p_2)}{d^3\hat{k}_2 d^3 p_2}.$$

In accordance with the calculations in Sections 3 and 4, the matrices A_{Jj} and B_{Ii} are given by

$$\begin{aligned}
A_{Jj} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{pmatrix}, \\
B_{Ii} &= \begin{pmatrix} \cos\theta_1 & \sin\theta_1 & 0 \\ -\sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix},
\end{aligned} \tag{50}$$

$$I, J = L, T, N; \quad i, j = l, t, n.$$

If we write the hadron–hadron spin correlations entering the Born cross section as

$$\varepsilon_2 E_2 \frac{d\sigma^B_{ji}}{d^2k_2 d^3p_2} = \frac{\alpha^2 V^4 X_{ji}}{16(2\pi)^3 2(2S_A+1) q^4}, \tag{51}$$

the quantities X_{ji} can be written as

$$\begin{aligned}
X_{ll} &= 2\sqrt{\frac{f\tau_1}{\tau_2}}\Big\{\eta^2(R_{29} + \xi R_{33}) \\
&+ \frac{2}{V^2\tau_1}[b(F_1 + \xi F_6) - d(F_{15} + \xi F_{20})]\Big\},
\end{aligned} \tag{52}$$

$$\begin{aligned}
X_{lt} &= \eta^2\sqrt{\frac{f}{\tau_1[1]}}\Big[bR_{11} - dR_{25} + 2\tau_1 F_{37} \\
&- \frac{2\psi}{\eta^2 V^2 f}\sqrt{[1]}(2bF_6 - 2dF_{20} + \eta^2 V^2 \tau_1 R_{33})\Big],
\end{aligned} \tag{53}$$

$$\begin{aligned}
X_{ln} &= \frac{\eta}{\sqrt{\tau_1}}\Big[\psi(bR_{11} - dR_{25} + 2\tau_1 F_{37}) \\
&+ \frac{2}{V^2\sqrt{[1]}}(2bF_6 - 2dF_{20} + \eta^2 V^2 \tau_1 R_{33})\Big],
\end{aligned} \tag{54}$$

$$\begin{aligned}
X_{tl} &= \sqrt{\frac{f}{r\tau_2}}\Big\{\eta^2 d(R_{29} + \xi R_{33}) \\
&+ \frac{4}{V^2}[2r(F_{15} + \xi F_{20}) + \zeta(F_1 + \xi F_6)]\Big\},
\end{aligned} \tag{55}$$

$$\begin{aligned}
X_{tt} &= \eta^2\sqrt{\frac{f}{r[1]}}\Big[\zeta R_{11} + 2rR_{25} + dF_{37} \\
&- \frac{\psi\sqrt{[1]}}{\eta^2 V^2 f}(\eta^2 V^2 d R_{33} + 4\zeta F_6 + 8rF_{20})\Big],
\end{aligned} \tag{56}$$

$$\begin{aligned}
X_{tn} &= \frac{\eta}{\sqrt{r}}\Big[\psi(\zeta R_{11} + 2rR_{25} + dE_{37}) \\
&+ \frac{1}{V^2\sqrt{[1]}}(\eta^2 V^2 d R_{33} + 4\zeta F_6 + 8rF_{20})\Big],
\end{aligned} \tag{57}$$

$$X_{nl} = \eta\sqrt{\frac{f}{r\tau_2}}\left[\eta_1(R_{29}+\xi R_{33}) - \frac{4}{V^2}(F_1+\xi F_6)\right], \quad (58)$$

$$\begin{aligned} X_{nt} &= \frac{\eta}{\sqrt{fr}} \\ &\times\left[\psi\left(\frac{4}{V^2}F_6 - \eta_1 R_{33}\right) + \frac{f}{\sqrt{[1]}}(\eta_1 F_{37} - \eta^2 R_{11})\right], \end{aligned} \quad (59)$$

$$\begin{aligned} X_{nn} &= -\frac{\eta^2}{\sqrt{r}} \\ &\times\left[\frac{1}{\sqrt{[1]}}\left(\frac{4}{V^2}F_6 - \eta_1 R_{33}\right) - \psi\left(\frac{\eta_1}{\eta^2}F_{37} - R_{11}\right)\right], \end{aligned} \quad (60)$$

where we used the notation

$$b = 2z_1\tau_1 - z, \quad d = y(1+2x\tau_1),$$

$$f = z^2 - 4\tau_1\tau_2, \quad r = xy(1-y-xy\tau_1),$$

$$\zeta = z_2 - z_1(1-y-2xy\tau_1) - xyz,$$

$$\xi = \frac{z(z_1-z_2) - 2y\tau_2}{z^2 - 4\tau_1\tau_2}.$$

The functions R_l and F_l entering the expressions for X_{ji} are defined by the hadron structure functions f_n in Eq. (48) as

$$\begin{aligned} R_l &= (2-y)f_l + (z_1+z_2)f_{l+2} \\ &+ \frac{\lambda}{\eta}(\eta_1 f_{l+1} + \eta_2 f_{l+3}), \end{aligned} \quad (61)$$

$$\begin{aligned} F_l &= -\frac{2xy}{V}f_l + (1-y-xy\tau_1)f_{l+1} \\ &+ (z_1z_2 - xy\tau_2)f_{l+2} \\ &+ (z_2 + z_1(1-y) - xyz)f_{l+3} - \lambda\eta f_{l+4}. \end{aligned} \quad (62)$$

6. HADRONIC VARIABLES

There exists the experimental possibility of measuring the total 4-momentum of the hadron system X instead of recording the scattered electron in semi-inclusive reactions. In such experiments, the momentum q_h of the heavy intermediate photon that probes the hadron structure can be determined explicitly. The corresponding set of dynamical variables is usually referred to as the hadronic one.

For the hadronic variables, we must eliminate the phase space of the scattered electron and introduce the heavy photon phase space using the identities

$$\frac{d^3k_2}{\varepsilon_2} = 2x_2^2 x_h \frac{d^4q_h}{Q_h^2}\delta(x_1 - x_h),$$

$$\frac{d^4q_h}{Q_h^2} = \frac{dQ_h^2 dx_h dy_h dz_h}{4x_h^2|\eta_h|}, \quad x_h = -\frac{Q_h^2}{2k_1q_h},$$

$$y_h = \frac{2p_1q_h}{V}, \quad z_h = \frac{2p_2q_h}{V}, \quad Q_h^2 = -q_h^2, \quad (63)$$

$$\begin{aligned} \eta_h^2 = \frac{Q_h^2}{V}\Bigg[&(4\tau_1\tau_2 - z^2)\frac{Q_h^2}{x_h^2 V} + 2\left(1 - \frac{y_h}{x_h}\right)(zz_1 - 2\tau_2) \\ &+ 2\left(z_1 - \frac{z_h}{x_h}\right)(z - 2z_1\tau_1)\Bigg] - (z_h - z_1y_h)^2. \end{aligned}$$

Therefore, combining this with representation (3) for the leptonic tensor and also bearing in mind that the hadronic tensor is independent of x_2, we can express the quantity $L_{\mu\nu}d^3k_2/\varepsilon_2$ through the hadronic variables as

$$\begin{aligned} \frac{d^3k_2}{\varepsilon_2}L_{\mu\nu} &= \frac{D(x_h, Q_h^2)}{x_h^2}L^B_{\mu\nu}(\hat{k}_1, \hat{k}_1 - q_h, \lambda) \\ &\times\frac{dx_h dy_h dz_h dQ_h^2}{2|\eta_h|}. \end{aligned} \quad (64)$$

We note that, for the events with the undetected scattered electron, the lower integration limit with respect to x_2 in Eq. (3) is equal to 0. In accordance with the Kinoshita–Lee–Nauenberg theorem [22], the mass singularities caused by the final-state radiation must disappear in this case. In the language of the electron structure functions, this fact exhibits itself due to the relation

$$\int_0^1 D(x, Q^2)dx = 1,$$

which was used in writing Eq. (64).

In the Born approximation, the lepton tensor can be rewritten as

$$\begin{aligned} L^B_{\mu\nu}(k_1, k_1 - q_h) &= 2(k_1q_h)\tilde{g}_{\mu\nu} \\ &+ 4\tilde{k}_{1\mu}\tilde{k}_{1\nu} - 2i\lambda(\mu\nu k_1 q_h), \end{aligned} \quad (65)$$

and the physically justified parametrizations for S^j in process (1) and for P^j in process (32) remain stable with respect to the scaling transformation $k_1 \longrightarrow x_n k_1$. For example, one set can be chosen as in Eqs. (14) and (15) and the other as

$$S^L_{h\mu} = S^l_\mu, \quad S^T_{h\mu} = \frac{(z^2 - 4\tau_1\tau_2)q_{h\mu} + (2z_h\tau_1 - zy_h)p_{2\mu} + (2y_h\tau_2 - zz_h)p_{1\mu}}{\sqrt{V(z^2 - \tau_1\tau_2)[q_h]}},$$
$$S^N_{h\mu} = \frac{2(\mu q_h p_1 p_2)}{\sqrt{V^3[q_h]}}, \quad [q_h] = zz_hy_h + \frac{Q_h^2}{V}(z^2 - 4\tau_1\tau_2) - z_h^2\tau_1 - y_h^2\tau_2, \tag{66}$$

with the transverse component belonging to the plane ($\mathbf{q}_h$, $\mathbf{p}_2$) in the laboratory system.

Two physical sets of the target polarizations, each with the normal component perpendicular to the plane ($\mathbf{k}_1$, $\mathbf{q}_h$), can be chosen as

$$P^l_{h\mu} = \frac{2\tau_1 k_{1\mu} - p_{1\mu}}{M},$$
$$P^t_{h\mu} = \left[-q_{h\mu} + \left(y_h + \frac{2Q_h^2\tau_1}{x_hV}\right)k_{1\mu} - \frac{Q_h^2}{x_hV}p_{1\mu}\right]K^{-1}, \tag{67}$$
$$P^n_{h\mu} = -\frac{2(\mu k_1 q_h p_1)}{VK},$$

with the longitudinal component along $\mathbf{k}_1$ in the laboratory system, and

$$P^L_{h\mu} = \frac{2\tau_1 q_{h\mu} - y_h p_{1\mu}}{MG},$$
$$P^T_{h\mu} = \left[\left(y_h^2 + 4\tau_1\frac{Q_h^2}{V}\right)k_{1\mu} - \left(y_h + \frac{2Q_h^2\tau_1}{x_hV}\right)q_{h\mu} - \frac{Q_h^2}{V}\left(2 - \frac{y_h}{x_h}\right)p_{1\mu}\right](KG)^{-1}, \tag{68}$$
$$P^N_{h\mu} = P^n_{h\mu}, \quad K = \sqrt{Q_h^2\left(1 - \frac{y_h}{x_h} - \frac{Q_h^2\tau_1}{x_h^2V}\right)},$$
$$G = \sqrt{y_h^2 + 4\frac{Q_h^2\tau_1}{V}},$$

with the longitudinal component along $\mathbf{q}_h$. The different components of P_h^J in the laboratory system are

$$P_h^L = (0, \mathbf{n}_q), \quad P_h^T = \left(0, \frac{\mathbf{n}_1 - (\mathbf{n}_1 \cdot \mathbf{n}_q)\mathbf{n}_q}{\sqrt{1 - (\mathbf{n}_1 \cdot \mathbf{n}_q)^2}}\right),$$
$$P_h^N = \left(0, \frac{\mathbf{n}_q \times \mathbf{n}_1}{\sqrt{1 - (\mathbf{n}_1 \cdot \mathbf{n}_q)^2}}\right),$$
$$\mathbf{n}_q = \frac{\mathbf{q}_h}{|\mathbf{q}_h|}, \quad \mathbf{n}_1 = \frac{\mathbf{k}_1}{|\mathbf{k}_1|}.$$

All these sets of the proton spin and target polarizations given by Eqs. (66), (67), and (68) are stable with respect to the initial-state collinear radiation. This can be verified by replacing k_1 with x_hk_1, which implies

$$k_1 \longrightarrow x_hk_1, \quad x_h \longrightarrow 1, \quad y_h \longrightarrow \frac{y_h}{x_h},$$
$$z_h \longrightarrow \frac{z_h}{x_h}, \quad z \longrightarrow \frac{z}{z_h}, \quad V \longrightarrow x_hV, \quad \tau_{1,2} \longrightarrow \frac{\tau_{1,2}}{x_h}. \tag{69}$$

To make the invariance of P^j ($j = l, t, n$) and P^J ($J = L, T, N$) under replacement (69) more transparent, we express x_h in terms of Q_h^2 and (k_1q_h). Then, e.g.,

$$K = \sqrt{Q_h^2 + y_h2(k_1q_h) - \frac{4(k_1q_h)^2\tau_1}{V}},$$

and it is easy to see that this quantity is not changed under substitution (69). We also note that the quantity η_h can be derived using the rule

$$\eta_h = x_h\eta^*,$$

where η^* is determined from η with xy replaced by Q_h^2/V, z_2 replaced with $z_1 - z_h$, and with the subsequent replacement (69).

For hadronic variables, the cross section for both the spin-independent and the spin-dependent parts can therefore be written as

$$E_2\frac{d\sigma^j}{d^3p_2dQ_h^2dx_hdy_hdz_h} = \frac{D(x_h, Q_h^2)}{x_h^2}E_2\frac{d\hat{\sigma}_j^B}{d^3p_2dQ_h^2d\hat{y}_hd\hat{z}_h}, \tag{70}$$

where

$$E_2\frac{d\hat{\sigma}_j^B}{d^3p_2dQ_h^2d\hat{y}_hd\hat{z}_h} = \frac{\alpha^2C}{(2\pi)^3(2S_A + 1)\hat{V}Q_h^4 2|\eta^*|} \times L_{\mu\nu}(\hat{k}_1, \hat{k}_1 - q_h, \lambda)H_{\mu\nu}(q_h, p_1, p_2; S^j(P^j))$$

and C is equal to 1/2 or 1 for the respective process (1) or (32).

Representation (70) shows that using the hadron variables allows us to tag the initial-state radiated photon. Indeed, for a fixed 4-momentum P_x, we can reconstruct the 4-momentum q_h and, consequently, the variable x_h that is the energy fraction of the photon radiated by the initial electron (see Eq. (63)).

The Born cross section on the right-hand side of Eq. (70) has the form that is very similar to the corresponding cross section for the leptonic variables. We can formulate the following rules to write it:

(i) change the phase space differentials on the left-hand sides of the expressions valid for the leptonic variables as

$$\frac{\varepsilon_2}{d^3k_2} \longrightarrow \frac{2|\eta_{1h}|}{dQ_h^2dy_hdz_h}, \quad \eta_{1h} = \eta_h(x_h = 1);$$

(ii) apply the substitution

$$xy \longrightarrow \frac{Q_h^2}{V}, \quad y \longrightarrow y_h, \quad z_2 \longrightarrow z_1 - z_h$$

to the right-hand sides.

These rules lead, e.g., to the formula for the spin-dependent part of the cross section of process (1) in the case of the longitudinal polarization (which follows from Eq. (25))

$$E_2\frac{d\sigma_L^B}{d^3p_2dQ_h^2dy_hdz_h} = -\frac{\alpha^2V^3\eta_{1h}\sqrt{z^2-4\tau_1\tau_2}}{8m(2S_A+1)(2\pi)^3Q_h^42|\eta_{1h}|} \times\left[H_2^{(h)} + \frac{zz_h-2y_h\tau_2}{z^2-4\tau_1\tau_2}H_3^{(h)}\right], \tag{71}$$

$$H_2^{(h)} = (2-y_h)h_6 + (2z_1-z_h)h_8 + \frac{\lambda}{\eta_{1h}}(\eta_1^{(h)}h_7 + \eta_2^{(h)}h_9),$$

$$\eta_1^{(h)} = \frac{Q_h^2}{V}[2(2z_1-z_h)\tau_1 - z(2-y_h)] + z_1y_h^2 - z_hy_h,$$

$$\eta_2^{(h)} = \frac{Q_h^2}{V}[z(2z_1-z_h) - 2(2-y_h)\tau_2] - z_h^2 + z_1z_hy_h,$$

where $H_3^{(h)}$ is derived from $H_2^{(h)}$ by the replacement $h_i \longrightarrow h_{i+4}$.

For the normal target polarization that follows from Eq. (40), the spin-dependent part of the cross section of process (32) is given by

$$\begin{aligned}&E_2\frac{d\sigma_N^B}{d^3p_2dQ_h^2dy_hdz_h}\\ &= -\frac{\alpha^2V^3}{4(2S_A+1)(2\pi)^3Q_h^4K(x_h=1)2|\eta_{1h}|}\\ &\times\left\{\eta_{1h}^2G_2^{(h)} - \left[y_h(z_1y_h-z_h)\right.\right.\\ &\left.\left.+\frac{Q_h^2}{V}(2\tau_1(2z_1-z_h) - z(2-y_h))\right]G_4^{(h)}\right\}.\end{aligned} \tag{72}$$

The remaining spin-dependent and spin-independent parts of the cross sections for processes (1) and (32) can be obtained totally similarly using the above rules and the results in Sections 3 and 4.

The variable x_h characterizes the inelasticity of the initial-state electron and is equal to 1 in the absence of radiation. The electron structure function $D(x_h, Q_h^2)$ has a singularity at $x_h = 1$, and representation (70) shows that this singularity is such that

$$\lim_{x_h\to 1} D(x_h, Q_h^2)dx_h = 1, \tag{73}$$

because, in this limiting case, the left-hand side of Eq. (70) multiplied by dx_h must coincide with the Born cross section.

7. CONCLUSION

In this paper, we consider radiative corrections to the polarization observables in a wide class of semi-inclusive deep-inelastic processes. We restrict ourselves to the leading-log accuracy and neglect the contribution of the pair production in the singlet channel. This allows us to write compact formulas for the radiatively corrected spin-independent and spin-dependent parts of the corresponding cross sections in the form of the Drell–Yan representation in electrodynamics by means of the electron structure functions. The parametrization of the hadron spin 4-vectors in terms of the particle 4-momenta is very important in the calculations. If the momentum of the intermediate photon that probes the hadron structure is determined in terms of the hadronic variables, the traces of the final-state radiation disappear in the final result within the adopted approximation.

In practice, the corrections can be computed adopting some specific model for the structure functions. The correction then acquires some model dependence that can contribute to the systematic error in experimental measurements. Another possibility is related to some iteration procedure, where the fit of the processed experimental data is used for the chosen model. We note that the obtained leading-log formulas have a partly factorized form, which is very convenient for this procedure. The examples for the DIS case can be found in [20, 23].

Apart from the classes of experiments discussed above, the results can also be adopted to exclusive electroproduction processes, where the unobservable hadron state is one particle. In this case, the structure functions involve an additional δ function, and therefore, some analytical manipulations could be necessary.

Accuracy higher than the leading one sometimes becomes necessary. To go beyond the leading accuracy, one must modify the master representations. This modification affects both the electron structure function and the cross section (the hard part) that depends on the shifted variables. To improve the hard part, it suffices to take the radiation of a single additional noncollinear photon into account and to add the nonleading part of the one-loop correction. The corresponding procedure is described in [21] for the unpolarized deep-inelastic scattering and in the second of [20] for the quasi-elastic polarized electron–proton scattering. To be complete, one must also improve the structure functions by adding the second-order next-to-leading contributions of the double collinear photon emission and the pair production. The nonleading contributions to the D function caused by the one-loop corrected collinear single-photon emission and the two-loop correction must also be added properly. These contributions are different for symmetric and asymmetric parts of the leptonic tensor and can be extracted from the results in [16] (for the two-loop correction, see [24]). In this case, we must therefore distinguish between D and D_λ at the level of the nonsinglet channel contribution. The specific calculations will be done elsewhere.

REFERENCES

1. B. W. Filippone and X. Ji, hep-ph/0101224.
2. S. Frullani and J. Mougey, Adv. Nucl. Phys. **14**, 1 (1984).
3. J. Carlson and R. Schiavilla, Rev. Mod. Phys. **70**, 743 (1998).
4. J. Collins, Nucl. Phys. B **396**, 161 (1993); D. Yu. Ivanov, Phys. Rev. D **53**, 3564 (1996); J. Levelt and P. J. Mulders, Phys. Rev. D **49**, 96 (1994); Phys. Lett. B **338**, 357 (1994); R. D. Tangerman and P. J. Mulders, Phys. Lett. B **352**, 129 (1995); A. Kotzinian, Nucl. Phys. B **441**, 234 (1995).
5. D. de Florian, G. M. Shore, and G. Veneziano, Preprint CERN-TH/97-297 (1997); X. Ji, Phys. Rev. D **49**, 114 (1994); F. E. Close and R. G. Milner, Phys. Rev. D **44**, 3691 (1991); L. L. Frankfurt, M. I. Strikman, L. Mankiewicz, *et al.*, Phys. Lett. B **230**, 141 (1989); X. Artru and M. Mekhfi, Nucl. Phys. A **532**, 351 (1991).
6. S. Boffi, C. Giusti, and F. D. Pacati, Phys. Rep. C **226**, 1 (1993); J. J. Kelly, Adv. Nucl. Phys. **23**, 75 (1996).
7. M. P. Rekalo, G. I. Gakh, and A. P. Rekalo, Phys. Lett. **166**, 27 (1986); C. Giusti and F. D. Pacati, Nucl. Phys. A **504**, 685 (1989); A. Picklesimer and J. W. van Orden, Phys. Rev. C **40**, 290 (1989); C. Y. Cheng and R. M. Woloshyn, Phys. Lett. B **127B**, 147 (1983); M. P. Rekalo, G. I. Gakh, and A. P. Rekalo, Ukr. Fiz. Zh. **32**, 805 (1987); G. I. Gakh, Ukr. Fiz. Zh. **35**, 967 (1990).
8. A. V. Soroko and N. M. Shumeiko, Yad. Fiz. **49**, 1348 (1989) [Sov. J. Nucl. Phys. **49**, 838 (1989)]; Yad. Fiz. **53**, 1015 (1991) [Sov. J. Nucl. Phys. **53**, 628 (1991)]; I. Akushevich, Eur. Phys. J. C **8**, 457 (1999); I. Akushevich, N. Shumeiko, and A. Soroko, Eur. Phys. J. C **10**, 681 (1999).
9. I. Schienbein, Phys. Rev. D **59**, 013001 (1999).
10. J. A. Templon, C. E. Vellidis, R. E. Florizone, and A. J. Sarty, Phys. Rev. C **61**, 014607 (2000).
11. V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15**, 781 (1972) [Sov. J. Nucl. Phys. **15**, 438 (1972)]; Yad. Fiz. **15**, 1218 (1972) [Sov. J. Nucl. Phys. **15**, 675 (1972)].
12. I. V. Akushevich, A. V. Afanas’ev, and N. P. Merenkov, submitted to J. Phys. G.
13. A. V. Afanasev, I. Akushevich, and N. P. Merenkov, hep-ph/0009273.
14. S. D. Drell and T. M. Yan, Phys. Rev. Lett. **25**, 316 (1970).
15. S. Malov, K. Wijesooriya, F. T. Baker, *et al.*, nucl-ex/0001007; Phys. Rev. C **62**, 057302 (2000).
16. I. V. Akushevich, A. B. Arbuzov, and E. A. Kuraev, Phys. Lett. B **432**, 222 (1998); M. I. Konchatnij and N. P. Merenkov, Pis’ma Zh. Éksp. Teor. Fiz. **69**, 845 (1999) [JETP Lett. **69**, 893 (1999)]; G. I. Gakh, M. I. Konchatnij, and N. P. Merenkov, Pis’ma Zh. Éksp. Teor. Fiz. **71**, 328 (2000) [JETP Lett. **71**, 227 (2000)]; M. I. Konchatnij, N. P. Merenkov, and O. N. Schekhovzova, hep-ph/9903384; Zh. Éksp. Teor. Fiz. **118**, 5 (2000) [JETP **91**, 1 (2000)].
17. E. A. Kuraev and V. S. Fadin, Yad. Fiz. **41**, 733 (1985) [Sov. J. Nucl. Phys. **41**, 466 (1985)]; F. A. Berends, W. L. van Neervan, and G. J. H. Burgers, Nucl. Phys. B **297**, 429 (1988).
18. A. B. Arbuzov, V. S. Fadin, E. A. Kuraev, *et al.*, Nucl. Phys. B **485**, 457 (1997); S. Jadach, M. Skrzypek, and B. F. L. Ward, Phys. Rev. D **47**, 3733 (1993).
19. V. N. Baier, V. S. Fadin, and V. A. Khoze, Nucl. Phys. B **65**, 381 (1973).
20. I. V. Akushevich and N. M. Shumeiko, J. Phys. G **20**, 513 (1994); I. V. Akushevich, A. V. Afanas’ev, and N. P. Merenkov, hep-ph/0102086; submitted to Phys. Rev. D.
21. E. A. Kuraev, N. P. Merenkov, and V. S. Fadin, Yad. Fiz. **47**, 1593 (1988) [Sov. J. Nucl. Phys. **47**, 1009 (1988)].
22. T. Kinoshita, J. Math. Phys. B **271**, 267 (1962); T. D. Lee and M. Nauenberg, Phys. Rev. B **133**, 1549 (1964).
23. I. Akushevich, A. Ilichev, N. Shumeiko, *et al.*, Comput. Phys. Commun. **104**, 201 (1997).
24. R. Barbiery, J. A. Mignaco, and E. Remiddi, Nuovo Cimento A **11**, 824 (1972).

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 462–470.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 529–538.
Original Russian Text Copyright © 2001 by Kozlovskiĭ.

ATOMS, SPECTRA, RADIATION

Photon Emission from a Two-Level Atom Moving in a Cavity

A. V. Kozlovskiĭ*
Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 117924 Russia
**e-mail: kozlovsk@sci.lebedev.ru*

Received February 6, 2001

Abstract—The interaction of a two-level atom uniformly moving along a classical trajectory with a high-Q cavity quantum mode is analyzed. The dressed-state method is used to derive a recurrence formula for the transition probability of the atom with photon emission; the temporal dynamics of this probability qualitatively depends on the Doppler shift of the atomic transition frequency, on the Rabi frequency of the atom–field system, and on the detuning of the atomic transition frequency from the field mode frequency. The emission dynamics of a moving atom is very sensitive to the detuning. Rabi-type oscillations with a frequency equal to the Doppler shift can arise under certain conditions. At resonance, the emission probability of a moving atom can considerably exceed the emission probability of an atom at rest. A plane-parallel-mirror cavity and a confocal spherical-mirror cavity are considered. It is shown that the peculiarities of Doppler–Rabi oscillations must be taken into account in micromaser theory. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The effects attributable to the motion of the center of mass of an emitting atom or molecule include well-studied phenomena such as inhomogeneous spectral-line broadening for an ensemble of particles (Doppler broadening), hole burning in the frequency distribution of population inversion for a two-level system, and the Lamb dip in the output power dependence on the working frequency of a gas laser. The influence of the position of an atom interacting with the field in a cavity also manifests itself in spatial hole burning, which plays a major role in the generation dynamics of a multimode laser.

For a moving atom placed in a cavity whose eigenmodes are standing waves, both the Doppler shift of the atomic transition frequency and the spatial field dependence must be taken into account, which makes it much more difficult to theoretically study the above phenomena. Using the Jaynes–Cummings model [1], which describes the interaction of a two-level atom with a single-mode field, and extending it to a moving atom allows a number of mathematical difficulties that arise in this case to be overcome. At present, only a few papers are devoted to solving the quantum-mechanical problem of the emission of a classically moving atom. Most current studies of atomic motion in a field are concerned with ultracold atoms interacting with a cavity field and involve laser cooling. In this case, the atomic center-of-mass motion must be considered in terms of quantum mechanics with recoil. This adds further complications to the problem, which have not yet been completely overcome. However, even when the particle velocity and mass are large enough to use classical center-of-mass trajectories, the papers are few in number. By numerically solving the Schrödinger equation, Meystre [2] obtained separate data on the probability of an atom moving along a classical trajectory in a cavity being in an initial state. The author considered time variations of this probability for various Doppler shifts of the cavity mode frequency in a coordinate system associated with the atom. In [3, 4], the authors calculated fluorescence spectra for a two-level atom moving along a classical trajectory when the atomic transition frequency was equal to the cavity field mode frequency. In [5], the atomic center-of-mass motion was considered from a quantum-mechanical standpoint.

In this paper, we use the dressed-state method to derive a simple, convenient (for calculations) recurrence formula for the transition probability of an excited atom moving along a classical trajectory with photon emission into a high-Q cavity mode. An analysis of the data obtained with this formula has revealed qualitative dependences of the temporal field oscillations inside the cavity on the Doppler atomic frequency shift and on the detuning of the atomic and field frequencies. We show that, depending on the relations between system parameters, the periodic energy exchange between atom and field can be fundamentally different from the case of an atom at rest and qualitatively similar to ordinary Rabi oscillations. Further development of the theory for the micromaser [4–7], whose experimental studies are currently being carried out [8–10], requires that the motion be taken into account.

2. TRANSITION PROBABILITY

We consider the Hamiltonian of a nonconservative system composed of a two-level atom uniformly moving along a classical trajectory and a single mode of a quantized cavity electromagnetic field in the form

$$H(t) = \hbar\omega_c a^+ a + \frac{\hbar\omega_a}{2}\sigma^z + \hbar g(t)B, \qquad (1)$$

where the atom–field interaction operator in the rotating-wave approximation,

$$B = a^+\sigma^- + \sigma^+ a,$$

is written by using the following operators for the transition between upper ($|\uparrow\rangle$) and lower ($|\downarrow\rangle$) atomic states: $\sigma^- = |\downarrow\rangle\langle\uparrow|$ and $\sigma^+ = |\uparrow\rangle\langle\downarrow|$; the inversion operator is $\sigma^z = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|$; and a^+ and a are the field production and destruction operators, respectively. The coupling parameter of an atom moving along trajectory $x(t)$ with the field is

$$g(t) = g_0\cos[k_c x(t)], \quad g_0 = d\sqrt{\frac{2\pi\omega_a}{\hbar V}}, \qquad (2)$$

where ω_c and $k_c = \omega_c/c$ are the frequency and wave vector of the cavity mode, respectively; d is the dipole moment of the two-level atomic transition; and V is the cavity volume. This form of coupling parameter is characterized by the spatial field distribution in a plane-parallel-mirror cavity when the atom moves along the cavity axis. For the atomic center of mass uniformly moving at velocity v considered here, the coupling parameter is

$$\begin{aligned} g(t) &= g_0\cos(\Omega_D t + \Delta\Omega t), \\ \Omega_D &= \frac{v}{c}\omega_a, \quad \Delta\Omega = \frac{v}{c}\Delta\omega, \end{aligned} \qquad (3)$$

where Ω_D is the Doppler shift of the atomic transition frequency, and $\Delta\omega = \omega_c - \omega_a$ is the detuning of the cavity mode frequency from the atomic transition frequency.

We will also consider atomic motion across a confocal spherical-mirror cavity. In this case, the coupling parameter is

$$g(x(t)) = g_0\exp\left[-\left(\frac{x(t)}{w_0}\right)^2\right]; \qquad (4)$$

the coordinate origin is assumed to be at the cavity center. The quantity w_0 is the waist length of the Gaussian field distribution.

Hamiltonian (1) can be written as a sum,

$$H(t) = \hbar[C_1 + C_2(t)], \qquad (5)$$

of the terms

$$C_1 = \omega_c\left(a^+a + \frac{\sigma^z}{2}\right), \quad C_2(t) = -\frac{\Delta\omega}{2}\sigma^z + g(t)B. \qquad (6)$$

It is easy to verify that the operators C_1 and $C_2(t)$ commute with one another for any time t:

$$[C_1, C_2(t)] = 0. \qquad (7)$$

By applying the operators C_1 and $C_2(t)$ to the basis of state vectors $|n, j\rangle \equiv |n\rangle|j\rangle$, where $|n\rangle$ is the Fock basis of field states and $j = \downarrow, \uparrow$, it is easy to see that C_1 is diagonal in the basis $|n, j\rangle$ with the eigenvalue spectrum $\omega_c(n \pm 1/2)$, while $C_2(t)$ is nondiagonal. Using the basis of eigenvectors for C_1, we could find a basis in which both operators C_1 and $C_2(t)$ would be diagonal [1]. In this case, the Hamiltonian $H(t)$ would clearly also be diagonal. As such a basis, we use the following linear combinations of elements from the basis $|n, j\rangle$:

$$\begin{aligned} |+, n\rangle_t &= \cos\theta_n(t)|n+1, \downarrow\rangle + |n, \uparrow\rangle\sin\theta_n(t), \\ |-, n\rangle_t &= -\sin\theta_n(t)|n+1, \downarrow\rangle + |n, \uparrow\rangle\cos\theta_n(t), \end{aligned} \qquad (8)$$

where θ_n is the state mixing parameter.

In the absence of a field in the cavity at initial time $t = 0$, we have $n = 0$, $|+, 0\rangle_t$ and $|-, 0\rangle_t$. For the ground state, when there are no field quanta and when the atom is in a lower state, we have

$$C_1|0, \downarrow\rangle = -\frac{\omega_c}{2}|0, \downarrow\rangle.$$

Applying the operator C_1 to the elements of this basis of dressed states yields

$$C_1|\pm, n\rangle_t = \omega_c\left(n + \frac{1}{2}\right)|\pm, n\rangle_t; \qquad (9)$$

i.e., $|\pm, n\rangle_t$ are the degenerate eigenstates of this operator. Let us apply the operator $C_2(t)$ to $|\pm, n\rangle_t$ and require that $C_2(t)$ be diagonal in this basis, i.e.,

$$C_2(t)|\pm, n\rangle_t = \Omega^R_{\pm, n}(t)|\pm, n\rangle_t. \qquad (10)$$

Using definition (8), we then obtain from (10) with the aid of (6) for the Rabi frequency

$$\begin{aligned} &\Omega^R_{+, n}(t) = -\Omega^R_{-, n}(t) \\ &= \sqrt{\left(\frac{\Delta\omega}{2}\right)^2 + g^2(t)(n+1)} \equiv \Omega^R_n(t); \end{aligned} \qquad (11)$$

for the mixing angles in (8), we have

$$\theta_n(t) = \frac{1}{2}\arctan\left[\frac{2g(t)}{\Delta\omega}\sqrt{n+1}\right]. \qquad (12)$$

For the ground state,

$$C_2(t)|0, \downarrow\rangle = \frac{\Delta\omega}{2}|0, \downarrow\rangle.$$

In the basis of dressed states that diagonalize the Hamiltonian $H(t)$, we have

$$\begin{aligned} H(t)|\pm, n\rangle_t &= \left[\hbar\omega_c\left(n + \frac{1}{2}\right) \pm \hbar\Omega^R_n(t)\right]|\pm, n\rangle_t \\ &\equiv \hbar\Omega_{\pm, n}(t)|\pm, n\rangle_t, \end{aligned} \qquad (13)$$

$$H(t)|0,\downarrow\rangle = -\frac{\hbar\omega_a}{2}|0,\downarrow\rangle. \tag{14}$$

From (8) and from the property of the basis of dressed states being orthonormal, we derive a relation between the bases $|\pm, n\rangle_t$ and $|n, j\rangle$ at a fixed time:

$$\begin{aligned} |n+1,\downarrow\rangle &= \cos\theta_n(t)|+,n\rangle_t - \sin\theta_n(t)|-,n\rangle_t, \\ |n,\uparrow\rangle &= \sin\theta_n(t)|+,n\rangle_t + \cos\theta_n(t)|-,n\rangle_t. \end{aligned} \tag{15}$$

The elements of the parametrically time-dependent basis of dressed states derived in this way for different times are not orthonormal, which must be borne in mind in subsequent calculations. The relations between the elements of the basis of dressed states for various times $t_1 \neq t_2$ are

$$\begin{aligned} |+,n\rangle_{t_1} &= \cos[\theta_n(t_2) - \theta_n(t_1)]|+,n\rangle_{t_2} \\ &\quad - \sin[\theta_n(t_2) - \theta_n(t_1)]|-,n\rangle_{t_2}, \\ |-,n\rangle_{t_1} &= \sin[\theta_n(t_2) - \theta_n(t_1)]|+,n\rangle_{t_2} \\ &\quad + \cos[\theta_n(t_2) - \theta_n(t_1)]|-,n\rangle_{t_2}. \end{aligned} \tag{16}$$

In the Schrödinger representation, the time evolution of an atom interacting with a field is described by the propagator (evolution operator)

$$U(t, t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} H(t')U(t', t_0)dt', \tag{17}$$

with the state vector at time t being

$$|\psi(t)\rangle = U(t, t_0)|\psi(t_0)\rangle. \tag{18}$$

Let us break down the finite time interval $[t_0, t]$ into a large number $M \gg 1$ of small segments Δt and assume that the time dependence of the system Hamiltonian within each of these segments may be disregarded. The time evolution of the system in each of these segments is then described by the propagator

$$\begin{aligned} |\psi(t_j)\rangle &= \exp\left[-\frac{i}{\hbar}H(t_j)\Delta t\right]|\psi(t_{j-1})\rangle, \\ t_j &= t_{j-1} + \Delta t, \quad j = 1, M, \end{aligned} \tag{19}$$

and the state vector of the system at the final time instant t is

$$\begin{aligned} |\psi(t)\rangle &= \exp\left[-\frac{i}{\hbar}H(t_M)\Delta t\right] \\ &\times \exp\left[-\frac{i}{\hbar}H(t_{M-1})\Delta t\right]\dots \\ &\dots\exp\left[-\frac{i}{\hbar}H(t_1)\Delta t\right]|\psi(t_0)\rangle, \quad \Delta t = \frac{t}{M}. \end{aligned} \tag{20}$$

Below, we consider the case where the atom at initial time $t_0 = 0$ was in an excited state and the field was in a vacuum state, i.e., $|\psi(t_0 = 0)\rangle = |0,\uparrow\rangle$.

Using (19), (20), (13) and applying (15), (16), we obtain the following recurrence formula for the system's state vector at a finite time:

$$\begin{aligned} &|\psi(t = t_M)\rangle \\ &= A_{+,M}|+,0\rangle_{t=t_M} + A_{-,M}|-,0\rangle_{t=t_M}, \end{aligned} \tag{21}$$

where $A_{\pm,M}$ are given by the recurrence relations

$$\begin{aligned} A_{+,t_j} &= \exp[-i\Omega_{+,n}(t_j)\Delta t] \\ &\times[\cos(\Delta\theta_{0,j})A_{+,t_{j-1}} + \sin(\Delta\theta_{0,j})A_{-,t_{j-1}}], \\ A_{-,t_j} &= \exp[-i\Omega_{-,n}(t_j)\Delta t] \\ &\times[-\sin(\Delta\theta_{0,j})A_{+,t_{j-1}} + \cos(\Delta\theta_{0,j})A_{-,t_{j-1}}], \\ &\Delta\theta_{0,j} \equiv \theta_0(t_j) - \theta_0(t_{j-1}) \end{aligned} \tag{22}$$

for any $1 < j \leq M$ and

$$\begin{aligned} A_{+,t_1} &= \sin[\theta_0(\Delta t)]\exp[-i\Omega_{+,0}(\Delta t)\Delta t], \\ A_{-,t_1} &= \cos[\theta_0(\Delta t)]\exp[-i\Omega_{-,0}(\Delta t)\Delta t]. \end{aligned} \tag{23}$$

The probability of the atomic transition to a lower state with photon emission into the cavity mode is then

$$\begin{aligned} P_{\downarrow}(t) &= |\langle 1,\downarrow|\psi(t)\rangle|^2 \\ &= |\cos[\theta_0(t)]A_{+,t} - \sin[\theta_0(t)]A_{-,t}|^2. \end{aligned} \tag{24}$$

It is easy to see that, in calculations with (24), $\Omega_{\pm,0}(t)$ may be substituted for $\pm\Omega_0^R(t)$ in (22) and (23) without affecting the result.

Our assumption of classical atomic motion in a cavity constrains the atomic velocity and mass. In our calculations using the above formulas, we assume the de Broglie wavelength ƛ $= h/mv$, where m is the atomic mass, to be much shorter than the cavity mode field wavelength $\lambda = 2\pi/k_c$; i.e., the atomic mass and/or center-of-mass velocity are large. The assumption of uniform atomic motion, i.e., disregarding the recoil momentum when a photon is emitted, also requires that the binding energy $\hbar g_0$ be small compared to the kinetic energy of the atomic center of mass $mv^2/2$. Applicability of the classical approximation to the problem of the interaction of a moving atom in a cavity was analyzed in detail, for example, by Ren *et al.* [3].

We also assume that the cavity Q factor is high enough for the atom–field interaction time to be smaller than γ^{-1}, where γ is the rate of field dissipation through mirror losses. The conditions under consideration, in particular, correspond to the micromaser configuration from [6, 7], where the standard Jaynes–Cummings model was used.

The iteration formula for the evolution operator (17) in the interaction representation is

$$U(t, t_0) = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{t_0}^{t} dt_n \int_{t_0}^{t} dt_{n-1} \dots \int_{t_0}^{t} dt_1 \times T\{C_2(t_n)C_2(t_{n-1})\dots C_2(t_1)\} \qquad (25)$$

$$= T\left\{\exp\left(-i\int_{t_0}^{t} C_2(t')dt'\right)\right\},$$

where T denotes the operation of time ordering of the product of the operators under the integrals. Since the operator $C_2(t)$ commutes with itself for any $t_i \neq t_j$, the time ordering in the right-hand part of Eq. (25) is not necessary in our case.

Let us now assume that the atom–field coupling parameter is given by (3). Using (6) and (25), we then derive for the transition probability at $t_0 = 0$

$$P_{\downarrow}(t) = \left|\beta(t)\sum_{n=0}^{\infty}(-1)^n \frac{\delta^n(t)}{(2n+1)!}\right|^2, \qquad (26)$$

where we designated

$$\delta(t) \equiv \left(\frac{\Delta\omega}{2}t\right)^2 + \beta^2(t), \quad \beta(t) \equiv \frac{g_0}{\Omega_D + \Delta\omega}\sin(\Omega_D t).$$

If the detuning of the atomic transition frequency from the field mode frequency is $\Delta\omega \ll g_0$, then, in our case of a vacuum initial field state in the cavity, the following simple relation can be derived from (26) for the transition probability, which is equal to the average number of photons in the cavity:

$$P_{\downarrow}(t) = \langle a^+(t)a(t)\rangle = \sin^2\left[\frac{g_0}{\Omega}\sin(\Omega_D t)\right]. \qquad (27)$$

This relation is in close agreement with the results from [2, 3]. As follows from (27), the probability of photon emission by an atom at rest in an electromagnetic vacuum oscillates in time with the Rabi vacuum frequency

$$\Omega_R = \sqrt{\left(\frac{\Delta\omega}{2}\right)^2 + g_0^2}$$

as

$$P_{\downarrow}(t) = |\langle 1, \downarrow|\psi(t)\rangle|^2 = \frac{4g_0^2}{\Delta\omega^2 + 4g_0^2}\sin^2(\Omega_R t). \qquad (28)$$

Let us now consider the case where a moving atom crosses a confocal spherical-mirror cavity along its transverse x axis. Assuming the atomic motion along this axis passing through the cavity center to be described by $x(t) = vt$, we place the coordinate origin ($x = 0$) at distance $L/2$ from the cavity center. The time dependence of coupling parameter (4) then takes the form

$$g(x(t)) = g_0 \exp\left[-\frac{1}{w_0^2}\left(vt - \frac{L}{2}\right)^2\right]. \qquad (29)$$

Applying Eq. (26) and using (29) in (6), we find that the transition probability can be written as a series

$$P_{\downarrow}(t) = \left|I(t)\sum_{n=0}^{\infty}(-1)^n \frac{[I^2(t) + \delta^2(t)]^n}{(2n+1)!}\right|^2, \qquad (30)$$

where

$$I(t) \equiv \frac{g_0 w_0}{v}\int_{-L/2w_0}^{(vt - L/2)/w_0} e^{-\tau^2} d\tau, \qquad (31)$$

$$\delta(t) \equiv -\frac{\Delta\omega}{2}t.$$

For an exact resonance, $\Delta\omega = 0$, we derive from (30) and (31)

$$P_{\downarrow}(t) = \sin^2(I(t)). \qquad (32)$$

Since the cavity field decreases exponentially with increasing distance from the cavity center, we may take $L \longrightarrow \infty$ ($L \gg 2w_0$), $0 \le vt \le L$ in the expression for $I(t)$. This expression can then be written as

$$I(t) = \frac{g_0 w_0}{v}\frac{\sqrt{\pi}}{2}\left[1 + \operatorname{erf}\left(\frac{vt - L/2}{w_0}\right)\right], \qquad (33)$$

where $\operatorname{erf} z$ is the error integral; $\operatorname{erf} 0 = 0$ and $\operatorname{erf}\infty = 1$.

Let us now calculate the probability of photon emission by an atom as it passes through the cavity, i.e., $P_{\downarrow}(t)$ for $t = L/v$ on condition that $L \gg 2w_0$, by using (32). Clearly, this probability is

$$P_{\downarrow}(t = \infty) = \sin^2\left(\sqrt{\pi}\frac{g_0 w_0}{v}\right); \qquad (34)$$

it thus follows that the atom leaves the cavity in a lower state if $g_0 w_0 / v\sqrt{\pi} = n + 1/2$, $n = 0, 1, 2, \dots$, while, for $g_0 w_0 / v\sqrt{\pi} = n$, the state of the atom after its passage through the cavity does not change.

When $g_0 w_0 \sqrt{\pi}/v \ll 1$ and $v \gg g_0 w_0 \sqrt{\pi}$, the energy exchange between the excited atom and the cavity cannot take place, and the atomic and field states remain unchanged during the entire passage of the atom through the cavity.

3. DOPPLER–RABI OSCILLATIONS

As our calculations with formula (24) show, the time dependence of the transition probability for a moving atom basically depends on the center-of-mass velocity, coupling constant, and the detuning of the transition

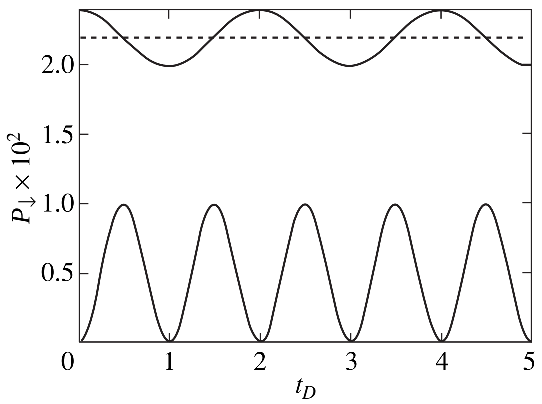


Fig. 1. Probability $P_\downarrow$ for the transition of an excited atom to a lower state with photon emission as it moves in a plane-parallel-mirror cavity versus dimensionless time $t_D = \Omega_D t/\pi$, $\Omega_D = v\omega_a/c$ at $\Delta\omega \ll g_0$, $\xi = \Omega_R/\Omega_D = 0.1$. The upper curve represents the dependence $\cos(k_c x(t_R))$.

frequency from the cavity mode frequency. Depending on the ratio $\xi = \Omega_R/\Omega_D$ of the Doppler and Rabi frequencies, $\Omega_R = \sqrt{g_0^2 + (\Delta\omega/2)^2}$, there can be no Rabi oscillations in the time dependence of the emission probability. For $\xi \ll 1$ and $\Delta\omega \ll g_0$, the probability of photon emission by the atom and photon escape from the cavity in the lowest state can be low even at large coupling constants (Fig. 1). In this case, the frequency of oscillations in the transition probability corresponds to the 1/ξth Rabi harmonic, while the probability decreases as ξ^2. As follows from Eq. (27), for $\xi \ll 1$ and $\Delta\omega \ll g_0$, the transition probability

$$P_\downarrow(t) = \langle a^+(t)a(t)\rangle = \xi^2 \sin^2(\Omega_D t)$$

oscillates with the frequency equal to the Doppler shift rather than with the Rabi frequency.

For a small detuning, the condition for the presence of Rabi-type probability oscillations is a Doppler resonance, when $\xi = 1$ (Fig. 2); in this case, $P_\downarrow \le \sin^2(1\ \text{rad})$.

If $\xi > 1$ and $\Delta\omega \ll g_0$, i.e., if the Doppler shift is smaller than the Rabi frequency, the transition probability at maxima is unity, as with an atom at rest. As we see from Fig. 3a, there are no Rabi oscillations for $\xi = 2$. However, Rabi oscillations are restored near the field antinodes (Figs. 3b and 3c). As one might expect, ordinary Rabi oscillations take place for $\xi \gg 1$, but only near the field antinodes in the cavity at a finite atomic velocity.

When the detuning of the atomic transition frequency from the field mode eigenfrequency is comparable in magnitude to the coupling constant, $\Delta\omega \ge g_0$, the average number of photons in the cavity generally decreases with increasing detuning and exhibits a complex time dependence on detuning. In Fig. 4, $P_\downarrow(t)$ is plotted for a small detuning, $\Delta\omega = g_0$, at $\xi = 2.2$. As we see from the plot, the maximum number of photons differs little from the corresponding value for an atom at rest. The shape of the oscillations in photon number proves to be very sensitive to the detuning even if the latter is small but comparable to the coupling constant g_0.

If $\xi \ll 1$ and $\Delta\omega > g_0$, as with a small detuning, the transition probability oscillates in time with Ω_D. In this case, the transition probability exhibits a dependence on the Doppler frequency shift, coupling constant, and time similar to (28) for an atom at rest with the change $\Omega_R \longrightarrow \Omega_D$, i.e.,

$$P_\downarrow(t) = \frac{g_0^2}{\Omega_D^2 + g_0^2}\sin^2(\Omega_D t). \tag{35}$$

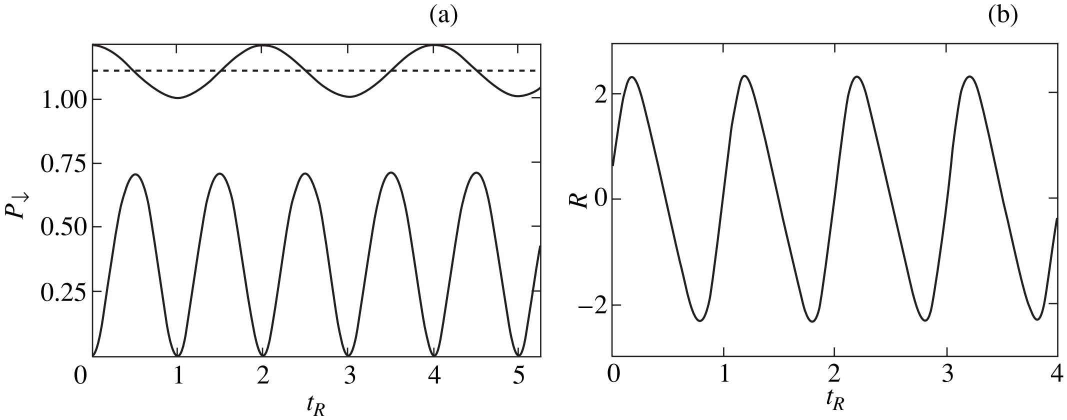


Fig. 2. (a) Probability $P_\downarrow$ for the transition of an excited atom to a lower state with photon emission as it moves in a plane-parallel-mirror cavity versus dimensionless time $t_R = \Omega_R t/\pi$, $\Omega_R = \sqrt{g_0^2 + (\Delta\omega/2)^2}$ at Doppler–Rabi resonance, $\xi = \Omega_R/\Omega_D = 1$. The upper curve represents the dependence $\cos(k_c x(t_R))$. (b) Transition rate $R(t_R) = dP_\downarrow(t_R)/dt_R$.

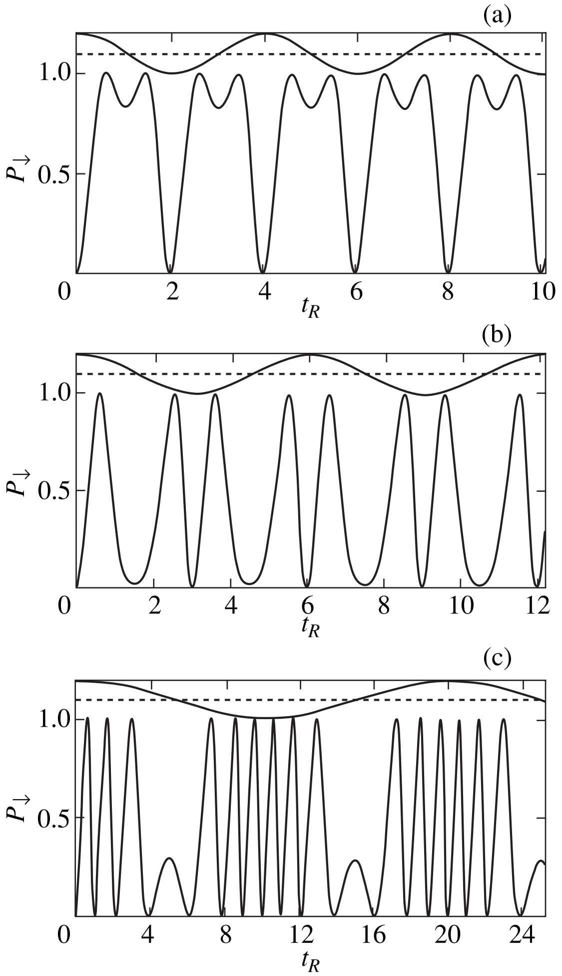


Fig. 3. Same as in Fig. 2a, but for $\xi = 2$ (a), 3 (b), and 10 (c).

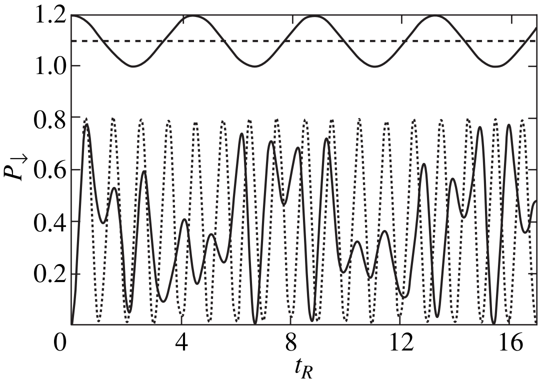


Fig. 4. Same as Fig. 2a for the detuning $\Delta\omega = g_0$ at $\xi = 2.2$. Dashed curve shows the Rabi oscillations of a resting atom for the same g_0 and $\Delta\omega$ values.

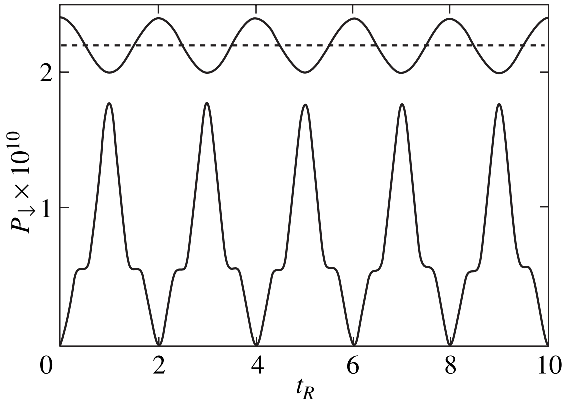


Fig. 5. Same as in Fig. 2a, but for the detuning $\Delta\omega = 2 \times 10^5 g_0$ at $\xi = 1$.

As we see from (35), the frequency of energy exchange between atom and field is determined by the Doppler shift of the atomic transition frequency if $\xi \ll 1$.

At resonance, $\xi = 1$, the transition probability decreases as the detuning increases and the oscillation frequency for $\Delta\omega \gg g_0$ is $\Omega_R/2$ (Fig. 5). At the same time, when there is a detuning and $\xi < 1$, oscillations similar in shape to Rabi oscillations can be restored on reaching a resonance $\Delta\omega \approx \Omega_D$(or $\Delta\omega \approx \omega_c v/c$ [2]). Figure 6 shows the photon emission probability for an exact resonance, $\Delta\omega = \Omega_D = 10^2 g_0$, at $\xi = 0.5$. As we see from Fig. 6a, there are Rabi-type oscillations with a unit amplitude but with a frequency that is a factor of $\Delta\omega/g_0$ lower than the frequency of ordinary Rabi oscillations at a given detuning. Consequently, at Doppler resonance $\Delta\omega \approx \Omega_D$, the Doppler frequency shift effectively offsets the effect of detuning on the transition probability.

As we see from Fig. 6b, the dependence $P_\downarrow(t)$ is modulated by shallow oscillations with twice the Rabi frequency, which results in a deep modulation of the transition rate $dP_\downarrow/dt$ (Fig. 6c).

Figure 7 shows a plot of $P_\downarrow(t)$ for the detuning $\Delta\omega = 20 g_0$ at $\xi = 10$. In this typical case for $\xi \gg 1$, ordinary Rabi oscillations turn out to be modulated in amplitude with the period $2/\Delta\omega$ much as the frequency modulation takes place in the absence of a detuning (see Fig. 3).

Our calculations of the transition probability by the summation of series (26) show that the alternating series converges very slowly for large $\delta(t)$ and at a large detuning $\Delta\omega$ or coupling constant g_0. In these cases, calculations can easily be performed with the recurrence formula (24).

Currently, the Jaynes–Cummings model for an atom at rest is used in micromaser theory [6–9]. A comparison of theory with experiment [10–12] shows that predictions of this theory are not always consistent with experiment. For example, by experimentally analyzing the photon statistics for a micromaser, Rempe *et al.* [12] pointed out qualitative disagreement between theory and experiment for the atomic transition frequency ω_a = 21.5 GHz, cavity length L = 2.4 cm, coupling parameter g_0= 10 kHz, and the time it takes for the atom to pass through the cavity L/v = 50 μs. At the same time, agreement with theory is achieved by changing

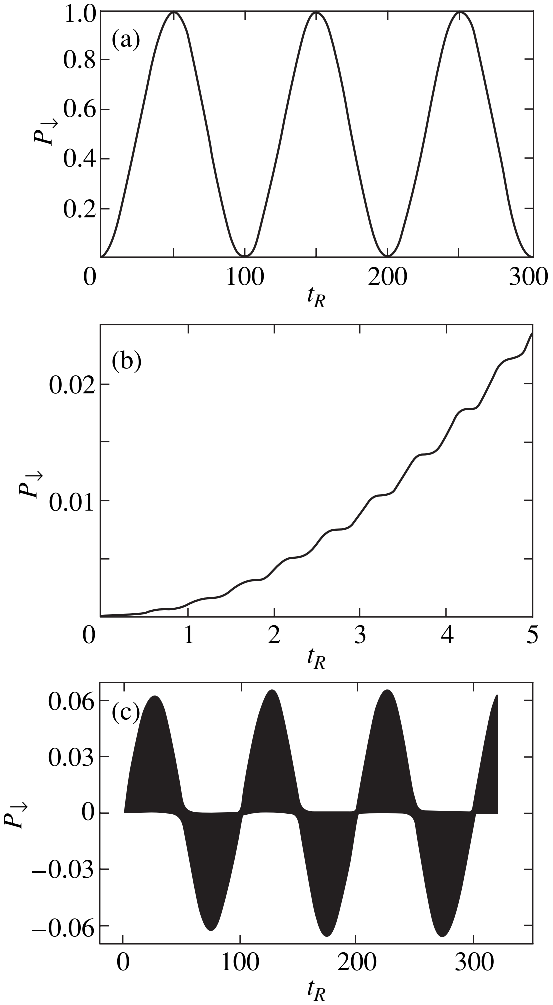


Fig. 6. (a, b) Same as in Fig. 2a, but for the detuning $\Delta\omega = 100g_0$ at $\xi = 0.5$; (c) same as Fig. 2b.

micromaser parameters, decreasing the atom–field interaction time to $L/v = 35$ μs and increasing the coupling parameter to $g_0 \approx 44$ kHz. Calculations with formula (24), which includes the effect of atomic center-of-mass motion on the atom–field interaction, indicate that, in the latter case, the system's state is close to Doppler–Rabi resonance ($\xi \approx 1$) and the variations in transition probability follow a law similar to ordinary Rabi oscillations (see Fig. 2) for an atom at rest. At the same time, for the other experimental conditions given above, the ratio of the Rabi frequency to the Doppler shift is much smaller ($\xi \approx 0.3$), and a pattern of energy exchange between atom and field (similar to that in Fig. 1 at $\xi \ll 1$) that differs sharply from the pattern at resonance ($\xi = 1$) takes place. The detuning of the cavity frequency from the atomic transition frequency significantly affects the pattern of oscillations in transition probability. The time dependence of probability in Fig. 4, which corresponds to micromaser conditions [12], shows that the transition probability is very sensitive to the detuning for a moving atom.

The above analysis leads us to conclude that the effects attributable to atomic motion in the cavity, in our view, must be taken into account in micromaser theory and for some microlaser schemes.

Next, consider the situation when a moving atom crosses a confocal spherical-mirror cavity through its center transversally. In our calculations, we use the cavity parameters, atomic transition frequency, coupling constant, and atomic velocities from [13], where the quantum Rabi oscillations of Rydberg atoms were studied experimentally.

Let the detuning be negligible, $\Delta\omega \ll g_0$ [formulas (32) and (33)]. The time dependence of the transition probability (average photon number) for an initially excited atom passing through the cavity (see Fig. 8) illustrates the resonance conditions found in the preceding section. Figure 8a shows the transition probability when

$$\frac{g_0 w_0}{v\sqrt{\pi}} = n + \frac{1}{2}, \quad n = 6.5;$$

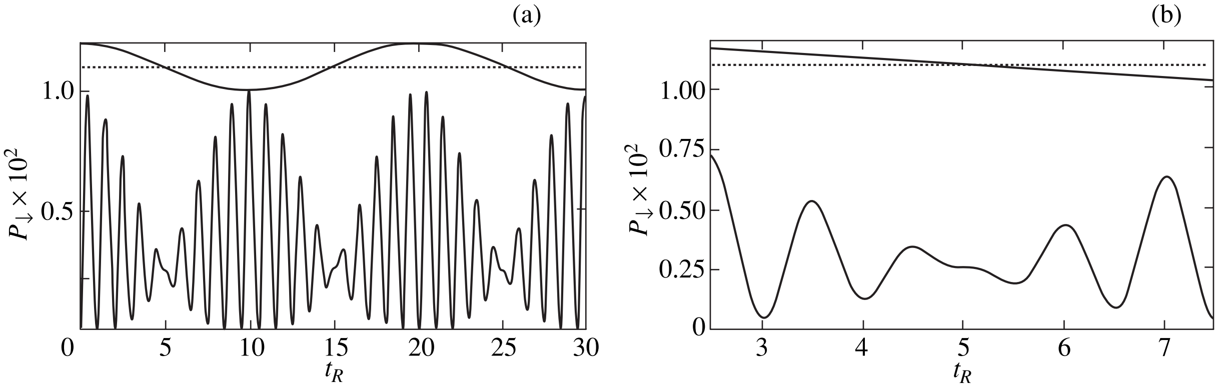


Fig. 7. (a) Same as in Fig. 2a, but for the detuning $\Delta\omega = 20g_0$ at $\xi = 10$; (b) the same dependence near $t_R = 5$.

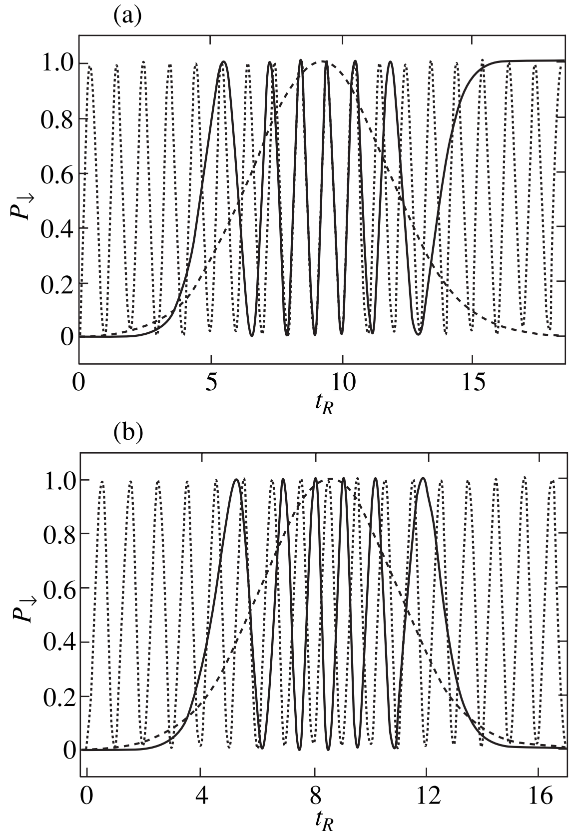


Fig. 8. Probability $P_\downarrow$ for the transition of an excited atom to a lower state with photon emission versus dimensionless time $t_R = \Omega_R t/\pi$, $\Omega_R = \sqrt{g_0^2 + (\Delta\omega/2)^2}$, as the atom passes through a confocal spherical-mirror cavity with a Gaussian transverse field distribution for the atomic center-of-mass velocities (a) $v = 8.122 \times 10^3$ cm s^{-1}, $\Omega_D = 8.67 \times 10^4$ s^{-1} and (b) $v = 8.799 \times 10^3$ cm s^{-1}, $\Omega_D = 9.39 \times 10^4$ s^{-1}. The dashed curves represent the time dependence of the cavity transverse field distribution, and the dotted curves represent the Rabi oscillations of an atom at rest for the same coupling constant g_0 and detuning $\Delta\omega \ll g_0$.

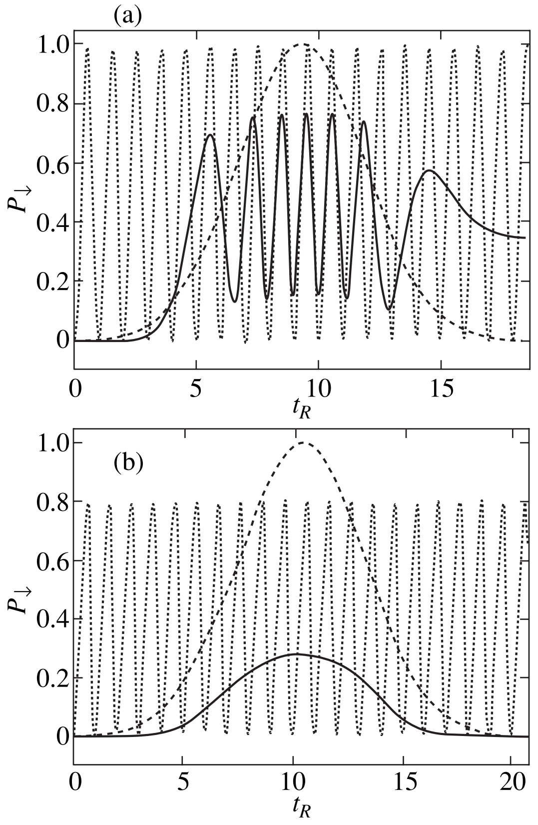


Fig. 9. Same as in Fig. 8, but at $v = 8.122 \times 10^3$ cm s^{-1} for various detunings: (a) $\Delta\omega = 0.2g_0$ ($\Omega_R = 1.58 \times 10^5$ s^{-1}) and (b) $\Delta\omega = g_0$ ($\Omega_R = 1.76 \times 10^5$ s^{-1}).

in this case, the atom after its passage through the cavity proves to be in a lower state and the cavity field contains a single photon with a unit probability. Figure 8b shows the dynamics of the transition probability for a different resonance condition,

$$\frac{g_0 w_0}{v\sqrt{\pi}} = n, \quad n = 6;$$

in this case, the cavity is transparent for the excited atom; i.e., the state of the atom does not change as it leaves the cavity.

The dynamics of the transition probability when there is a detuning of the cavity mode frequency from the atomic transition frequency is shown in Fig. 9 for various detunings. As we see from the plots, the transition probability is very sensitive to the detuning: even a small detuning compared to the coupling constant, $\Delta\omega = 0.2g_0$, causes the transition probability to decrease sharply both inside the cavity and at its exit (Fig. 9a). A comparison indicates that the maximum transition probability for an atom at rest decreases with increasing detuning much more slowly than for a moving atom. For the detuning $\Delta\omega = g_0$, the deexcitation probability of the atom as it leaves the cavity is nearly zero (Fig. 9b) for any values of the coupling constant g_0 and atomic velocity.

4. CONCLUSION

We have analyzed the interaction of a uniformly moving two-level atom with a high-Q cavity field. Our calculations for a plane-parallel-mirror cavity lead us to the following conclusions.

We considered the conditions for spontaneous emission when the atom was initially in an excited state and the cavity field was in a vacuum state. We established that, if the ratio of the Rabi frequency to the Doppler shift of the atomic transition frequency, $\xi = \Omega_R/\Omega_D$, is much less than unity, then the average number of photons in the cavity is small when the cavity mode frequency is close to the atomic frequency, $\Delta\omega \ll g_0$. In this case, the average number of photons (transition probability) oscillates in time with Doppler frequency rather than with Rabi frequency; the oscillation amplitude is proportional to $\xi^2 \ll 1$.

For a Doppler resonance ($\xi = 1$), the photon number oscillates with Rabi frequency, with the average number of photons being no larger than $\sin^2(1 \text{ rad}) \approx 0.708$.

If the Rabi frequency is larger than the Doppler frequency shift ($\xi > 1$), then the average number of photons in the cavity can reach unity when the atom is near the antinodes of the spatial field distribution in the cavity and is much smaller than unity near the field nodes. The ordinary Rabi oscillations characteristic of an atom at rest take place for finite atomic center-of-mass velocities if $\xi \gg 1$ only near the field antinodes.

When there is a detuning of the atomic transition frequency from the field mode eigenfrequency, $\Delta\omega \geq g_0$, the average number of photons in the cavity decreases with increasing detuning and exhibits a complex time dependence for each specific detuning. The shape of the oscillations in photon number proves to be very sensitive to the detuning even if the latter is small but comparable to the coupling constant g_0.

The photon emission probability for an exact resonance ($\Delta\omega = \Omega_D$, $\xi < 1$) increases sharply. Under these conditions, there are Rabi-type oscillations with a unit amplitude but with a frequency that is a factor of $\Delta\omega/g_0$ lower than the frequency of ordinary Rabi oscillations for a given detuning. Thus, at Doppler resonance $\Delta\omega \approx \Omega_D$, the Doppler frequency shift effectively offsets the effect of detuning on the transition probability.

Rabi-type oscillations with twice the period and with an amplitude much smaller than unity occur at $\xi = 1$ and $\Delta\omega \gg g_0$.

Our calculations for a Gaussian confocal-cavity mode showed that the probability of photon emission by the atom and its deexcitation after it passes through the cavity decreases sharply as the detuning increases and is nearly zero at $\Delta\omega = g_0$, irrespective of the atomic velocity.

Our analysis for the two types of cavities leads us to conclude that the effects attributable to atomic center-of-mass motion must be taken into account to theoretically interpret experiments with micromasers [13, 14] and microlasers.

REFERENCES

1. E. T. Jaynes and F. W. Cummings, Proc. IEEE **51**, 89 (1963).
2. P. Meystre, Opt. Commun. **90**, 41 (1992).
3. W. Ren, J. D. Cresser, and H. J. Carmichael, Phys. Rev. A **46**, 7162 (1992).
4. M. Wilkens and P. Meystre, Opt. Commun. **94**, 66 (1992).
5. W. Ren and H. J. Carmichael, Phys. Rev. A **51**, 752 (1995).
6. P. Filipowicz, J. Javanainen, and P. Meystre, Phys. Rev. A **34**, 3077 (1986).
7. L. A. Lugiato, M. O. Scully, and H. Walther, Phys. Rev. A **36**, 740 (1987).
8. H.-J. Briegel, B.-G. Englert, and M. O. Scully, Phys. Rev. A **54**, 3603 (1996).
9. M. O. Scully, G. M. Meyer, and H. Walther, Phys. Rev. Lett. **76**, 4144 (1996).
10. D. Meschede, W. Walther, and G. Müller, Phys. Rev. Lett. **54**, 551 (1985).
11. G. Rempe, H. Walther, and N. Klein, Phys. Rev. Lett. **58**, 353 (1987).
12. G. Rempe, F. Schmidt-Kaler, and H. Walther, Phys. Rev. Lett. **64**, 2783 (1990).
13. M. Brune, F. Schmidt-Kaler, A. Maali, *et al.*, Phys. Rev. Lett. **76**, 1800 (1996).
14. M. Brune, P. Nussenzweig, F. Schmidt-Kaler, *et al.*, Phys. Rev. Lett. **72**, 3339 (1994).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 471–476.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 539–545.
Original Russian Text Copyright © 2001 by Martynenko, Faustov.

ATOMS, SPECTRA, RADIATION

The *g* Factors of Bound Particles in Quantum Electrodynamics

A. P. Martynenko* and R. N. Faustov**

**Samara State University, Samara, 443011 Russia*
e-mail: mart@info.ssu.samara.ru

*** Research Council Kibernetika, Russian Academy of Sciences, Moscow, 117333 Russia*
e-mail: faustov@theory.npi.msu.su

Received February 9, 2001

Abstract—A quasipotential method is formulated for calculating relativistic and radiative corrections to the magnetic moment of a two-particle bound state in the case of particles of arbitrary spin. It is shown that expressions for the *g* factors of bound particles contain terms of order $O(\alpha^2)$ that depend on the spin of particles. Numerical values of the *g* factor of an electron in the hydrogen and deuterium atoms are obtained. *© 2001 MAIK "Nauka/Interperiodica".*

The study of electromagnetic properties of hydrogen-like atoms and ions in quantum electrodynamics is one of the basic problems in the theory of two-particle bound states. The experimental verification of the calculation of the *g* factors of particles in a bound state has been carried out over many years [1, 2]. The measured values of the *g* factors of an electron in hydrogen, deuterium, and helium ($^4He^+$) are in good agreement with the relevant theoretical results. Recently, the field of experimental investigations of hydrogen-like ions has been substantially expanded [3, 4]. These experiments call for new theoretical calculations of different contributions to the *g* factors of bound particles [5–7]. The most precise measurements of the electron *g* factor have been performed for the hydrogen-like ion of carbon, $^{12}C^{5+}$ ($Z = 6$) [3, 4, 7]:

$$g_e^{\exp}(^{12}\mathrm{C}^{5+}) = 2.001\ 041\ 596\ 4(8)(6)(40), \qquad (1)$$

where the numbers in parentheses stand for the statistical error (8), systematic error (6), and the error associated with the mass of electron (40). The theoretical investigations of the electromagnetic properties of hydrogen-like atoms carried out in [8–14] have shown that the gyromagnetic factors of bound particles can be represented by the series

$$g(\mathrm{H\ atom}) = 2 + \Delta g_{\mathrm{rel}} + \Delta g_{\mathrm{rad}} + \Delta g_{\mathrm{rec}} + \ldots. \qquad (2)$$

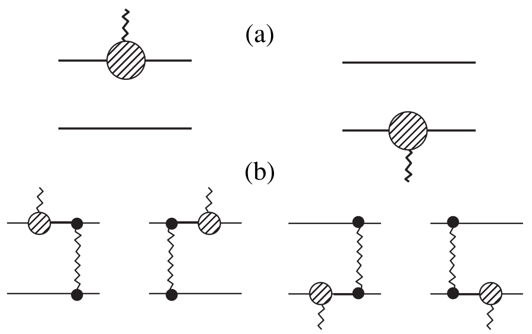
In [8, 9], the relativistic (Δg_{rel}), radiative (Δg_{rad}), and recoil (Δg_{rec}) corrections were calculated up to the terms of order $\alpha^3(m/M)$ and $\alpha^2(m/M)^2$ within the quasipotential method for spin-1/2 particles that form a bound system. The dots in (2) stand for other possible terms in the *g* factor. At the same time, the experiments with deuterium and with hydrogen-like ions whose nuclei have different spins call for the methods for calculating the *g* factors in this case. Eides and Grotch [15] proposed a method, based on the Bargmann–Michel–Telegdi (BMT) equation [16], for calculating the corrections to the gyromagnetic factors and suggested that these factors are independent of the spins of the constituents. In this paper, we formulate a quasipotential method for calculating the magnetic moment of the bound state of two particles of arbitrary spin and calculate the basic contributions to the corrections in (2) of order $O(\alpha^2)$ and $O(\alpha^3)$ that arise in the approximation of one-photon interaction of particles in the bound state (see figure).

The interaction of massive particles of arbitrary spin with an electromagnetic field have been investigated by various methods over a long period of time [17–27]; however, this problem is still far from its final solution. In [18–20], it was shown that, within a tree approximation, a particle of arbitrary spin must have a gyromagnetic factor of $g = 2$. In the general case, the matrix element of electromagnetic current for a particle of spin S is determined by $(2S + 1)$ form factors (the charge, magnetic, quadrupole, etc., form factors). When studying the magnetic moments of simple atomic systems, one can restrict the analysis to the form factors with the lowest multipolarity that determine the distributions of charge and magnetic moment by representing the matrix element J_μ of the operator of electromagnetic current between states with momenta p and q as

$$J_\mu = \bar{U}(p)\left\{\Gamma_\mu F_1^D + \frac{1}{2m}\Sigma_{\mu\nu}k^\nu F_2^P\right\}U(q), \qquad (3)$$

where $k^\nu = (p - q)^\nu$. The wave function $U(p)$ of a particle of arbitrary spin involved in (3) can be represented as follows [26, 27]:

$$U = \begin{pmatrix}\xi\\ \eta\end{pmatrix} = \begin{pmatrix}\xi^{\alpha_1\alpha_2\ldots\alpha_p}_{\dot\beta_1\dot\beta_2\ldots\dot\beta_q}\\ \eta^{\beta_1\beta_2\ldots\beta_q}_{\dot\alpha_1\dot\alpha_2\ldots\dot\alpha_p}\end{pmatrix}, \quad p + q = 2S, \qquad (4)$$



Generalized vertex functions Γ_μ of two particles: (graph *a*) $\Gamma_\mu^{(0)}$ and (graph *b*) $\Gamma_\mu^{(1)}$. The heavy line represents the negative-frequency part of the propagator of a particle.

where the spin tensors ξ and η are symmetric with respect to the upper and lower indices. For a particle of half-integer spin, we have

$$p = S + 1/2, \quad q = S - 1/2,$$

whereas, in the case of an integer spin, we have

$$p = q = S.$$

The Lorentz transformation for the spinors ξ and η is given by [27, 28]

$$\xi = \exp\left(\frac{\mathbf{\Sigma}\cdot\boldsymbol{\phi}}{2}\right)\xi_0, \quad \eta = \exp\left(-\frac{\mathbf{\Sigma}\cdot\boldsymbol{\phi}}{2}\right)\xi_0; \tag{5}$$

here, ξ_0 is the spinor in the rest frame of reference, the direction of the vector $\boldsymbol{\phi}$ coincides with the velocity direction of the particle, $\tanh\phi = v$, and the generator $\mathbf{\Sigma}$ of the Lorentz transformation is given by

$$\mathbf{\Sigma} = \sum_{i=1}^{p}\boldsymbol{\sigma}_i - \sum_{i=p+1}^{p+q}\boldsymbol{\sigma}_i. \tag{6}$$

The matrix $\boldsymbol{\sigma}_i$ acts on the ith index of the spinor ξ_0 as follows:

$$\boldsymbol{\sigma}_i\xi_0 = (\boldsymbol{\sigma}_i)_{\alpha_i\beta_i}(\xi_0)_{\ldots\beta_i\ldots}. \tag{7}$$

The components of the antisymmetric tensor $\Sigma_{\mu\nu}$ in (3) represent the generators of boosts and rotations [27, 28]:

$$\Sigma_{n0} = \begin{pmatrix} \Sigma_n & 0 \\ 0 & -\Sigma_n \end{pmatrix},$$
$$\Sigma_{mn} = -2i\epsilon_{mnk}\begin{pmatrix} s_k & 0 \\ 0 & s_k \end{pmatrix}, \quad s = \frac{1}{2}\sum_{i=1}^{2S}\boldsymbol{\sigma}_i. \tag{8}$$

In the standard representation, which is introduced by analogy with the case of spin 1/2, the wave function (4) of a free particle is expressed up to terms of order $(v/c)^2$ as

$$U(p) = \begin{pmatrix} \left[1 + \dfrac{(\mathbf{\Sigma}\cdot\mathbf{p})^2}{8m^2}\right]\xi_0 \\ \dfrac{\mathbf{\Sigma}\cdot\mathbf{p}}{2m}\xi_0 \end{pmatrix}. \tag{9}$$

The magnetic moment of a bound state of two particles of arbitrary spin is determined as [8, 9]

$$\mathbf{M} = -\frac{i}{2}\left[\frac{\partial}{\partial\mathbf{\Delta}} \times \langle\mathbf{K}_A|\mathbf{J}(0)|\mathbf{K}_B\rangle\right], \quad \mathbf{\Delta} = \mathbf{K}_A - \mathbf{K}_B, \tag{10}$$

where the matrix element of the operator of electromagnetic current between two bound states is expressed in terms of the wave functions $\Psi_{\mathbf{K}_B}(\mathbf{p})$ of the bound system and the generalized vertex function Γ_μ shown in the figure:

$$\langle\mathbf{K}_A|J_\mu(0)|\mathbf{K}_B\rangle = \int\frac{d\mathbf{p}_1 d\mathbf{p}_2}{(2\pi)^3}\delta(\mathbf{p}_1 + \mathbf{p}_2 - \mathbf{K}_A)$$
$$\times\Psi^*_{\mathbf{K}_A}(\mathbf{p})\Gamma_\mu(\mathbf{p}, \mathbf{q}, E_A, E_B) \tag{11}$$
$$\times\Psi_{\mathbf{K}_B}(\mathbf{q})\delta(\mathbf{q}_1 + \mathbf{q}_2 - \mathbf{K}_B)\frac{d\mathbf{q}_1 d\mathbf{q}_2}{(2\pi)^3}.$$

The vertex function Γ_μ is determined by the five-point function

$$R_\mu = \langle 0|\psi_1(t, \mathbf{x}_1)\psi_2(t, \mathbf{x}_2)J_\mu(0)\bar{\psi}_1(\tau, \mathbf{y}_1)\bar{\psi}_2(\tau, \mathbf{y}_2)|0\rangle, \tag{12}$$

projected onto positive-frequency states:

$$\Gamma_\mu = G^{-1}R_\mu^{(+)}G^{-1}, \quad R_\mu^{(+)} = U_1^*U_2^*R_\mu U_1 U_2, \tag{13}$$

where G is the two-particle Green's function. Since we deal with a weakly bound two-particle system, the quantities Γ, R, and G^{-1} can be expanded in series according to perturbation theory:

$$\Gamma = \Gamma^{(0)} + \Gamma^{(1)} + \ldots, \quad R = R_0 + R_1 + \ldots,$$
$$G^{-1} = G_0^{-1} - V_1 - \ldots, \tag{14}$$

$$\Gamma^{(0)} = G_0^{-1}R_0G_0^{-1}, \tag{15}$$

$$\Gamma^{(1)} = G_0^{-1}R_1G_0^{-1} - V_1G_0\Gamma^{(0)} - \Gamma^{(0)}G_0V_1, \ldots, \tag{16}$$

where G_0 is the Green's function of two noninteracting particles and V_1 is a quasipotential of one-photon interaction (see Eq. (19)).

When passing from the rest frame to a moving frame of reference with the momentum $\mathbf{K}_B$, the wave function

$\Psi_{\mathbf{K}_B}(\mathbf{p})$ of a system of particles with spins s_1 and s_2 is transformed as [29]

$$\delta(\mathbf{p}_1+\mathbf{p}_2-\mathbf{K}_B)\Psi_{\mathbf{K}_B}(\mathbf{p}) = D_1^{S_1}(R_W)D_2^{S_2}(R_W) \times\sqrt{\frac{\epsilon_1^0\epsilon_2^0 M}{\epsilon_1\epsilon_2 E}}\Psi_0(\mathbf{p}^0)\delta(\mathbf{p}_1^0+\mathbf{p}_2^0), \tag{17}$$

where $D^S(R)$ are rotation matrices and R_W is the Wigner rotation associated with the Lorentz transformation $\Lambda_{\mathbf{K}_B}$:

$$(E, \mathbf{K}_B) = \Lambda_{\mathbf{K}_B}(M, 0); \quad (\epsilon, \mathbf{p}) = \Lambda(\epsilon^0, \mathbf{p}^0),$$

$$E = \sqrt{M^2+\mathbf{K}_B^2}, \quad \epsilon(\mathbf{p}) = \sqrt{\mathbf{p}^2+m^2}.$$

The exact expression for the rotation matrix is [9]

$$D^S(R_W) = S^{-1}(\mathbf{p})S(\mathbf{K}_B)S(\mathbf{p}^0), \tag{18}$$

where $S(\mathbf{p})$ is the matrix of the Lorentz transformation of the spinor wave function (4). The quasipotential wave function $\Psi_0(\mathbf{p}^0)$ in the rest frame of the bound state satisfies the following quasipotential equation [30]:

$$\begin{aligned} G_0^{-1}\Psi &\equiv \left(\frac{b^2}{2\mu_R}-\frac{\mathbf{p}^{02}}{2\mu_R}\right)\Psi_0(\mathbf{p}^0) \\ &= \int V(\mathbf{p}^0, \mathbf{q}^0, M)\Psi_0(\mathbf{q}^0)\frac{d\mathbf{q}^0}{(2\pi)^3}, \end{aligned} \tag{19}$$

where μ_R is the relativistic reduced mass:

$$\mu_R = \frac{E_1E_2}{M} = \frac{M^4-(m_1^2-m_2^2)^2}{4M^3},$$

$$E_{1,2} = \frac{M^2-m_{2,1}^2+m_{1,2}^2}{2M},$$

$M = E_1 + E_2$ is the mass of the bound state, and

$$b^2(M) = \frac{[M^2-(m_1+m_2)^2][M^2-(m_1-m_2)^2]}{4M^2}.$$

In the nonrelativistic limit, Eq. (19) reduces to the Schrödinger equation with the Coulomb potential. Using (5), we can obtain the following approximate expressions for the functions D^S in (18):

$$\begin{aligned} D^S(R_W) \approx 1 &+ \frac{\mathbf{p}^{02}-(\mathbf{\Sigma}\cdot\mathbf{p}^0)(\mathbf{\Sigma}\cdot\mathbf{p}^0)}{4m^2} \\ &+ \frac{\mathbf{K}_B^2-(\mathbf{\Sigma}\cdot\mathbf{K}_B)(\mathbf{\Sigma}\cdot\mathbf{K}_B)}{4M^2} \\ &+ \frac{\mathbf{p}^0\cdot\mathbf{K}_B-(\mathbf{\Sigma}\cdot\mathbf{p}^0)(\mathbf{\Sigma}\cdot\mathbf{K}_B)}{4mM}. \end{aligned} \tag{20}$$

The main contribution to the vertex function Γ_μ, which is determined by graph *a* in the figure, can be represented (in the Breit frame) as follows:

$$\begin{aligned} \mathbf{\Gamma}^{(0)}(\mathbf{p}, \mathbf{q}) &= \bar{U}_1(p_1)e_1\left\{\mathbf{\Gamma}_1+\frac{i\kappa_1}{m_1}[\mathbf{S}_1\times\mathbf{\Delta}]\right\} \\ &\times U_1(q_1)\delta(\mathbf{p}_2-\mathbf{q}_2)+(1\longleftrightarrow 2), \end{aligned} \tag{21}$$

$$\mathbf{S}_1 = \begin{pmatrix} \mathbf{s}_1 & 0 \\ 0 & \mathbf{s}_1 \end{pmatrix}, \quad \mathbf{\Delta} = \mathbf{p}_1-\mathbf{q}_1,$$

where $F^D_{1,2}(0) = e_{1,2}$, $F^P_{1,2}(0) = e_{1,2}\kappa_{1,2}$, and the matrix

$$\mathbf{\Gamma} = \begin{pmatrix} 0 & \mathbf{\Sigma} \\ -\mathbf{\Sigma} & 0 \end{pmatrix} \tag{22}$$

is a natural generalization of the Dirac γ matrix for spin 1/2 in standard representation. To simplify individual terms in (20), it is convenient to apply the following commutation relations [27]:

$$[\Sigma_i, \Sigma_j] = 4i\epsilon_{ijk}s_k, \quad [\Sigma_i, s_j] = i\epsilon_{ijk}\Sigma_k. \tag{23}$$

When constructing the vertex function $\mathbf{\Gamma}^{(0)}(\mathbf{p}, \mathbf{q})$ with regard to the terms of order $(v/c)^2$, we avail ourselves of the explicit expression of the wave function (9) and transform the individual terms of the matrix element (21) using the equations of motion for the spinors $U(p)$. Taking into account $\delta(\mathbf{p}_2 - \mathbf{q}_2)$, we obtain

$$\begin{aligned} &\bar{U}_1(\mathbf{p}_1)\frac{\mathbf{p}_1+\mathbf{q}_1}{2m_1}U(\mathbf{q}_1) \\ &= 2\mathbf{p}^0-\frac{\epsilon_2}{M}\mathbf{\Delta}+\frac{i\mathbf{p}^0(\mathbf{S}_1\cdot[\mathbf{p}^0\times\mathbf{\Delta}])}{m_1^2}, \end{aligned} \tag{24}$$

$$\begin{aligned} &\bar{U}_1(\mathbf{p}_1)\frac{\epsilon_1(\mathbf{p}_1)-\epsilon_1(\mathbf{q}_1)}{2m_1}\mathbf{A}_1U_1(\mathbf{q}_1) \\ &= \frac{2\mathbf{p}^0\cdot\mathbf{\Delta}}{m_1^2}i[\mathbf{S}_1\times\mathbf{p}^0], \end{aligned} \tag{25}$$

$$\begin{aligned} &\bar{U}_1(\mathbf{p}_1)[\mathbf{S}_1\times\mathbf{\Delta}]U_1(\mathbf{q}_1) = [\mathbf{S}_1\times\mathbf{\Delta}] \\ &-\frac{1}{2m_1^2}\{\mathbf{p}^0(\mathbf{S}_1\cdot[\mathbf{p}^0\times\mathbf{\Delta}])+[\mathbf{p}^0\times\mathbf{S}_1](\mathbf{p}^0\cdot\mathbf{\Delta})\}. \end{aligned} \tag{26}$$

The bound-state effects in the vertex function Γ_μ are determined by graph *b* in the figure. Taking into consideration the iterative terms of the quasipotential, we can conveniently represent the corresponding expression as follows [8, 9]:

$$\begin{aligned}\mathbf{\Gamma}^{(1)}(\mathbf{p},\mathbf{q}) &= U_1^*(\mathbf{p}_1)U_2^*(\mathbf{p}_2)\frac{e_1}{2m_1}\\ &\times\{\mathbf{A}_1\Lambda_1^{(-)}(\mathbf{p}_1')\mathcal{B}_1\mathcal{B}_2\hat{V}(\mathbf{q}_2-\mathbf{p}_2)\\ &+\mathcal{B}_1\mathcal{B}_2\hat{V}(\mathbf{p}_2-\mathbf{q}_2)\Lambda_1^{(-)}(\mathbf{q}_1')\mathbf{A}_1\}\\ &\times U_1(\mathbf{q}_1)U_2(\mathbf{q}_2)+(1\longleftrightarrow 2),\end{aligned}\tag{27}$$

$$\begin{aligned}\hat{V}(\mathbf{k}) &= \mathcal{B}_1\mathcal{B}_2\left\{\left(1+\frac{\kappa_1}{2m_1}\mathbf{\Gamma}_1\cdot\mathbf{k}\right)\left(1-\frac{\kappa_2}{2m_2}\mathbf{\Gamma}_2\cdot\mathbf{k}\right)\right.\\ &\left.-\left(\mathbf{A}_1+\frac{\kappa_1}{m_1}\mathcal{B}_1 i[\mathbf{S}_1\times\mathbf{k}]\right)\left(\mathbf{A}_2-\frac{\kappa_2}{m_2}\mathcal{B}_2 i[\mathbf{S}_2\times\mathbf{k}]\right)\right\}\frac{e_1e_2}{k^2},\end{aligned}\tag{28}$$

where the one-particle operator of projection onto negative-frequency states is given by

$$\Lambda^{-}(\mathbf{p})\approx\frac{1-\mathcal{B}}{2}-\frac{\mathbf{A}\cdot\mathbf{p}}{2m},$$

and $\kappa_{1,2}$ are anomalous magnetic moments of the particles. The matrices $\mathbf{A}_{1,2}$ and $\mathcal{B}_{1,2}$ also represent natural generalizations of $\boldsymbol{\alpha}_{1,2}$ and $\beta_{1,2}$ to the case of spin-1/2 particles (by analogy with (22)):

$$\mathbf{A} = \begin{pmatrix}0 & \mathbf{\Sigma}\\ \mathbf{\Sigma} & 0\end{pmatrix},\quad \mathcal{B} = \begin{pmatrix}I & 0\\ 0 & -I\end{pmatrix}.\tag{29}$$

To order $(v/c)^2$, both terms of potential (28) contribute to the magnetic moment of the system. Substituting (27), (28), and (9) into (11) and calculating the derivative with respect to $\mathbf{\Delta}$ in (10), we obtain

$$\begin{aligned}\mathbf{M} &= \frac{1}{(2\pi)^3}\int d\mathbf{p}\Psi_0^*(\mathbf{p})\frac{e_1}{2\epsilon_1(\mathbf{p})}\left\{2(1+\kappa_1)\right.\\ &+\mathbf{s}_1[1+N_1+N_2]+(1+4\kappa_1)\frac{[\mathbf{p}\times[\mathbf{s}_1\times\mathbf{p}]]}{2m_1^2}\\ &+(1+\kappa_2)\frac{[\mathbf{p}\times[\mathbf{s}_2\times\mathbf{p}]]}{m_1m_2}\frac{\mathbf{\Sigma}_1^2}{3}-\frac{\epsilon_2(\mathbf{p})}{M}\\ &\times\left[1+N_1+N_2+\frac{(M-\epsilon_1-\epsilon_2)}{m_2}\frac{\mathbf{\Sigma}_1^2}{3}\right]i\left[\mathbf{p}\times\frac{\partial}{\partial\mathbf{p}}\right]\\ &\left.+\frac{1}{2M}\left[\mathbf{p}\times\left[\mathbf{p}\times\left(\frac{\mathbf{s}_1}{m_1}-\frac{\mathbf{s}_2}{m_2}\right)\right]\right]\right\}\Psi_0(\mathbf{p})+(1\longleftrightarrow 2),\end{aligned}\tag{30}$$

where

$$N_i = \frac{\mathbf{p}^2-(\mathbf{\Sigma}_i\cdot\mathbf{p})(\mathbf{\Sigma}_i\cdot\mathbf{p})}{2m_i^2}.\tag{31}$$

In the case of *S* states, expression (30) is substantially simplified:

$$\mathbf{M} = \frac{1}{2}g_{1\,\mathrm{bound}}\frac{e_1}{m_1}\langle\mathbf{s}_1\rangle+\frac{1}{2}g_{2\,\mathrm{bound}}\frac{e_2}{m_2}\langle\mathbf{s}_2\rangle,\tag{32}$$

where the *g* factors of bound particles are given by

$$\begin{aligned}g_{1\,\mathrm{bound}} &= g_1\left\{1-\frac{\langle\mathbf{p}^2\rangle}{3m_1^2}\left[1-\frac{3\kappa_1}{2(1+\kappa_1)}\right]\right.\\ &+\frac{\langle\mathbf{p}^2\rangle}{2m_1^2}\left[1-\frac{\langle\mathbf{\Sigma}_1^2\rangle}{3}+\frac{m_1^2}{m_2^2}\left(1-\frac{\langle\mathbf{\Sigma}_2^2\rangle}{3}\right)\right]\\ &\left.+\frac{e_2}{e_1}\frac{\langle\mathbf{p}^2\rangle}{3m_2^2}\frac{\langle\mathbf{\Sigma}_2^2\rangle}{3}-\frac{\langle\mathbf{p}^2\rangle}{(1+\kappa_1)6m_1(m_1+m_2)}\left(1-\frac{e_2m_1}{e_1m_2}\right)\right\},\end{aligned}\tag{33}$$

$$g_{2\,\mathrm{bound}} = g_{1\,\mathrm{bound}}(1\longleftrightarrow 2),\quad \frac{1}{2}g_{1,2} = 1+\kappa_{1,2}.$$

For a hydrogen-like ion (index 1 corresponds to the electron, and index 2 to the nucleus), we have

$$\begin{aligned}&e_1 = -e,\quad e_2 = Ze,\quad \langle\mathbf{p}^2\rangle = \frac{m_1^2m_2^2(Z\alpha)^2}{(m_1+m_2)^2},\\ &\frac{1}{2}g_e = 1+\kappa_1,\quad \frac{1}{2}g_N = 1+\kappa_2,\\ &K_{s_1} = \frac{\langle\mathbf{\Sigma}_1^2\rangle}{3} = 1,\quad K_{s_2} = \frac{\langle\mathbf{\Sigma}_2^2\rangle}{3}\end{aligned}\tag{34}$$

$$= \begin{cases}\dfrac{4s_2}{3}, & s_2 \text{ is an integer spin of the nucleus}\\[2ex] \dfrac{4s_2+1}{3}, & s_2 \text{ is a half-integer spin of the nucleus,}\end{cases}$$

so that the *g* factors of the electron and nucleus in the bound state are given by

$$\begin{aligned}g_{e\,\mathrm{bound}} &= g_e\left\{1-\frac{m_2^2(Z\alpha)^2}{3(m_1+m_2)^2}\left[1-\frac{3\kappa_1}{2(1+\kappa_1)}\right.\right.\\ &-\frac{3}{2}(1-K_{s_1})-\frac{3m_1^2}{2m_2^2}\left(1-K_{s_2}-\frac{2}{3}ZK_{s_2}\right)\\ &\left.\left.+\frac{m_1}{2(m_1+m_2)(1+\kappa_1)}\left(1+Z\frac{m_1}{m_2}\right)\right]\right\},\end{aligned}\tag{35}$$

$$\begin{aligned}g_{N\,\mathrm{bound}} &= g_N\left\{1-\frac{m_1^2(Z\alpha)^2}{3(m_1+m_2)^2}\left[1-\frac{3\kappa_2}{2(1+\kappa_2)}\right.\right.\\ &-\frac{3}{2}(1-K_{s_2})-\frac{3m_2^2}{2m_1^2}\left(1-K_{s_1}-\frac{2}{3Z}K_{s_1}\right)\\ &\left.\left.+\frac{m_2}{2(m_1+m_2)(1+\kappa_2)}\left(1+\frac{m_2}{Zm_1}\right)\right]\right\},\end{aligned}\tag{36}$$

where we retained the coefficients K_{s_1} and K_{s_2} in the general form to demonstrate the symmetry of relations (35) and (36) with respect to the replacement $1 \longleftrightarrow 2$. Expressions (35) and (36), which are obtained from graphs *a* and *b* in the figure, contain the corrections of order $O(\alpha^2)$ and $O(\alpha^3)$ associated with the bound-state effects. Here, the interaction of particles in the bound state are considered in the one-photon approximation. The corrections of order $O(\alpha^3)$ are attributed to the terms of the quasipotential that are proportional to the anomalous magnetic moments κ_1 and κ_2 of the particles. Formulas (30), (35), and (36) are the generalizations of the expression for the magnetic moment of a hydrogen-like atom and gyromagnetic factors for spin-1/2 particles, obtained in [8, 9], to the case of particles of arbitrary spin. Expression (30) for the magnetic moment contains the terms corresponding to the interaction of free nonrelativistic charged point particles with an external electromagnetic field, as well as a number of corrections to these terms associated with the bound-state effects. Some of these corrections, quadratic in the spin operator $\mathbf{\Sigma}$ (the terms proportional to $\sim N_i$), are determined by the transformation properties of the spinors ξ and η (5) and can be interpreted as the relativistic corrections for a particle of spin *s*. Other corrections, which are associated with the general structure of the matrix element of current J_μ in (3), relate to the dipole interaction. Our calculations by formulas (35) and (36) show that the terms of order $O(\alpha^3)$ in the expressions for $g_{e\,\mathrm{bound}}$ and $g_{N\,\mathrm{bound}}$ depend on the spin of the second particle, the nucleus, in contrast to [15], where no such dependence was observed. From the experimental point of view, of special interest are the electron *g* factors of the hydrogen, deuterium, and tritium atoms, as well as the ratios of these factors [1]. The experimental value of the ratio $g_{e\,\mathrm{H}}/g_{e\,\mathrm{D}}$ obtained in [13] is given by

$$r^{\exp} = \left[\frac{g_{e\,\mathrm{H}}}{g_{e\,\mathrm{D}}}\right]^{\exp} = 1 + 7.22(3) \times 10^{-9} \qquad (37)$$

to a high degree of accuracy. An analytical expression for this ratio is obtained from (35):

$$r^{th} = \left[\frac{g_{e\,\mathrm{H}}}{g_{e\,\mathrm{D}}}\right]^{\mathrm{th}} = 1 + \alpha^2\left[\frac{1}{4}\frac{m_1}{m_2} - \frac{25}{72}\frac{m_1^2}{m_2^2} - \frac{\alpha}{\pi}\left(\frac{m_1}{24m_2} - \frac{1}{16}\frac{m_1^2}{m_2^2}\right)\right]. \qquad (38)$$

The numerical value of (38) is given by

$$r^{\mathrm{th}} = 1 + 7.237 \times 10^{-9},$$

which is in good agreement with (37).

The problem of relativistic description of the interaction of a massive particle of arbitrary spin with an electromagnetic field has been studied over a long period of time [17–25]; however, this problem has not yet been solved completely. It is well known that the application of the minimal coupling principle

$$\partial_\mu \phi \longrightarrow \partial_\mu \phi + ieA_\mu \phi = D_\mu \phi \qquad (39)$$

to relativistic equations for particles of spin $s \geq 1$ leads to the following fundamental difficulties [17–27]:

1. The system of partial differential equations describing a spin particle may become inconsistent after substitution (39). For example, such a situation occurs with the Proca equation for a vector particle.

2. The introduction of minimal coupling into the equation describing a free particle of fixed spin *s* may lead to an equation that cannot be interpreted as the equation of motion of an object with spin *s*, since the corresponding wave function has redundant components.

3. The equation for a particle of spin $s > 1$ that minimally interacts with an electromagnetic field is relativistically invariant; however, it still describes the propagation of a wave with velocity $v > c$.

In the nonrelativistic limit, different relativistic wave functions for a particle of spin *s* lead to different Pauli-type equations. As we noted above, Eides and Grotch [15] proposed a method for calculating the contributions of different orders to the magnetic moment of a weakly coupled system on the basis of a relativistic semiclassical equation of motion of spin. The Hamiltonian of interaction of a particle of arbitrary spin with an external electromagnetic field constructed in [15] on the basis of the above equation yields the *g* factors of particles in the bound state that are independent of their spins. The BMT equation is approximate: it is linear in the spin of the particle and the field $F_{\mu\nu}$; the dependence of the field on the coordinates is neglected. When a particle of spin *s* is in a bound state in a uniform external magnetic field, the terms neglected in the approximation of the BMT equation may make a certain contribution to the *g* factors of bound particles. In this paper, when calculating the contributions (depending on the spin of the nucleus) to the *g* factors of bound particles of a hydrogen-like ion, we applied the method proposed in [26, 27] for describing the interaction of a particle of arbitrary spin with an electromagnetic field. New (as compared with [8]) contributions to (35), (36), and (38) have appeared as a result of the substitution of operators (29) for ordinary generators of boosts α in the case of a spin-1/2 particle. Thus, our approach to the description of the interaction of a particle of arbitrary spin with an electromagnetic field is based on the application of the matrix element of electromagnetic current (11) and the Lorentz transformation of the spinor wave functions ξ and η (5). We obtained additional spin-dependent terms in the expressions for the *g* factors of bound particles (see (33)) that arise when one takes into account the terms nonlinear in the spin operator $\mathbf{\Sigma}$ in the operator of interaction of particles. The Hamiltonian obtained in [15] does not contain such nonlinear terms of order $O((v/c)^2)$ for particles of arbitrary spin s. The terms thus calculated in the expressions for the gyromagnetic factors of particles actually represent the contribution of contact terms of the quadrupole interaction investigated in [26, 27]. In the case of spin-1/2 particles, these terms correctly reproduce the well-known result obtained independently in [8, 10, 11, 22] by different approaches. The correction to r^{th} associated with the spin of a deuteron, such that $I = 1$, $Z = 1$, $m = 2m_2$ (m_2 is the mass of a proton), is given by

$$\Delta r^{\text{th}} = 5\alpha^2 m_1^2/72 m_2^2 = 0.001 \times 10^{-9}$$

and, as follows from (37), lies within experimental error bounds. The corrections to (35) and (36) associated with the spin of the nucleus also depend on Z and N, the number of nucleons in the nucleus. In spite of the fact that these terms increase ($\propto Z^3$) with Z, an increase in the number N of nucleons in the nucleus leads to the opposite effect. Therefore, in the case of ions of the type $^{12}C^{5+}$ with the spin $I \neq 0$, the numerical values of these corrections also lie beyond experimental error bounds. The spin I of the $^{12}C^{5+}$ ion itself is equal to zero; therefore, $K_I = 0$, and the corresponding spin correction also vanishes. At present, there are experimental data on the electron g factor in $^{16}O^{7+}$ and $^{32}S^{14+}$ ions [4], whose nuclei also have spin $I = 0$. In our opinion, it would be interesting to measure the g factors of particles using Penning traps [3, 4] for such ions in which, on the one hand, the spin $I \neq 0$ and, on the other hand, the ratio Z^3/N^2 attains large values. The $^{59}Co^{26+}$ ion with $I = 7/2$ and $Z^3/N^2 \approx 5.65$ and the spin-dependent correction to (35) equal to 0.1×10^{-9} can serve as one of such ions. As was pointed out in the report made by Quint [4], one can expect that the g factor of a bound electron will be measured to an accuracy higher than 1 ppb in the nearest future.

ACKNOWLEDGMENTS

We are grateful to V.M. Shabaev and W. Quint for providing new experimental data about the g factors of bound particles and S.G. Karshenboim and I.B. Khriplovich for useful discussions concerning the problem of the magnetic moment of bound states.

This work was supported by the Russian Foundation for Basic Research (project no. 00-02-17771) and by the Program "Universities of Russia: Fundamental Research" (project no. 990192).

REFERENCES

1. P. J. Mohr and B. N. Taylor, Rev. Mod. Phys. **72**, 351 (2000).
2. T. Beier, Phys. Rep. **339**, 79 (2000).
3. N. Hermanspahn, H. Haffner, H.-J. Kluge, *et al.*, Phys. Rev. Lett. **84**, 427 (2000).
4. W. Quint *et al.*, in *Proceedings of the 2nd Workshop on Quantum Electrodynamics and Physics of the Vacuum, QED 2000*, Ed. by D. Cantatore, AIP Conf. Proc. **564**, 144 (2001).
5. V. M. Shabaev, Can. J. Phys. **76**, 907 (1998).
6. S. G. Karshenboim, Phys. Lett. A **266**, 380 (2000).
7. A. Czarnecki, K. Melnikov, and A. Yelkhovsky, Phys. Rev. A **63**, 012509 (2001).
8. R. N. Faustov, Phys. Lett. B **33B**, 422 (1970).
9. R. N. Faustov, Nuovo Cimento A **69**, 37 (1970).
10. H. Grotch, Phys. Rev. Lett. **24**, 39 (1970).
11. H. Grotch and R. A. Hegstrom, Phys. Rev. A **4**, 59 (1971).
12. S. A. Blundell, K. T. Cheng, and J. Sapirstein, Phys. Rev. A **55**, 1857 (1997).
13. F. G. Walther, W. D. Phillips, and D. Kleppner, Phys. Rev. Lett. **28**, 1159 (1972).
14. H. Persson, S. Salomonson, P. Sunnergren, and I. Lindgren, Phys. Rev. A **56**, R2499 (1997); T. Beier, I. Lindgren, H. Persson, *et al.*, Phys. Rev. A **62**, 032510 (2000).
15. M. I. Eides and H. Grotch, Ann. Phys. **260**, 191 (1997).
16. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics* (Nauka, Moscow, 1989).
17. L. P. S. Singh and C. R. Hagen, Phys. Rev. D **9**, 898, 910 (1974).
18. S. Weinberg, in *Lectures on Elementary Particles and Quantum Field Theory*, Ed. by S. Deser, M. Grisaru, and H. Pendleton (MIT Press, Cambridge, 1970).
19. I. B. Khriplovich, Zh. Éksp. Teor. Fiz. **96**, 385 (1989) [Sov. Phys. JETP **69**, 217 (1989)].
20. S. Ferrara, M. Poratti, and V. L. Telegdi, Phys. Rev. D **46**, 3529 (1992).
21. S. D. Brodsky and J. R. Primack, Ann. Phys. **52**, 315 (1969).
22. F. E. Close and H. Osborn, Phys. Lett. B **34B**, 400 (1971).
23. V. I. Fushich and A. G. Nikitin, *Symmetry of Equations of Quantum Mechanics* (Nauka, Moscow, 1990).
24. S. Deser, V. Pascalutsa, and A. Waldron, Phys. Rev. D **62**, 105031 (2000).
25. Yu. M. Zinoviev, in *Proceedings of the XVII Seminar on High Energy Physics and Field Theory, Protvino, 1994.*
26. I. B. Khriplovich, A. I. Milstein, and R. A. Sen'kov, Zh. Éksp. Teor. Fiz. **111**, 1935 (1997) [JETP **84**, 1054 (1997)].
27. A. A. Pomeranskiĭ and I. B. Khriplovich, Zh. Éksp. Teor. Fiz. **113**, 1537 (1998) [JETP **86**, 839 (1998)].
28. Yu. B. Rumer and A. I. Fet, *Group Theory and Quantized Fields* (Nauka, Moscow, 1977).
29. R. N. Faustov, Ann. Phys. **78**, 176 (1973).
30. A. P. Martynenko and R. N. Faustov, Theor. Math. Phys. **64**, 765 (1985).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 477–484.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 546–554.
Original Russian Text Copyright © 2001 by Karshenboim, Ivanov, Shabaev.

ATOMS, SPECTRA, RADIATION

Vacuum Polarization in a Hydrogen-like Relativistic Atom: *g* Factor of a Bound Electron

S. G. Karshenboim[a], *, V. G. Ivanov[b], and V. M. Shabaev[c]

[a]*State Research Center Mendeleev All-Russia Research Institute of Metrology, St. Petersburg, 198005 Russia*
**e-mail: sgk@vniim.ru, sek@mpq.mpg.de*

[b]*Pulkovo Observatory, Russian Academy of Sciences, Pulkovskoe shosse 65, St. Petersburg, 196140 Russia*

[c]*St. Petersburg State University, Ul'yanovskaya ul. 1, Petrodvorets, 198904 Russia*

Received April 12, 2001

Abstract—The contribution of vacuum polarization to the *g* factor of a bound electron is considered for the ground state of a hydrogen-like atom. A final expression for the correction is obtained in terms of generalized hypergeometric functions and represents a function of parameter $Z\alpha$ and the ratio of the mass of the orbital particle (electron or muon) to the mass of a particle in the vacuum loop. Different asymptotic forms of this expression are derived for both common and muon atoms. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

There exist several characteristics of simple atoms of theoretical interest that can be determined experimentally to a high degree of accuracy. Until recently, one considered only the Lamb shift and hyperfine splitting among these characteristics, as well as certain associated intervals of fine and gross structure of the atomic spectrum. Recently, precision measurements of the *g* factor of a bound electron have been carried out in the hydrogen-like ion of carbon ($^{12}C^{5+}$) [1]. A number of corrections to the *g* factor of a bound electron are presented in [2–6] (see also the recent studies [7, 8] and references therein). However, the state of the art of the theory cannot be regarded as quite satisfactory. In particular, this circumstance is attributed to the fact that the experiments in [1, 9] allow one to obtain the results for different ions; at present, the processing of data for a hydrogen-like oxygen is approaching the end [10]. The requirements imposed on the theory also depend on the nuclear charge Z. For instance, the accuracy of the experiments carried out about 30 years ago with hydrogen [11], deuterium [11, 12], tritium [13], and the helium ion [14] was sufficiently high: the relative error was 10^{-8} for hydrogen ($Z = 1$) [11, 15] and 3×10^{-7} for helium ($Z = 2$) [14, 16]. Since the nuclear charge Z in this case is small, the nontrivial corrections of order $\alpha(Z\alpha)^4$ or higher associated with the quantum electrodynamics (QED) of bound states did not play any appreciable role. However, the accuracy of the experiments with hydrogen-like ions of carbon and oxygen [9] is several times higher (the error amounts to 2×10^{-9}), and, which is more important, for $Z = 6$ and 8, the nontrivial QED corrections of order $\alpha(Z\alpha)^4$ prove to be three orders of magnitude greater. This fact radically changes the situation, and the *g* factor of a bound electron in such ions proves to be highly sensitive to the effects of quantum electrodynamics of bound states. Further experimental plans involve the hydrogen-like ions of calcium ($Z = 20$); in this case, the measurements of the *g* factor of a bound electron are as sensitive to higher order QED corrections as the standard experiments on the determination of the Lamb shift in hydrogen [17].

To facilitate the comparison of the theory with experiment, one should analyze a number of higher order corrections and carry out calculations without expanding in terms of the parameter $Z\alpha$. Similar calculations were carried out numerically in the one-loop approximation [5, 6, 8]. At the same time, it is known that various phenomena associated with the free polarization of vacuum admit analytical calculations for a Dirac electron [18–21]. The present study is devoted to the exact (without expanding in terms of $Z\alpha$) analytical calculation of the contribution of the vacuum polarization to the *g* factor of a bound electron (some preliminary results were briefly presented in [22]) and is a continuation of our studies [19, 21],[1] where we obtained expressions for the contribution of vacuum polarization to the Lamb shift and the hyperfine structure of a hydrogen-like atom. In this paper, we present an exact (with respect to $Z\alpha$) analytical result concerning the correction, induced by the free polarization of vacuum, to the *g* factor of a bound particle in the ground state of a two-particle atomic system. All calculations are performed under the assumption of an infinitely heavy point nucleus. For medium values of the nuclear charge Z that correspond to the experiments of [1, 9], one can easily take into account, when necessary, the finiteness of the nucleus size as a perturbation [23]. The results obtained apply to both electron and muon atoms.

---

[1] Certain misprints made in [19] are discussed in the Appendix.

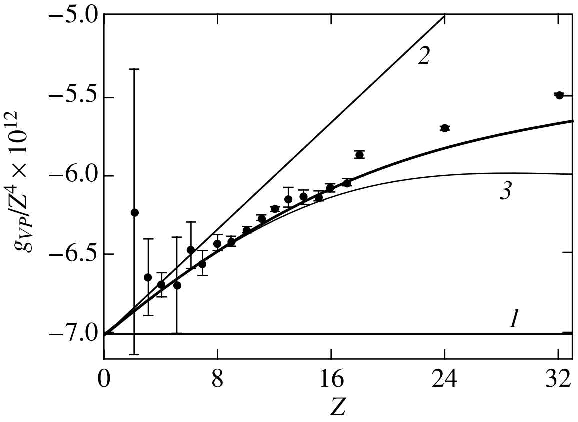


Fig. 1. Comparison of numerical data [6] (dots) for $g_{VP}(Z)$ and analytical results obtained in this paper; the heavy curve represents the full analytical expression (see (17)), while the thin lines represent the asymptotics (*1*) $g^{(a)} = -16\alpha(Z\alpha)^4/15\pi$ [23], (*2*) $g^{(b)} = g^{(a)} + 5\alpha(Z\alpha)^5/9$, and (*3*) $g^{(c)} = g^{(b)} + [(16/15)\ln(2Z\alpha) - 2012/525]\alpha(Z\alpha)^6/\pi$ (cf. (1)).

In this work, we also found certain useful asymptotic expressions. In particular, the polarization correction to the g factor of an electron in the ground state for medium values of Z can be represented as

$$\begin{aligned} g_{VP} = \frac{\alpha}{\pi}\Bigg[&-\frac{16}{15}(Z\alpha)^4 + \frac{5\pi}{9}(Z\alpha)^5 \\ &+ (Z\alpha)^6\left(\frac{16}{15}\ln(2Z\alpha) - \frac{2012}{525}\right) \\ &+ (Z\alpha)^7\left(-\frac{5\pi}{9}\ln\left(\frac{Z\alpha}{2}\right) + \frac{125\pi}{216}\right)\Bigg] \end{aligned} \qquad (1)$$

to a sufficient degree of accuracy. This expression reproduces the known contribution of order $\alpha(Z\alpha)^4$ [23] and contains new expansion terms. The results obtained agree with the numerical results for an electron atom [6]. This fact is illustrated in Fig. 1, which reproduces the analytical results obtained below and the results of numerical calculations performed in [6]. One can easily see that asymptotics (1) is in good agreement both with the full analytical expression and with the numerical results. The latter results involve an additional contribution due to the Wichmann–Kroll potential. For small Z, this contribution is of order $\alpha(Z\alpha)^6$, which explains the small difference between our data and the numerical results (they have the same order of magnitude). In view of certain physical reasons, one should expect that this contribution, which takes into account the photon–photon interaction (the so-called scattering of light by light), has small numerical coefficients. It is this fact that explains the good agreement between asymptotics (1) and the results of analytical and numerical calculations.

The structure of this paper is as follows. In Section 2, we derive a general expression for the correction to the g factor of a bound electron (muon) as a function of two parameters: the nuclear charge Z and the ratio of the mass m of the particle bound in the atom (electron or muon) to the mass m_e of the particle in the polarization loop (electron). In this paper, we follow the notation of [19, 21]. In particular, we use the following quantities instead of the two aforementioned parameters:

$$\epsilon = 1 - \sqrt{1-(Z\alpha)^2} \approx \frac{(Z\alpha)^2}{2}$$

and

$$\kappa = \frac{Z\alpha m}{m_e},$$

which better correspond to the physical sense of the problem. In Section 3, we analyze various asymptotics corresponding to various limit values of the parameters ϵ and κ. In this paper, we use the relativistic system of units $\hbar = c = 1$.

2. GENERAL EXPRESSION

The energy levels of an electron in an external magnetic field **H** are determined by the effective Hamiltonian

$$\Delta H = -g\mu_B(\mathbf{j}\cdot\mathbf{H}), \qquad (2)$$

where μ_B is the Bohr magneton of the electron (muon); **j** is the total angular momentum of the electron; and g is its g factor, which depends on the electron configuration. In this paper, we consider the ground state of a hydrogen-like ion with the spinless nucleus, so that the total momentum of the ion, the total momentum of the electron, and its spin are identical.

In the ground state of a hydrogen-like atom, the g factor of a relativistic bound electron

$$g = 2(1 + a + b) \qquad (3)$$

differs from its free value

$$g^{(0)}_{\text{Dirac}} = 2, \qquad (4)$$

which is determined by the Dirac equation, even in the absence of radiative phenomena such as the vacuum polarization. This situation is attributed to the fact that the effective Hamiltonian (2) has a nonrelativistic form; therefore, all relativistic phenomena are contained in the parameter g, which depends on the state of the system. In general, Hamiltonian (2) also has off-diagonal matrix elements; however, in the problem of relativistic corrections to the electron energy in a weak external magnetic field, it suffices to restrict the analysis to the diagonal elements.

For the ground state, the leading correction for the binding phenomena is given by the Breit formula [2]

$$\begin{aligned} g_{Br} &= -\frac{8m}{3}\int_0^\infty f_{1s} g_{1s} r^3 dr \\ &= \frac{2(3-2\epsilon)}{3} \approx 2\left(1-\frac{1}{3}(Z\alpha)^2+\ldots\right). \end{aligned} \tag{5}$$

The consideration of the QED phenomena changes the value of the g factor in the case of a free electron,

$$a = a_1 + a_2 + \ldots = \frac{\alpha}{2\pi} - 0.328\ldots\left(\frac{\alpha}{2\pi}\right)^2 + \ldots \tag{6}$$

(where a_n is an n-loop contribution to the anomalous magnetic moment of electron [24]), as well as in the case of a bound electron [3, 4],

$$b = \left(-\frac{1}{3} + \frac{\alpha}{12\pi} + \frac{m}{2M} + \ldots\right)(Z\alpha)^2 + \ldots \tag{7}$$

(where M is the nuclear mass). It should be noted that all corrections of order $(Z\alpha)^2$ have a purely kinematic character [3, 4] and represent a solution of a two-particle problem in which a free electron (muon) has a nonzero anomalous magnetic moment a and the nuclear mass M is finite. The nontrivial QED phenomena corresponding to the quantum electrodynamics of bounded states only conribute to the terms of order $\alpha(Z\alpha)^4$; therefore, it suffices to take into account these terms in the limit of external field, when $m/M = 0$. In particular, one of such corrections is attributed to the effects of vacuum polarization.

The correction to the g factor of a bound electron induced by the free polarization of vacuum corresponds to the Feynman graphs shown in Fig. 2. These graphs correspond to the second-order terms of perturbation theory; one should apply the Coulomb Green's function of a relativistic electron to calculate these graphs. However, instead of calculating the second-order cross contribution for two perturbations, those of an external magnetic field and of the Uehling potential, we can reduce the problem to two first-order calculations (cf. the calculation of the correction to the hyperfine splitting in [21]). Namely, we first solve the problem of perturbation of the Dirac equation by an external homogeneous time-independent magnetic field **H** and determine the correction linear in this field to the wave functions and then calculate the matrix element of the Uehling potential over the perturbed wave functions (cf. (15) in [21]).

2.1. Wave Function with Regard to the External Magnetic Field

The correction linear in an external magnetic field to the relativistic Coulomb wave function of an s state in the hydrogen atom, which is given by [25]

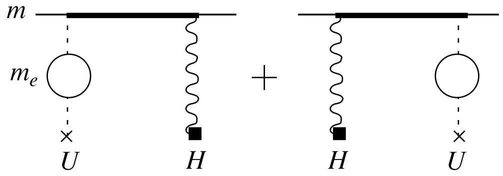


Fig. 2. The contribution of the vacuum polarization to the g factor of a bound electron. The heavy line corresponds to the reduced Green's function of a particle in the Coulomb field of the nucleus; the wavy line with a square attached to its end corresponds to the external dc magnetic field H, and the dashed line represents the Coulomb field perturbed by the Uehling potential (U).

$$\Psi_{nsm}(\mathbf{x}) = \begin{pmatrix} f_{ns}(r) \\ -i(\boldsymbol{\sigma}\cdot\mathbf{n})g_{ns}(r) \end{pmatrix}\Omega_{1/2,\,0,\,m}(\mathbf{n}),$$

in the absence of the magnetic field, can be determined by the method of generalized virial relations. This method was developed in [26]; we used it in [21]. The correction to the wave function, induced by an external dc magnetic field **H**, contains terms with different angular dependence ($l = 0$ and $l = 2$). Note that the correction to the wave function is needed in this study only for calculating the matrix element of the central potential; therefore, it suffices to consider only the s component of the correction in the first order of perturbation due to the magnetic field **H**. This component of the correction to the wave function is given by

$$\begin{aligned} (\delta\Psi_{ns,\,m})^s &= -\frac{e}{3}\begin{pmatrix} X_{ns}(r) \\ -i(\boldsymbol{\sigma}\cdot\mathbf{n})Y_{ns}(r) \end{pmatrix} \\ &\quad\times(\boldsymbol{\sigma}\cdot\mathbf{H})\Omega_{1/2,\,0,\,m}(\mathbf{n}), \end{aligned} \tag{8}$$

where

$$\begin{aligned} X_{ns} &= -\left[\frac{2\mathscr{E}_{ns}+3m}{2m^2}r + \frac{Z\alpha}{m^2}\right]g_{ns} - \frac{3}{2m^2}f_{ns}, \\ Y_{ns} &= \left[\frac{2\mathscr{E}_{ns}-m}{2m^2}r + \frac{Z\alpha}{m^2}\right]f_{ns} + \frac{1}{2m^2}g_{ns}, \end{aligned} \tag{9}$$

and

$$\begin{aligned} \mathscr{E}_{ns} &= \frac{m}{\sqrt{1+\dfrac{(Z\alpha)^2}{(n-\epsilon)^2}}} \\ &\approx m\left(1-\frac{(Z\alpha)^2}{2n^2} + \left(-\frac{1}{2n^3}+\frac{3}{8n^4}\right)(Z\alpha)^4\right) \end{aligned} \tag{10}$$

is the energy of level ns. In particular, for the radial components of the ground state, we obtain

$$\begin{aligned} X_{1s} &= -\left[\frac{5-2\epsilon}{2m}r + \frac{Z\alpha}{m^2}\right]g_{1s} - \frac{3}{2m^2}f_{1s}, \\ Y_{1s} &= -\left[\frac{1-2\epsilon}{2m}r + \frac{Z\alpha}{m^2}\right]f_{1s} + \frac{1}{2m^2}g_{1s}. \end{aligned} \tag{11}$$

2.2. Matrix Element for the g *Factor of a Bound Electron*

The next step consists in calculating the matrix element of the Uehling potential

$$U(r) = \frac{\alpha}{\pi}\int_0^1 dv \frac{v^2(1-v^2/3)}{1-v^2}\left(-\frac{Z\alpha}{r}e^{-\lambda r}\right), \tag{12}$$

where

$$\lambda = \frac{2m_e}{\sqrt{1-v^2}}, \tag{13}$$

over the wave functions perturbed by an external dc magnetic field. The correction to the g factor for an arbitrary ns state can be expressed as

$$\frac{g_{VP}}{g_{Br}} = \frac{\int_0^\infty (f_{ns}X_{ns} + g_{ns}Y_{ns})U(r)r^2 dr}{\int_0^\infty f_{ns}g_{ns}r^3 dr}. \tag{14}$$

In particular, for $n = 1$, we can rewrite the numerator of the right-hand side of (14) as (cf. [19, 21])

$$\begin{aligned}
&\frac{\alpha}{\pi}\frac{2^{1-2\epsilon}(Z\alpha m)^{4-2\epsilon}}{m(1-\epsilon)} \\
&\times \int_0^1 dv \frac{v^2(1-v^2/3)}{1-v^2}\frac{2Z\alpha m + \lambda(3-2\epsilon)}{(\lambda + 2Z\alpha m)^{3-2\epsilon}} \\
&= \frac{\alpha(Z\alpha)^2}{\pi \, 2m}\frac{3-2\epsilon}{\kappa(1-\epsilon)}\int_0^1 dy \frac{\sqrt{1-y^2}}{y^2} \\
&\times \left(1 - \frac{1-y^2}{3}\right)\left(\frac{\kappa y}{1+\kappa y}\right)^{3-2\epsilon}\left(1 + \frac{\kappa y}{3-2\epsilon}\right)
\end{aligned} \tag{15}$$

and represent the correction in terms of the base integral introduced in [18, 19]:

$$\begin{aligned}
I_{abc} &= \int_0^1 dy \frac{(1-y^2)^{a-1/2}}{y^{b-1}}\left(\frac{\kappa y}{1+\kappa y}\right)^{c-2\epsilon} \\
&= \frac{1}{2}\kappa^{c-2\epsilon}B(a+1/2, 1-b/2+c/2-\epsilon) \\
&\times {}_3F_2(c/2-\epsilon, c/2+1/2-\epsilon, 1-b/2+c/2-\epsilon; \\
&\qquad 1/2, a+3/2-b/2+c/2-\epsilon; \kappa^2) \\
&- \frac{c-2\epsilon}{2}\kappa^{c+1-2\epsilon}B(a+1/2, 3/2-b/2+c/2-\epsilon) \\
&\times {}_3F_2(c/2+1-\epsilon, c/2+1/2-\epsilon, 3/2-b/2+c/2-\epsilon; \\
&\qquad 3/2, a+2-b/2+c/2-\epsilon; \kappa^2).
\end{aligned} \tag{16}$$

The final result for the ground level is given by

$$\begin{aligned}
\frac{g_{VP}}{g_{Br}} &= -\frac{2\alpha(Z\alpha)^2}{\pi(1-\epsilon)\kappa} \\
&\times \left\{I_{133} - \frac{1}{3}I_{233} + \frac{\kappa}{3-2\epsilon}\left(I_{123} - \frac{1}{3}I_{223}\right)\right\}.
\end{aligned} \tag{17}$$

3. ASYMPTOTICS

The result obtained contains two free parameters: the nuclear charge Z and the ratio of the mass of the orbital particle to that of the particle in the vacuum loop. This result can be applied to common (electron) atoms,

$$\kappa_e = Z\alpha, \tag{18}$$

as well as to muon atoms,

$$\kappa_\mu = \frac{Z\alpha m_\mu}{m_e} \approx 1.5Z. \tag{19}$$

Expression (17) is rather cumbersome; therefore, it is expedient to consider certain of its asymptotic forms.

3.1. Corrections to the g *Factor for* $Z\alpha \ll 1$

We begin with the asymptotic form of the expression obtained for small values of $Z\alpha$ corresponding to nonrelativistic kinematics. Expanding I_{abc} in powers of ϵ (cf. [19, 21]),

$$I_{abc} = I_{abc}^{(0)} - 2\epsilon I_{abc}^{(1)} + \mathcal{O}(\epsilon^2), \tag{20}$$

we obtain

$$\begin{aligned}
\frac{g_{VP}}{g_{Br}} &= -\frac{2\alpha(Z\alpha)^2}{\pi\kappa}\left\{-\frac{4\pi}{9\kappa^2} - \frac{24-20\kappa^2+5\kappa^4}{27\kappa(\kappa^2-1)}\right. \\
&- \frac{8-12\kappa^2+3\kappa^4-2\kappa^6}{9\kappa^2(\kappa^2-1)}\mathcal{A}(\kappa) \\
&- 2\epsilon\left[-\frac{\pi(4-13\kappa^2+9\kappa^4)}{108\kappa^2(\kappa^2-1)} + \frac{12-37\kappa^2+43\kappa^4}{162\kappa(\kappa^2-1)}\right. \\
&\left.\left. + \frac{4-15\kappa^2+15\kappa^4-10\kappa^6}{54\kappa^2(\kappa^2-1)}\mathcal{A}(\kappa) + J(\kappa)\right] + \mathcal{O}(\epsilon^2)\right\},
\end{aligned} \tag{21}$$

where we introduced the following notation (cf. [27, 28]):

$$\mathcal{A}(\kappa) = \begin{cases} \dfrac{\arccos\kappa}{\sqrt{1-\kappa^2}}, & \kappa < 1 \\[2ex] \dfrac{\ln(\kappa + \sqrt{\kappa^2-1})}{\sqrt{\kappa^2-1}}, & \kappa > 1, \end{cases}$$

and (cf. [28])

$$J(\kappa) = \frac{\kappa^3}{9}\int_0^1 dy \frac{y\sqrt{1-y^2}(2+y^2)(3+\kappa y)}{(1+\kappa y)^3}\ln\left(\frac{\kappa y}{1+\kappa y}\right).$$

In the case of an ordinary (electron) atom (18), from the general expression (17), we can readily obtain the following asymptotic expansion for small $Z\alpha$:

$$\begin{aligned}\frac{g_{VP}}{g_{Br}} = \frac{\alpha}{\pi}\bigg[&-\frac{8}{15}(Z\alpha)^4 + \frac{5\pi}{18}(Z\alpha)^5 \\ &+ (Z\alpha)^6\left(\frac{8}{15}\ln(2Z\alpha) - \frac{3298}{1575}\right) \qquad (22) \\ &+ (Z\alpha)^7\left(-\frac{5\pi}{18}\ln\left(\frac{Z\alpha}{2}\right) + \frac{55\pi}{144}\right) + \mathbb{O}((Z\alpha)^8\ln(Z\alpha))\bigg].\end{aligned}$$

The first coefficient of the expansion reproduces the earlier result [23], while the higher order terms are obtained for the first time. Figure 1 shows the results of numerical calculations of $g_{VP}(Z)$ for a bound electron, an exact analytical expression, and several approximate formulas for $Z\alpha \ll 1$. The first approximation corresponds to the term of order $(Z\alpha)^4$ [23], while the second and third approximations contain the terms of order $(Z\alpha)^5$ and $(Z\alpha)^6$, respectively.

3.2. The g Factor of a Bound Muon: $\kappa \gg 1$

In the case of muon atoms, we are interested in the asymptotics of the expression obtained for large ratio of the mass of the orbital particle to that of the electron. Passing to the limit in the exact formula (17), we readily obtain the following asymptotics for large κ:

$$\begin{aligned}\frac{g_{VP}}{g_{Br}} &= -\frac{2\alpha(Z\alpha)}{\pi(1-\epsilon)} \\ &\times\left\{\frac{2}{3(3-2\epsilon)}\left(\ln(2\kappa) + \psi(1) - \psi(3-2\epsilon) - \frac{5}{6}\right)\right. \qquad (23) \\ &\left. + \frac{1}{3(1-\epsilon)} + \frac{1-\epsilon}{2\kappa^2} + \mathbb{O}\left(\frac{\ln\kappa}{\kappa^3}\right)\right\}.\end{aligned}$$

A particular case of formula (23) is the expression for the double limit for $Z\alpha \ll 1$ and $\kappa \gg 1$, which is of practical interest for moderately light muon atoms (19):

$$\begin{aligned}\frac{g_{VP}}{g_{Br}} = -\frac{2\alpha(Z\alpha)^2}{\pi}\bigg\{&\left(\frac{2}{9}\ln(2\kappa) - \frac{5}{27} + \frac{1}{2\kappa^2}\right) \\ &+ (Z\alpha)^2\left(\frac{5}{27}\ln(2\kappa) - \frac{61}{162} + \frac{\pi^2}{27}\right) \qquad (24) \\ &+ (Z\alpha)^4\left(\frac{53}{324}\ln(2\kappa) + \frac{13\pi^2}{324} - \frac{679}{1944} - \frac{1}{9}\psi''(3)\right) \\ &+ \mathbb{O}((Z\alpha)^6\ln\kappa) + \mathbb{O}\left(\frac{\ln\kappa}{\kappa^3}\right)\bigg\},\end{aligned}$$

where $\psi(x) = \Gamma'(x)/\Gamma(x)$ is the logarithmic derivative of the gamma function. The first two terms of this asymptotics (the leading contribution and the correction of relative order $((Z\alpha)^2)$ can also be obtained by passing to the limit for large κ in expression (21).

3.3. Logarithmic Contributions for $\kappa \gg 1$

Expressions (24) and (23) contain a number of logarithmic terms. An independent calculation of these terms may serve as another method for verifying the asymptotics. All logarithmic terms have a transparent physical meaning in the limit $\ln\kappa \gg 1$ and can easily be determined from renormalization-group considerations; this proves to be useful for the verification of the asymptotics obtained. Indeed, to calculate these terms, it suffices to take into account the known expression for the running coupling constant

$$Z\alpha(\kappa) \approx Z\alpha\left(1 + \frac{2\alpha}{3\pi}\ln\kappa\right)$$

in the relativistic g factor (5). In the nonrelativistic case $Z\alpha \ll 1$, we obtain the following expression for the logarithmic part of expansion (24):

$$\begin{aligned}g_{VP}^{(\log)} &= -\frac{\alpha}{\pi}(Z\alpha)^2 \\ &\times\left(\frac{8}{9} + \frac{4}{9}(Z\alpha)^2 + \frac{4}{3}(Z\alpha)^4 + \ldots\right)\ln\kappa,\end{aligned} \qquad (25)$$

which coincides with the result of direct calculation.

Similarly, the logarithmic part g_{VP} of formula (23), which is exact in $Z\alpha$, proves to be equal to (cf. [19, 21])

$$g_{VP}^{(\log)} = -\frac{8\alpha(Z\alpha)^2\ln\kappa}{9\pi(1-\epsilon)} = \frac{\partial g_{Br}}{\partial\ln(Z\alpha)}\frac{2\alpha}{3\pi}\ln\kappa. \qquad (26)$$

The reproduction of the logarithmic terms justifies the asymptotics (23) and (24) obtained above.

4. DISCUSSION OF THE RESULTS

Above, we obtained general expression (17), which applies both to ordinary (electron) and muon atoms. The results for the g factor of a bound electron obtained here are in agreement with the analytical [23] and numerical [6] results obtained earlier. We derived compact asymptotics for the g factor of a bound electron (22) and muon (24) in two-particle atomic systems.

In the case of a muon atom, the correction obtained is a leading quantum-electrodynamic correction. For moderately large values of the nuclear charge (i.e., when $Z\alpha \ll 1$ and $Z\alpha m/m_e \gg 1$), the g factor of a muon atom is given by

$$g_\mu = 2\left\{1 + \left[-\frac{1}{3} + \frac{1}{2}\frac{m}{M} - \frac{1+Z}{2}\left(\frac{m}{M}\right)^2 + \frac{\alpha}{2\pi}\left(\frac{1}{6} - \frac{1}{3}\frac{m}{M}\right)\right]\right.$$

$$\times (Z\alpha)^2 - \frac{1}{12}(Z\alpha)^4 - \frac{2\alpha(Z\alpha)^2}{\pi} \qquad (27)$$

$$\times\left[\frac{2}{9}\ln\left(\frac{2Z\alpha m}{m_e}\right) - \frac{5}{27} + \frac{m_e^2}{2(Z\alpha m)^2} + \mathcal{O}(m/M)\right]\Bigg\}.$$

Let us briefly discuss the importance of studying the g factor of a bound electron in hydrogen-like systems for experimental investigations. The most precise experimental results were obtained for the carbon ion ($^{12}C^{5+}$) [1], while the experiment with the oxygen ion ($^{16}O^{7+}$) [10] is nearing completion. The comparison of the theory with experimental results is illustrated in the table, which summarizes all theoretical expressions for the contributions to the g factor. The anomalous magnetic moment a of a free electron involved in expres-

Contributions to the g factor of a bound electron in hydrogen-like ions of carbon ($^{12}C^{5+}$) and oxygen ($^{16}O^{7+}$)

Contribution	Expression	Reference	$g(^{12}C^{5+})$, 10^{-9}	$g(^{16}O^{7+})$, 10^{-9}
$b_1^{\rm rel}$	$-\frac{1}{3}(Z\alpha)^2 - \frac{1}{12}(Z\alpha)^4 - \frac{1}{24}(Z\alpha)^6 - \frac{5}{192}(Z\alpha)^8$	[2]	−639322.8	−1136998.5
$b_1^{\rm rec}$	$\left[\frac{1}{2}\frac{m}{M} - \frac{1+Z}{2}\left(\frac{m}{M}\right)^2\right](Z\alpha)^2$	[4]	43.8	58.4
$a_1 b_a$	$\frac{1}{2}\frac{\alpha}{\pi}(Z\alpha)^2\left[\frac{1}{6} - \frac{1}{3}\frac{m}{M}\right]$	[3]	371.0	659.7
$a_2 b_a$	$-\frac{0.328\ldots}{6}\left(\frac{\alpha}{\pi}\right)^2(Z\alpha)^2$	[3, 7]	−0.6	−1.0
$b_1^{\rm h.o.}$	$\pm(Z\alpha)^4\frac{m}{M}$	Estimate [7]	±0.2	±0.4
b'_{NS}	$\frac{4}{3}(Z\alpha)^4(mR_N)^2$	[23, 7]	0.2	0.8
b'_{VP}	$\frac{\alpha}{\pi}\left[-\frac{8}{15}(Z\alpha)^4 + \frac{5\pi}{18}(Z\alpha)^5\right.$ $+ (Z\alpha)^6\left(\frac{8}{15}\ln(2Z\alpha) - \frac{1006}{525}\right)$ $\left.+ (Z\alpha)^7\left(-\frac{5\pi}{18}\ln\left(\frac{Z\alpha}{2}\right) + \frac{125\pi}{432}\right)\right]$	This paper	−4.3	−13.3
b'_{WK}	$\left(\frac{38}{45} - \frac{2\pi^2}{27}\right)\frac{\alpha(Z\alpha)^6}{\pi} \pm \ldots$	[23]	0.0	0.0
b'_{SE}	$\frac{\alpha}{\pi}(Z\alpha)^4 B_{SE}(Z\alpha)$	[29, 8]	55.3(4)	151.5(5)
b'_{VP2}	$-\frac{164}{81}\left(\frac{\alpha}{\pi}\right)^2(Z\alpha)^4$	[23]	−0.0	−0.1
b'_{SE2}	$\pm\left(\frac{\alpha}{\pi}\right)^2(Z\alpha)^4 B_{SE}(Z\alpha)$	Estimate [7]	±0.1	±0.3
b	Total contribution	This paper	−638857.3(5)	−1136142.5(7)
a	Total contribution	[24]	1159652.2	1159652.2
	(g(th) − 2)/2	This paper	520794.9(5)	23509.7(7)
	(g(exp) − 2)/2		520798(2), [1]	23509(30), [10]

Note: In certain cases, an error is not explicitly given in the cited papers. In [29, 8], the uncertainties associated with the calculation of the recoil corrections and one- and two-loop contributions were presented separately; we added them as independent errors. The value of the contribution b'_{WK} is borrowed from [23], whereas the error is estimated by the value $\alpha(Z\alpha)^7$ of the contribution b'_{VP} obtained in this paper. In our calculations, we used the value α^{-1} = 137.03599976(50) [30] and the values of nuclear charge radii [31].

sion (3) is well known (see, for example, the survey [24]). The correction b due to the binding phenomena can be represented as a sum of several terms:

$$\begin{aligned} b &= b_1(Z\alpha, m/M) + ab_a(Z\alpha, m/M) \\ &\quad + b'(\alpha, Z\alpha, m/M, mR_N), \end{aligned} \tag{28}$$

which have a simple physical sense: the first two terms correspond to ordinary kinematic corrections to the Dirac magnetic moment (b_1) [2, 3] and to the anomalous magnetic moment (b_a) [4]. Originally, the calculations of the corrections [4] were performed for the hydrogen, deuterium, and helium-3 ion, which were measured to a high degree of accuracy [11–14]; therefore, since the contribution of the term ab_a is small, the first term of the expansion, $\alpha/2\pi$, was substituted into the final expressions as the quantity a. This could make an impression that a contribution of order $\alpha^2(Z\alpha)^2$ is not known; therefore, the theoretical error was substantially overestimated (see, for example, [29]). However, one can easily verify [7] that all contributions of order $(Z\alpha)^2$ have a purely kinematic character, and the correction $\alpha^2(Z\alpha)^2$ is actually contained in [4].

The last term in (28) involves all nontrivial corrections

$$b' = b'_{NS} + b'_{VP} + b'_{WK} + b'_{SE} + b'_{2-\mathrm{loop}}, \tag{29}$$

such as the corrections to the finite size of the nucleus (b'_{NS}) and to the one-loop free polarization of vacuum (b'_{VP}), the contribution of the graphs involving the blocks of scattering of light by light (b'_{WK}) and the one-loop self-energy of electron (b'_{SE}), and the contributions of certain two-loop graphs ($b'_{2-\mathrm{loop}} = b'_{VP2} + b'_{SE2}$). Explicit analytical expressions and numerical values of the contributions to the g factors of hydrogen-like ions of carbon and oxygen are presented in the table for all types of corrections.

ACKNOWLEDGMENTS

We are grateful to G. Werth, W. Quint, and Th. Beier for useful discussions and the communication of their results before publication.

This work was supported by the Russian Foundation for Basic Research (project no. 00-02-16718) and the Programs “Fundamental Metrology” and “Universities of Russia: Fundamental Research” (project no. 3930).

APPENDIX

Certain Misprints Made in [19]

The base integrals necessary for calculating the matrix elements of the Uehling potential were considered in [19], where we made several misprints. In particular, the recurrence relation (43) must have the form

$$\begin{aligned} R(n, n\epsilon_n, \kappa_n) &= \frac{\kappa_n^{2n(1-\epsilon_n)}}{(2n-1-2n\epsilon_n)(2n-2-2n\epsilon_n)} \\ &\quad \times \frac{\partial^2}{\partial\kappa_n^2}\frac{R(n-1, n\epsilon_n, \kappa_n)}{\kappa_n^{2(n-1)(1-n\epsilon_n)}}, \end{aligned} \tag{A.1}$$

whereas the general expression for the Lamb shift of an arbitrary circular state in terms of the result for the ground state (44) has the form

$$\begin{aligned} R(n, n\epsilon_n, \kappa_n) &= \kappa_n^{2n(1-\epsilon_n)}\frac{\Gamma(2-2n\epsilon_n)}{\Gamma(2n-2n\epsilon_n)} \\ &\quad \times \frac{\partial^{(2n-2)}}{\partial\kappa_n^{(2n-2)}}\frac{R(1, n\epsilon_n, \kappa_n)}{\kappa_n^{2(1-n\epsilon_n)}}. \end{aligned} \tag{A.2}$$

In [19], we omitted ϵ_n in both exponents of κ_n in both expressions. These equalities were not used in the sequel; therefore, these misprints did not affect the subsequent calculations. However, an error was made in the terms of order $1/\kappa^2$ in the asymptotics of different contributions to R ([19], formulas (36)–(38)); this error resulted in an incorrect expression for this term in the case of R(n) (formula (46)). In particular, the asymptotics of the incomplete beta function is incorrect (the denominator of the last term in the expression for $B_{1-\delta}(c - 2\epsilon, 0)$ must be equal to $4\kappa^2$). In the general case, the corrected expression (46) is as follows:

$$\begin{aligned} R(n, n\epsilon_n, \kappa_n) &= \left[\frac{2}{3}\ln(2\kappa_n)\right. \\ &\quad \left.+\frac{2}{3}(\psi(1)-\psi(2n(1-\epsilon_n)))-\frac{5}{9}\right]+\frac{\pi}{2}\frac{n(1-\epsilon_n)}{\kappa_n} \\ &\quad +\left[-\frac{n(1-\epsilon_n)}{2}-n^2(1-\epsilon_n)^2\right]\frac{1}{\kappa_n^2}+\mathcal{O}(1/\kappa_n^3). \end{aligned} \tag{A.3}$$

The corrected expression (38) is obtained from (A3) for n = 1. We are grateful to Joan Soto for drawing our attention to this misprint.

REFERENCES

1. H. Häffner *et al.*, Phys. Rev. Lett. **85**, 5308 (2000).
2. G. Breit, Nature **122**, 649 (1928).
3. H. Grotch, Phys. Rev. Lett. **24**, 39 (1970).
4. R. N. Faustov, Nuovo Cimento A **69**, 37 (1970); Phys. Lett. B **33B**, 422 (1970); H. Grotch, Phys. Rev. A **2**, 1605 (1970); H. Grotch and R. A. Hegstrom, Phys. Rev. A **4**, 59 (1971); M. I. Eides and H. Grotch, Ann. Phys. **260**, 191 (1997).
5. S. A. Blundell, K. T. Cheng, and J. Sapirstein, Phys. Rev. A **55**, 1857 (1997).
6. H. Persson, S. Salomonson, P. Sunergren, and I. Lindgren, Phys. Rev. A **56**, R2499 (1997).

7. S. G. Karshenboim, in *Hydrogen Atom: Precision Physics of Simple Atomic Systems*, Ed. by S. G. Karshenboim *et al.* (Springer-Verlag, Berlin, Heidelberg, 2001), p. 561; E-print Archives hep-ph/0008227
8. Th. Beier *et al.*, in *Hydrogen Atom: Precision Physics of Simple Atomic Systems*, Ed. by S. G. Karshenboim *et al.* (Springer-Verlag, Berlin, Heidelberg, 2001), p. 605.
9. G. Werth *et al.*, in *Hydrogen Atom: Precision Physics of Simple Atomic Systems*, Ed. by S. G. Karshenboim *et al.* (Springer-Verlag, Berlin, Heidelberg, 2001), p. 204.
10. J. Verdu *et al.*, submitted to European Conference on Atomic and Molecular Physics, Berlin (2001), http://www.ecamp7.de/ECAMP7subads.htm
11. W. M. Hughes and H. G. Robinson, Phys. Rev. Lett. **23**, 1209 (1969).
12. D. J. Larson, P. A. Valberg, and N. F. Ramsey, Phys. Rev. Lett. **23**, 1369 (1969); F. G. Walther, W. D. Phillips, and D. Kleppner, Phys. Rev. Lett. **28**, 1159 (1972).
13. D. J. Larson and N. F. Ramsey, Phys. Rev. A **9**, 1543 (1974).
14. C. E. Johnson and H. G. Robinson, Phys. Rev. Lett. **45**, 250 (1980).
15. J. S. Tideman and H. G. Robinson, Phys. Rev. Lett. **39**, 602 (1977).
16. G. M. Keiser, H. G. Robinson, and C. E. Johnson, Phys. Rev. Lett. **35**, 1223 (1975); Phys. Rev. A **16**, 822 (1977).
17. S. G. Karshenboim, in *Proceedings of the 17th International Conference om Atomic Physics, ICAP, 2000*, Ed. by E. Arimondo *et al.* (American Inst. of Physics, New York, 2001), p. 238.
18. S. G. Karshenboim, Can. J. Phys. **76**, 169 (1998).
19. S. G. Karshenboim, Zh. Éksp. Teor. Fiz. **116**, 1575 (1999) [JETP **89**, 850 (1999)].
20. S. G. Karshenboim, V. G. Ivanov, and V. M. Shabaev, Can. J. Phys. **76**, 503 (1998).
21. S. G. Karshenboim, V. G. Ivanov, and V. M. Shabaev, Zh. Éksp. Teor. Fiz. **117**, 67 (2000) [JETP **90**, 59 (2000)].
22. S. G. Karshenboim, V. G. Ivanov, and V. M. Shabaev, submitted to Can. J. Phys. (2001).
23. S. G. Karshenboim, Phys. Lett. A **266**, 380 (2000).
24. T. Kinoshita, in *Hydrogen Atom: Precision Physics of Simple Atomic Systems*, Ed. by S. G. Karshenboim *et al.* (Springer-Verlag, Berlin, Heidelberg, 2001), p. 157.
25. V. B. Berestetskiĭ, E. M. Lifshits, and L. P. Pitaevskiĭ, *Quantum Electrodynamics* (Nauka, Moscow, 1980).
26. V. M. Shabaev, J. Phys. B **24**, 4479 (1991).
27. G. E. Pustovalov, Zh. Éksp. Teor. Fiz. **32**, 1519 (1957) [Sov. Phys. JETP **5**, 1234 (1957)]; D. D. Ivanenko and G. E. Pustovalov, Usp. Fiz. Nauk **61**, 27 (1957).
28. S. G. Karshenboim, U. Jentschura, V. G. Ivanov, and G. Soff, Eur. J. Phys. D **2**, 209 (1998).
29. T. Beier *et al.*, Hyperfine Interact. **127**, 339 (2000).
30. P. J. Mohr and B. N. Taylor, Rev. Mod. Phys. **72**, 351 (2000).
31. G. Fricke *et al.*, At. Data Nucl. Data Tables **60**, 177 (1995).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 485–490.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 555–561.
Original English Text Copyright © 2001 by Krainov, Smirnov.

ATOMS, SPECTRA, RADIATION

Nuclear Fusion Induced by a Super-Intense Ultrashort Laser Pulse in a Deuterated Glass Aerogel¶

V. P. Krainov[a] **and M. B. Smirnov**[b]
[a]*Moscow Institute of Physics and Technology, Dolgoprudny, Moscow oblast, 141700 Russia*
e-mail: krainov@cyberax.ru
[b]*Max-Born Institute, Berlin, 12489 Germany*

Received April 20, 2001

Abstract—A SiO_2 aerogel with absorbed deuterium is proposed as a target for the fusion reaction $d + d \longrightarrow He^3 + n$ induced by a superintense ultrashort laser pulse. The multiple inner ionization of oxygen and silicon atoms in the aerogel skeleton occurs in the superintense laser field. All the formed free electrons are heated and removed from the aerogel skeleton by the laser field at the front edge of the laser pulse. The subsequent Coulomb explosion of the deuterated charged aerogel skeleton propels the deuterium ions up to kinetic energies of ten keV and higher. The neutron yield is estimated at up to 10^5 neutrons per laser pulse for ~200–500 ps if the peak intensity is 10^{18} W/cm^2 and the pulse duration is 35 fs. © *2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

A new laser technique with the generation of femtosecond superintense pulses is a basis of new methods for the generation of neutrons. Several schemes have been suggested [1–5] with the usage of tabletop lasers. Schwoerer *et al.* [5] demonstrated a two-step scheme of neutron generation. At the first stage, X-ray photons are emitted by a target irradiated by a superintense laser pulse, and at the second stage, X-ray photons create neutrons as a result of the interaction with the nucleus Be^9. The approach of Ditmire *et al.* [1–4] used the fusion reaction between two deuterium nuclei (deuterons) that proceeds in accordance with the classical scheme

$$d + d \longrightarrow He^3 + n.$$

In this method, large clusters of deuterium molecules are irradiated by a superintense ultrashort laser pulse, and explosion of the produced positively charged clusters consisting of deuterons only leads to the formation of a plasma where the electrons and deuterons have the energy of several keV. The fusion reaction between deuterons proceeds both during the explosion of the charged clusters and after their decay before the total plasma expansion out of the laser focal volume.

The neutron yield up to $n_n \sim 10^4$ neutrons per laser pulse was observed in experiments [3, 4] at the peak laser intensity $I \sim 10^{17}$ W/cm^2 and the pulse width 35 fs. For the laser focal volume V with the focal spot $2r = 200$ μm and the length $l = 2$ mm [4], we have

$$V = 6.3 \times 10^{-5} \text{ cm}^3;$$

the number density of deuterons is then found to be

$$N_d = (2–3) \times 10^{19} \text{ cm}^{-3}.$$

Hence, the total number of deuterons in this plasma filament is of the order $n_d \sim 10^{15}$. Thus, only one of the $n_d/n_n = 10^{11}$ deuterons takes part in the nuclear fusion reaction! This extremely low efficiency is explained by a small tunneling rate constant of the fusion reaction and by a small lifetime of the formed plasma involving the deuteron filament (~200–500 ps).

The neutron yield could be increased by increasing the typical kinetic energy of deuterons, which is small compared to the classical threshold energy 180 keV of the fusion reaction involving two deuterons. In turn, the cluster expansion leads to the formation of fast deuterons under the action of the positive electric potential of the cluster consisting of deuterons. But there is an optimum cluster size for given parameters of the laser pulse. Indeed, a small cluster size leads to a small electric potential and, hence, to a small energy of deuterons in a plasma, whereas the laser signal cannot fully ionize a large deuterium cluster. In particular, under conditions of the experiments [3, 4], the optimum cluster radius is approximately $R \sim 25$ Å, which corresponds to the cluster charge ~+3000e after the field removal of all electrons and provides the typical kinetic energy of deuterons of several keV [6] (although the fusion reactions are mainly produced from clusters with the radii 80 Å and greater).

One can increase the typical deuteron kinetic energy by increasing the laser intensity. We now suggest an alternative approach where the aerogel with absorbed deuterium is used as a target for the laser irradiation instead of the deuterium cluster beam. The character of

---

¶This article was submitted by the authors in English.

processes with the formation of fast deuterons is similar in both cases, but the aerogel method can provide a higher deuteron kinetic energy, in our opinion. In this paper, we analyze the processes resulting from the irradiation of a typical SiO_2 aerogel with absorbed deuterium by an ultrashort superintense laser pulse.

2. PROPERTIES OF THE AEROGEL TARGET

We thus consider an aerogel with absorbed deuterium as a target for a power laser pulse. An aerogel can be described within a simple model where the aerogel matter consists of bound solid balls of identical radii a. These balls form a stable rigid skeleton due to contacts between the neighboring balls. At small distances from a ball, the aerogel has a fractal structure with the fractal dimension D. The aerogel matter is a homogeneous structure starting from a distance $R_c \gg a$ called the correlation radius. The aerogel consists of a fractal matter at distances $r < R_c$. These quantities are the basic aerogel parameters [7].

We note that the radius a of the individual ball is related to the specific internal aerogel surface S (which is usually measured in m^2/g) by [7]

$$S = \frac{3}{a\rho_0}, \tag{1}$$

where ρ_0 is the (solid) mass density of the individual aerogel balls. Another relation between aerogel parameters that follows from its fractal structure is given by

$$\frac{\rho}{\rho_0} = \left(\frac{a}{R_c}\right)^{3-D}, \tag{2}$$

where $\rho \ll \rho_0$ is the average aerogel mass density. For definiteness, we use the typical SiO_2 aerogel parameters [7] in what follows:

$$S = 715\ \text{m}^2/\text{g}, \quad \rho = 0.012\ \text{g/cm}^3, \quad D = 2.3.$$

Because the solid density of the SiO_2 aerogel material is $\rho_0 = 2.1$ g/cm^3, we obtain $a = 20$ Å from Eq. (1) and $R_c = 3$ μm from Eq. (2).

We note that the maximum amount of absorbed deuterium matter obviously corresponds to its solid mass density $\rho_d = 0.17$ g/cm^3 inside the aerogel, because pores occupy the main part of the aerogel volume. The maximum amount of absorbed deuterium is therefore given by $\rho_d/\rho = 14 g_d/g_a$ (gram of deuterium matter per aerogel gram). However, the optimum amount of absorbed deuterium must be chosen at a much smaller value. We assume that the internal aerogel surface S can be covered by three deuterium layers. This assumption agrees with the surface laws in physical chemistry [8]. Because the average distance between deuterium molecules is 3.5 Å in solid deuterium matter, the thickness of deuterium matter is approximately equal to $l_d = 10.5$ Å on the surface of the aerogel. The amount of absorbed deuterium is then given by

$$\rho_d l_d S = 0.13 g_d/g_a, \tag{3}$$

and the laser radiation is mainly absorbed by the aerogel rather than deuterium. Each individual SiO_2 ball is covered by

$$n_0 = 4\pi\rho_d \frac{(l_d + a)^3 - a^3}{3M_d} = 2150$$

deuterium molecules. Here, $M_d = 6.7 \times 10^{-24}$ g is the mass of one deuterium molecule.

The individual aerogel elements can also be described by approximating these elements by cylindrical fibers that have common knots. In this model, the fiber radius a_f follows from the fractal approach relation (instead of Eq. (1)),

$$S = \frac{2}{a_f\rho_0}.$$

For the same specific area $S = 715$ m^2/g of the internal aerogel surface, we then obtain the value $a_f \approx 13$ Å. The amount of absorbed deuterium is given by the same expression (3). We assume that the average aerogel mass density is the same as in the ball model, i.e., $\rho = 0.012$ g/cm^3. Using Eq. (2), we find the correlation radius $R_c = 2.1$ μm. The total length L of all the fibers in the plasma filament having the volume V is found from the relation

$$\rho_0 \pi a_f^2 L = \rho V.$$

This gives $L = 68$ km. We now estimate the total length l_c of all the fibers inside the correlation sphere with the radius R_c. We have the obvious relation

$$l_c = \frac{(4\pi R_c^2/3)}{V} L.$$

Inserting the above values of V, L, and R_c, we find that $l_c = 4.2$ cm.

We now estimate the average distance $\delta \gg a_f$ between the neighboring fibers in the correlation sphere. The average length of one fiber in this sphere is R_c. The quantity $\delta^2 R_c$ is the volume referring to one fiber in this sphere. The quantity

$$N_f = \frac{R_c^3}{\delta^2 R_c}$$

is the number of fibers in this sphere, and therefore, $l_c = R_c N_f$. We thus find

$$\delta = R_c\left(\frac{R_c}{l_c}\right)^{1/2} \approx 0.015\ \mu\text{m}.$$

In what follows, we consider the aerogel model consisting of individual balls and cylindrical fibers.

3. AEROGEL IN A LASER FIELD

As a result of irradiating the aerogel by a superintense ultrashort laser pulse, the following processes proceed. After a certain period of time, electrons are liberated from the aerogel skeleton by the laser field and occupy aerogel pores. The aerogel skeleton then consists of multicharged atomic ions of silicon and oxygen; deuterium nuclei (deuterons) are located on the internal skeleton surface. At the next evolution stage, all the atomic ions fly into the surrounding space forming a uniform plasma that fills all the aerogel space; deuterons fly first because they are light particles. They are located on the skeleton surface, and therefore, deuterons have the maximum kinetic energy.

The general character of the interaction of a power laser pulse with the aerogel system is similar to that for deuterium clusters. In what follows, we thus use the estimates that were obtained for the explosion of deuterium clusters under the action of a superintense ultrashort laser pulse [6].

The mass of the individual SiO_2 ball is equal to

$$m = \rho_0(4\pi a^3/3) = 7.0\times 10^{-20}\ \text{g},$$

and the SiO_2 molecule mass is $m_0 = 1.0 \times 10^{-22}$ g. The number density of SiO_2 molecules in the ball is

$$N_0 = \frac{\rho_0}{m_0} = 2.1\times 10^{22}\ \text{cm}^{-3}.$$

Hence, one ball contains approximately $n = m/m_0 =$ 700 SiO_2 molecules, and $n_0 = 2150$ deuterium molecules are located on its surface (see the previous section). We assume that the laser peak intensity is $I_{\max} = 10^{18}$ W/cm^2, the laser wavelength is $\lambda = 800$ nm, and the pulse width is $\tau = 35$ fs (defined as the full width at half maximum; see [9]). The laser field strength $F(t)$ is then defined as

$$f(t) = F_{\max}\exp[-(t/30)^2]$$

(with t measured in fs), and the peak laser field strength is $F_{\max} = 5.25$ a.u. (the linear polarization of the laser field is considered; the atomic units correspond to $e = m_e = \hbar = 1$). The laser focal volume is equal to $V = 6.3\times 10^{-5}$ cm^3 (see the Introduction). The mass of the aerogel in this volume is

$$M = \rho V = 7.6\times 10^{-7}\ \text{g}.$$

Hence, the number of the aerogel solid balls in the plasma filament can be estimated as

$$n_b = M/m = 1.1\times 10^{13}.$$

4. INNER AND OUTER IONIZATION OF THE AEROGEL

We now consider the processes that occur during the interaction of laser light with the aerogel. A multiple inner ionization of silicon, oxygen, and deuterium atoms first occurs in this laser field (of course, the molecular bonds are destroyed very quickly). Taking the known values of the ionization potentials E_Z of multicharged Si and O atomic ions into account (where Z is the charge multiplicity of the respective atomic ion), we find that the above-barrier inner multiple field ionization occurs in the laser field when the Bethe condition is satisfied [10]:

$$F(t) > \frac{E_Z^2}{4Z}. \qquad (4)$$

Here, $F(t)$ is the amplitude of the laser field strength at the time moment t.

Using the known values of the ionization potentials of atomic ions [11], we find from Eq. (4) that, at the leading edge of the laser pulse, the charge multiplicity is $Z = 6$ for oxygen atomic ions. Each oxygen atomic ion preserves only two electrons of the K shell because the K-shell ionization potentials are very high, 739 and 871 eV, respectively. Further, $Z = 9$ for silicon atomic ions, which means that each Si atomic ion preserves only two K-shell electrons and three L-shell electrons. The ionization potential of the Si atomic ion with $Z = 9$ is equal to 401 eV, and this quantity does not satisfy condition (4) even at $F(t) = F_{\max}$.

We thus conclude that $9 + 2 \times 6 = 21$ electrons are removed at the leading edge of the laser pulse from each SiO_2 molecule inside the individual ball during the inner atomic ionization process. We neglect the quantum-mechanical tunneling atomic ionization at the start of the laser pulse because its probability is too small compared to the probability of the classical above-barrier ionization [12]. Therefore, each individual ball contains

$$n_e = 21\times n + 2n_0 = 14\,700 + 4300 = 19\,000$$

free electrons after irradiation by the laser pulse (here, we also accounted for two electrons ejected from each deuterium molecule on the surface of the individual aerogel ball).

Simultaneously, the outer above-barrier ionization proceeds at the leading edge of the laser pulse, which means that the electrons leave the individual aerogel ball. We can calculate the number Q of electrons removed from the ball by the laser field by applying the Bethe model again. Instead of Eq. (4), we have a similar condition based on the Coulomb binding potential E_Q for electrons in the individual ball with the positive charge Q; for $Q \gg 1$, the potential is given by $E_Q = Q/a$, and therefore (see the review paper [13] for details),

$$F(t) > \frac{Q}{4a^2}. \qquad (5)$$

The thermal mechanism of the outer ionization for a femtosecond time range and for moderate dimensions of the individual aerogel balls gives an additional contribution (see estimates in [14] and the discussion below). The amplitude of electron oscillations in the laser field is

$$a_0 = F_{\max}/\omega^2 = 860\ \text{Å} \gg a,$$

and therefore, the electrons do not return to the individual ball after such oscillations with high probability.

We note that Eq. (5) for the outer ionization is not so obvious as Eq. (4) for the inner ionization. The relation $E_Q = Q/a$ becomes invalid when a few electrons remain in the center of the ball. Assuming that positive atomic ions are homogeneously distributed in the ball, we find that $E_Q = 3Q/2a$. Finally, if the stronger inequality $F(t) > Q/a^2$ is satisfied (the laser force is stronger than the Coulomb force), then electrons are definitely removed from the individual aerogel ball.

We now evaluate the number of electrons Q that are released from an aerogel ball under the action of the laser field. From Eq. (5), we obtain $Q = 30\,000$ released electrons for $F(t) = F_{\max}$; this means that all the $n_e = 19\,000$ free electrons are liberated in the outer ionization from each aerogel ball at the leading edge of the laser pulse. These free electrons are then uniformly distributed in the plasma filament. The total number of free electrons in the plasma filament is equal to

$$n_t = n_e n_b = 2.1 \times 10^{17}.$$

The concentration of these free electrons in the plasma filament is estimated as

$$N_e = n_t/V = 3.3 \times 10^{21}\ \text{cm}^{-3}.$$

Free electrons inside the individual ball are heated during the laser pulse. First, each electron acquires a large kinetic energy equal to the average oscillation energy $F^2/4\omega^2$. This quiver energy is equal to 58 keV for the peak intensity $I = I_{\max}$. However, after the end of the laser pulse, the electron loses all its quiver energy, because the kinetic energy of the electron adiabatically follows the envelope of the laser pulse. According to [12] in the case of the above-barrier ionization, the final kinetic energy of electrons is on the order of 10–20 eV only. Indeed, for the above-barrier ionization (and also for tunneling ionization), the real energy spectrum of electrons is determined by the simple exponential law [12]

$$w(E) \propto \exp\left(-\frac{2E_e\gamma^3}{3\omega}\right), \quad \gamma = \omega\sqrt{\frac{8a}{F(t)}}. \tag{6}$$

The quantity

$$\gamma = \frac{\omega\sqrt{2E_Q}}{F(t)}$$

is the so-called Keldysh parameter (see [12] for details) for the ejection of electrons from the individual ball by the laser field. For the maximum value $F = 5.25$ a.u., we obtain from Eq. (6) that $\gamma = 0.4$. Because $\gamma < 1$, we find that the above-barrier outer ionization by a quasistatic laser field is indeed realized. It then follows from Eq. (6) that the typical electron kinetic energy is

$$E_e \approx \frac{3\omega}{2\gamma^3} \approx 35\ \text{eV}.$$

Therefore, free electrons remain quite cold immediately after the outer ionization. But then free electrons can again significantly increase their kinetic energy in the ball expansion process. The Coulomb potential energy of electrons is transformed into their kinetic energy [6]. As a result, the electron temperature increases up to the final value $T_e \sim 1$ keV. In addition to this, the induced inverse bremsstrahlung [14, 15] in the plasma filament can contribute to the increase in the electron kinetic energy when electrons are scattered by the charged individual balls before the laser pulse is depleted. The amount of this contribution is unclear because of quick expansion of these balls.

We note that the Coulomb expansion of an individual ball proceeds slowly in comparison with the electron release; therefore, it is insignificant during the process of the outer field ionization. Indeed, the time t for doubling the ball radius can be estimated from a simple energy balance for the silicon atomic ion at the ball surface (see [15] for details),

$$\frac{(Q/2)}{a} - \frac{(Q/2)}{2a} = \frac{1}{2}M_0 v^2 = \frac{1}{2}M_0\left(\frac{2a}{t}\right)^2,$$

where M_0 is the mass of a single silicon atomic ion and v is its velocity at the time instance t. In this estimate, we used the typical time instance when a half of the free electrons ($Q/2$) are removed from an individual ball. It follows from this relation that the ball radius a is doubled during the time

$$t = \sqrt{\frac{8M_0a^3}{Q}} = 20\ \text{fs}. \tag{7}$$

Hence, we can neglect the expansion of the individual ball before the total ejection of all the $Q = 30\,000$ electrons from an individual ball at the leading edge of the ultrashort laser pulse. Estimates for the oxygen atomic ion lead to a similar conclusion.

The laser energy is absorbed by free electrons in the plasma filament. In accordance with Eq. (6), each electron acquires the energy $E_e \approx 35$ eV from the laser field (another part of the electron energy ~1 keV is acquired from the Coulomb potential energy of the electrons in the ball). If the laser energy in the pulse is, e.g., $E = 1$ J $= 6.2 \times 10^{18}$ eV, it follows that $n_l = E/E_e = 1.8 \times 10^{17}$ electrons absorb the entire energy of the laser pulse. But we have found above that the total number of free electrons in the plasma filament is $n_t = 2.1 \times 10^{17}$, which is approximately equal to n_l. Thus, we conclude that, for the high-intensity laser field, a low-frequency electromagnetic wave penetrates into the entire plasma filament.

The analysis leads to the following conclusions. First, a standard concept that the electromagnetic wave with the frequency $\omega < \omega_p$ cannot propagate in a plasma (see, e.g., [16]) is not valid at a high intensity of the electromagnetic field. Second, the propagation of a strong electromagnetic wave causes a redistribution of plasma charges, and the interaction of the electromagnetic wave with a forming nonuniform plasma may be important for plasma heating. We note that we neglected this interaction, in the above analysis, and formation of the plasma under the action of the electromagnetic wave results only in the separation of electrons and ions of the aerogel skeleton by the wave. A partial absorption of the electromagnetic wave by a forming plasma as a result of collective excitations in this plasma may contribute an additional heating of the plasma.

Deep penetration of the ultrahigh-intensity laser pulse into a dense plasma is also confirmed by theoretical results in [17].

5. FUSION PROCESSES

We finally estimate the number of fusion neutrons produced in the plasma after the end of the laser pulse. During the preliminary diffusion of the deuterium gas through the aerogel, the deuterium molecules penetrate inside the pores of the aerogel; then they adhere to the surface of individual balls. Of course, large pores in the fractal structure of the aerogel allow several layers of deuterium molecules to cover each of the SiO_2 balls. We estimated above that $n_0 \sim 2150$ deuterium molecules adhere to each SiO_2 ball surface. This estimate corresponds to three layers of deuterium molecules at the ball surface with the radius $a = 20$ Å. Deuterons are attracted to the free ends of the oxygen radicals. It must be noted that it is probably better to use heavy water instead of deuterium gas, because the polar D_2O molecule is better attached to the skeleton surface than the nonpolar D_2 molecule.

We now derive the number density of deuterons in the laser focal volume V:

$$N_d = \frac{2n_0n_b}{V} = 7.5\times10^{20}\ \text{cm}^{-3}.$$

It is ten times greater than in experiments [1–4] with deuterium clusters (see the Introduction).

At the Coulomb explosion of the individual aerogel balls, each deuteron acquires the maximum kinetic energy

$$\frac{Q}{a+l_d} = 13.5\ \text{keV}.$$

We note that the hydrodynamic expansion of the ball with the ion sound velocity is negligibly small compared to the Coulomb expansion, in contrast to the expansion of large Xe clusters [15].

We now consider the Coulomb explosion in the model of cylindrical fibers. The number density of SiO_2 molecules in the solid fiber matter (see above) is $N_0 = 2.1 \times 10^{22}$ cm^{-3}. After the inner ionization, the unit of the fiber length contains the electric charge

$$Z = 21N_0\pi a_f^2 = 2.3\times10^{10}\ e/\text{cm}.$$

According to the Gauss electrostatic theorem, the electric field strength $F(r)$ at the distance r from the fiber axis is found from the relation

$$2\pi rF(r) = 4\pi Z.$$

On the surface of the fiber, the field strength is

$$F_0 = F(a_f) = \frac{2Z}{a_f} = 10.0\ \text{a.e.}$$

Hence, the difference of electric potentials between the surface of the given fiber and the neighboring fiber is

$$\Delta\varphi = F_0a_f\ln(\delta/a_f) = 16\ \text{keV}.$$

This quantity is equal to the maximum kinetic energy of the deuteron. It is seen that the models of fibers and of balls give similar values of the deuteron kinetic energy.

The cross section of the fusion reaction

$$d + d \longrightarrow {}^3\text{He} + n$$

is $\sigma = 10^{-30}$ cm^2 for the deuteron kinetic energy $E_d = 10$ keV [18] (the reaction $d + d \longrightarrow t + p$ has the same cross section, but we are not interested in this reaction channel here). The rate for this nuclear reaction is $w = \sigma N_d v$, where v is the deuteron velocity. The time for the nuclear fusion is determined by the flight time T of the deuteron from the axis of the laser focal volume to its radial boundary, i.e., $T = r/v$. Hence, each deuteron produces

$$wT = \sigma N_d r = 7.5\times10^{-12} \ll 1\ \text{neutron}$$

at the collisions with other deuterons. The total yield of neutrons n_n per laser pulse can be obtained by multiplying this quantity with the number $n_d = n_0n_b$ of deuteron pairs in the laser focal volume:

$$n_n = wTn_0n_b \approx 10^5.$$

These estimates refer to the ball model of aerogel.

We now make a similar estimate in the fiber model of aerogel. If the quantity $S = 715$ m^2/g is the specific area of the aerogel (Section 2) and $M = \rho V$ is the mass of the aerogel ($\rho = 0.012$ g/cm^3 is the mass density of the aerogel and V is the volume of the plasma filament), then the quantity MS is the total area of the aerogel in the plasma filament. This area is covered by three layers of deuterium molecules. The thickness of this layer is $l_d = 10.5$ Å (Section 2). Hence, the quantity MSl_d is the volume of the deuterium layer, and the quantity ρ_dMSl_d is its mass ($\rho_d = 0.17$ g/cm^3 is the mass density of the solid deuterium matter). The number of deuterium molecules in the plasma filament is

$$n_d = \frac{\rho_d MSl_d}{M_d}$$

($M_d = 6.7 \times 10^{-24}$ g is the mass of one deuterium molecule). Thus, the number density of deuterons in the plasma filament is

$$N_d = \frac{n_d}{V} = \frac{\rho_d \rho Sl_d}{M_d} = 4.6 \times 10^{20}\ \mathrm{cm}^{-3}.$$

It is seen that this estimate is nearly the same as the corresponding estimate in the ball model of aerogel (see above). Hence, the estimate of n_n is also the same as above.

It should be noted that the cross section σ increases by several decimal orders compared to the case of the experiments of Ditmire *et al.* [1–4] with deuterium clusters, and the number of deuterons in the plasma filament $2n_0n_b = 4.7 \times 10^{16}$ is larger than the number of deuterons $n_d \approx (2–3) \times 10^{15}$ (see the Introduction) in the plasma filament of the same dimensions used in the experiments [1–4]. Nevertheless, the neutron yield increases by only ten times. In our opinion, the reason is that, in the experiments of Ditmire *et al.*, only deuterons with large radii $R \approx 80$ Å take part in the nuclear fusion, while, in our approach, the radius of the individual aerogel ball is only $a = 20$ Å.

We can conclude that the yield of neutrons is greater by approximately ten times compared to the yield of neutrons at the irradiation of deuterium clusters by a superintense ultrashort laser pulse observed in the experiments [1–4] (see also recent theoretical calculations for deuterated clusters in [19]). Thus, an aerogel saturated by deuterium can be used for the production of powerful sources of ultrashort pulses (about T = 200–500 ps) of monochromatic neutrons (2.45 MeV).

6. CONCLUSION

Developing the method by Ditmire *et al.* [3, 4] for the production of neutrons under the action of a superintense ultrashort laser pulse, we propose to use the aerogel skeleton with absorbed deuterium instead of the deuterium cluster in the experiments by Ditmire *et al.* This allows increasing the neutron yield per laser pulse by one order of magnitude because of a higher kinetic energy of deuterons liberated at the Coulomb explosion of the charged aerogel skeleton. It should be noted that an interesting theoretical approach was recently proposed in [20]: it is suggested to use heteronuclear clusters containing deuterium, e.g., clusters from D_2O molecules.

It is possible that the tungsten aerogel [21] has an advantage compared to the SiO_2 aerogel considered in our paper, due to a high charge of the tungsten nucleus.

We note in conclusion that the Lawson criterion is ~10^{-5} for deuterium clusters and ~10^{-4}–10^{-3} for the aerogel. The proposed method can be discussed as a version for the basis of the laser thermonuclear reactor.

ACKNOWLEDGMENTS

The authors are grateful to V.E. Fortov, T. Ditmire, P. Parks, and I.Yu. Skobelev for fruitful discussions. This work was supported by the Russian Foundation for Basic Research (project nos. 01-02-16056 and 01-02-06514) and by the Humboldt Foundation (Germany).

REFERENCES

1. T. Ditmire, J. Zweiback, V. P. Yanovsky, *et al.*, Nature **398**, 489 (1999).
2. J. Zweiback, R. A. Smith, V. P. Yanovsky, *et al.*, in *Proceedings of the 8th International Conference on Multiphoton Processes, ICOMP VIII*, Ed. by L. F. DiMauro, R. R. Freeman, and K. C. Kulander (American Inst. of Physics Press, New York, 2000), p. 534.
3. J. Zweiback, R. A. Smith, T. E. Cowan, *et al.*, Phys. Rev. Lett. **84**, 2634 (2000).
4. J. Zweiback, T. E. Cowan, R. A. Smith, *et al.*, Phys. Rev. Lett. **85**, 3640 (2000).
5. H. Schwoerer, H. Gibbon, S. Düsterer, *et al.*, Phys. Rev. Lett. **86**, 2317 (2001).
6. V. P. Krainov and M. B. Smirnov, Zh. Eksp. Teor. Fiz. **119**, 719 (2001) [JETP **92**, 626 (2001)].
7. B. M. Smirnov, Phys. Rep. **188**, 1 (1990); *Physics of Fractal Clusters* (Nauka, Moscow, 1993).
8. K. Christmann, *Introduction to Surface Physical Chemistry* (Springer-Verlag, New York, 1991).
9. T. Ditmire, Phys. Rev. A **57**, R4094 (1998).
10. H. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two Electron Atoms* (Rosetta, New York, 1977).
11. *Handbook of Chemistry and Physics*, Ed. by D. R. Lide (CRC, Boca Raton, 1998–1999, 79th ed.).
12. N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms* (Springer-Verlag, Berlin, 2000).
13. W. A. de Heer, Rev. Mod. Phys. **65**, 611 (1993).
14. V. P. Krainov and M. B. Smirnov, Usp. Fiz. Nauk **170**, 969 (2000).
15. T. Ditmire, T. Donnelly, A. M. Rubenchik, *et al.*, Phys. Rev. A **53**, 3379 (1996).
16. B. M. Smirnov, *Physics of Ionized Gases* (Wiley, New York, 2001).
17. E. Lefebre and G. Bonnaud, Phys. Rev. Lett. **74**, 2002 (1995).
18. E. Teller, *Fusion* (Academic, New York, 1981).
19. P. B. Parks, T. E. Cowan, R. B. Stephens, and E. M. Campbell, Phys. Rev. A **63**, 063203 (2001).
20. I. Last and J. Jortner, Phys. Rev. Lett. **87**, 033401 (2001).
21. C. D. Decker and R. A. London, Phys. Rev. A **57**, 1395 (1998).

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 491–498.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 562–569.
Original Russian Text Copyright © 2001 by Agre.

ATOMS, SPECTRA, RADIATION

The Scattering of Partially Polarized Light by Oriented Atoms

M. Ya. Agre*
International Solomon University, Kiev, 01135 Ukraine
**e-mail: markag@aport2000.ru*

Received April 24, 2001

Abstract—A compact expression is derived for the cross section of scattering of arbitrarily polarized light by oriented atomic systems, in which the dependence on the geometric parameters and on the Stokes parameters preassigning the state of partial polarization of incident radiation is explicitly separated. It is found that the cross section of any photoprocess accompanied by photon absorption (stimulated emission) contains the sum of the products of the circular and linear dichroisms of the process by the respective Stokes parameters. The effect of the atomic orientation and of the dissipation of light energy on the polarization singularities and angular distribution of scattered light is investigated. In particular, it is demonstrated that, in the case of an open dissipation channel, the angular distribution remains dependent on the atomic orientation even in the case of zero degree of circular polarization of scattered radiation. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Under regular conditions of free orientation, the magnetic sublevels of an atom are populated uniformly. Such a nonpolarized atom is a symmetric system. The nonuniformity of populations of states with different values of momentum projection (atomic polarization) disturbs the symmetry, which affects considerably the processes of interaction between a polarized atom and electromagnetic radiation. An atom polarizes during light absorption, collisions, and other processes. A special method of optical pumping has been developed for the polarization of atoms (see review papers [1, 2]), and some elementary photoprocesses on polarized atoms (photoeffect, light emission and absorption) have been well studied (see [3] and the references given there, as well as [4, 5]).

It will be recalled that, in the general case, a polarized atom is in a mixed quantum-mechanical state. The decomposition of its density matrix to irreducible components referred to as multipoles of state (statistical tensors) [3] enables one to identify different types of polarization. A multipole of state ρ_{KQ}, with $K = 0, 1, \ldots, 2j_1$ (where j_1 is the quantum number of total momentum of atom in polarized state) and $Q = -K, -K + 1, \ldots, K$, is an irreducible tensor of the Kth order. In this case,

$$\rho_{00} = (2j_1 + 1)^{-1/2},$$

and, in the absence of polarization, all of the remaining multipoles of state go to zero. An atom is referred to as oriented if the tensor ρ_{1Q} proportional to the spherical components of the mean momentum $\mathbf{j}$ of a polarized atom is other than zero. At $\rho_{2Q} \neq 0$, an atom is referred to as aligned.

The general theory of light scattering by a polarized atom was developed in [6]. It has been demonstrated, in particular, that the differential cross section of dipole scattering may contain multipoles of state up to the fourth order inclusive, and the total cross section, up to the second order. In the expression for scattering cross section, the dependence on geometric parameters was separated; however, the cross section was written in a fairly cumbersome form containing irreducible tensors made up of polarization vectors of incident and scattered photons (see formula (2) below). With this representation of the cross section, it proved difficult to analyze numerous effects which may be observed experimentally. Subsequent studies dealt with the so-called dissipatively induced effects during light scattering by oriented [7] and aligned [8] atoms, caused by channels of dissipation of light energy opened in the process of scattering. The respective cross sections for the case of complete polarization of incident light could be represented in a simpler form containing both scalar and vector products of vectors.

I have derived a compact expression for the cross section of scattering of partially polarized light by oriented atoms. In the cross section, the dependence on the Stokes parameters preassigning the polarization of light and on the geometric parameters is separated. The effect of the atomic orientation and of dissipatively induced effects on the polarization singularities and on the angular distribution of scattered light is investigated. In particular, it is demonstrated that the orientation brings about the dependence of the angular distribution of scattered radiation on the degree of circular polarization of incident light, and that the orientation-related dependence of angular distribution on the degree of linear polarization of light is dissipatively induced. For definiteness, in what follows, we will refer to atoms, but the results of this study relate to the scattering of light by oriented molecules as well, because the entire information about the internal structure of the target is contained in the reduced matrix elements of the scattering tensor.

2. SCATTERING OF COMPLETELY POLARIZED LIGHT

In [7], an expression was derived for the cross section of light scattering by an oriented atom in the case when all quantum numbers of the total momentum of atom are in the initial and final states $j_1 = j_2 = 1/2$ (in the case of $j_1 = 1/2$ and $j_2 = 3/2$, the respective expression is given only for the linear polarization of incident and scattered light). Therefore, we will treat the general case of arbitrary values of j_1 and j_2 permitted by the selection rules.

The formula for light scattering by a polarized atom given in [6] assumes that the atom is polarized axisymmetrically relative to the symmetry axis preassigned by the unit vector $\mathbf{n}$. In the coordinate system with the z axis directed along the vector $\mathbf{n}$, only zero components $\rho_{K0} = \rho_K^n$ of all multipoles of state are other than zero, and, accordingly, the density matrix of the axisymmetrically polarized atom is diagonal over the projection of momentum m [3]. The state of the polarized atom in this case is an incoherent mixture of states with different values of the projections of momentum m onto the direction $\mathbf{n}$. The polarization of this type apparently arises when the external polarizing effect is axisymmetric. The mean momentum of atom $\mathbf{j}$ in the case of axisymmetric polarization is collinear with the vector $\mathbf{n}$ and related to the orientation ρ_1^n by the relation

$$\rho_1^n \mathbf{n} = \sqrt{3}[(2j_1+1)(j_1+1)j_1]^{-1/2}\mathbf{j}. \qquad (1)$$

Note right away that all formulas derived for the case of axisymmetric polarization, which describe orientation effects, retain their form in the general case of asymmetric atomic polarization. In accordance with Eq. (1), the vector $\mathbf{n}$ is a unit vector collinear with the mean momentum $\mathbf{j}$ of oriented atom. In the coordinate system with the z axis directed along $\mathbf{n}$, the components of the multipoles of state of the first order are $\rho_{1,\pm1} = 0$. However, nonzero components of higher multipoles of state are, generally speaking, other than zero, so that the atomic density matrix will not be diagonal over m.[1]

We will use the expression for the differential cross section of light scattering by a polarized atom, obtained in [6],

$$\frac{d\sigma}{d\Omega'} = (4\pi)^{1/2}\omega\omega'^3\alpha^4 \sum_{K,k,k'} \rho_K^n (-1)^{j_1+j_2+k+K}$$

$$\times \begin{Bmatrix} k & k' & K \\ j_1 & j_1 & j_2 \end{Bmatrix} (2K+1)^{-1/2} T_k T_{k'}^* \qquad (2)$$

$$\times \sum_Q Y_{KQ}^*(\mathbf{n}) \{\{\mathbf{e}'^* \otimes \mathbf{e}\}_k \otimes \{\mathbf{e}' \otimes \mathbf{e}^*\}_{k'}\}_{KQ}.$$

Here, ω (ω') is the frequency, $\mathbf{e}$ ($\mathbf{e}'$) is the unit vector of polarization of incident (scattered) light, α is the fine structure constant (use is made of the atomic system of units), and

$$T_k = \langle \nu_2 j_2 \| t_k \| \nu_1 j_1 \rangle, \quad k = 0, 1, 2,$$

are reduced matrix elements of irreducible parts of the scattering tensor, which arise upon separation of the dependence from the magnetic quantum numbers in the irreducible components of the scattering tensor t_{kq} using the Wigner–Eckart theorem. The set of atomic quantum numbers in the initial and final states (except for the momentum and its projection) is denoted by $\nu_{1,2}$. Formula (2) includes also the spherical function $Y_{KQ}(\mathbf{n})$ and irreducible tensors made up of polarization tensors. The irreducible tensor of the Kth order made up of the irreducible tensors $A_{k_1q_1}$ and $B_{k_2q_2}$ of the orders k_1 and k_2 is defined as

$$\{A_{k_1} \otimes B_{k_2}\}_{KQ} = \sum_{q_1,q_2} C^{KQ}_{k_1q_1k_2q_2} A_{k_1q_1} B_{k_2q_2},$$

where $C^{KQ}_{k_1q_1k_2q_2}$ is the Klebsch–Gordan coefficient. For the vector $\mathbf{a}$, we have $a_{1q} = a_q$, where a_q denotes its spherical components,

$$a_0 = a_z, \quad a_{\pm1} = \mp\frac{1}{\sqrt{2}}(a_x \pm i a_y).$$

The irreducible components of the scattering tensor have the form

$$t_{kq} = \sum_{q_1,q_2} C^{kq}_{k_1q_1k_2q_2} \qquad (3)$$

$$\times (d_{q_1}\hat{G}_{E_1+\omega} d_{q_2} + d_{q_2}\hat{G}_{E_1-\omega'} d_{q_1}),$$

where d_q denotes the spherical components of the atomic dipole moment,

$$\hat{G}_E = \sum_n \frac{|n\rangle\langle n|}{E_n - E - i0} \qquad (4)$$

is the resolvent of the atomic Hamiltonian, and E_1 is the energy of the atom in the initial state.

---

[1] The exception is provided by the case of $j_1 = 1/2$, when the system is characterized by only one multipole of state of nonzero order ρ_{1Q}, and the polarization reduces to orientation. The direction of the z axis over the vector $\mathbf{n}$ diagonalizes automatically the density matrix of the system.

Assume that the atom is oriented only, i.e., $\rho^n_{K>1} = 0$. We will represent the light scattering cross section given by Eq. (2) in a compact form containing scalar and vector products of the polarization vectors **e** and **e**', the unit vectors **k** and **k**' preassigning the direction of propagation of incident and scattered photons, and the unit vector **n** collinear with the mean momentum **j** of the oriented atom. We will write the differential cross section of light scattering by an oriented atom in the form of the sum of two terms,

$$\frac{d\sigma}{d\Omega'} = \frac{d\sigma^{(unp)}}{d\Omega'} + \frac{d\sigma^{(or)}}{d\Omega'}. \tag{5}$$

Here,

$$\frac{d\sigma^{(unp)}}{d\Omega'} = \frac{\omega\omega'^3\alpha^4}{2j_1+1}$$
$$\times\left\{\frac{1}{2}|T_0|^2|\mathbf{e}'^*\cdot\mathbf{e}|^2 + \frac{1}{6}|T_1|^2(1-|\mathbf{e}'\cdot\mathbf{e}|^2) \right. \tag{6}$$
$$\left. + \frac{1}{10}|T_2|^2\left(1+|\mathbf{e}'\cdot\mathbf{e}|^2 - \frac{2}{3}|\mathbf{e}'^*\cdot\mathbf{e}|^2\right)\right\}$$

is the cross section of light scattering by a nonpolarized atom (see [6] and [9, Para. 60]), consisting of scalar, antisymmetric, and symmetric parts. The presence of each one of these parts, which are proportional to the squares of modules of reduced matrix elements T_k of the respective irreducible parts of the scattering tensor, $k = 0, 1, 2$, is possible only if the condition of triangle $\Delta(j_1, j_2, k)$ is valid. The second term in Eq. (5) proportional to ρ^n_1 defines the addition to the cross section caused by the atomic orientation.

The spherical function $Y_{1Q}(\mathbf{n})$ is proportional to the spherical component of the unit vector,

$$Y_{1Q}(\mathbf{n}) = \sqrt{\frac{3}{4\pi}}n_Q.$$

Therefore, at $K = 1$, the sum over Q in expression (2) is a scalar product of the vector **n** and the vector (irreducible tensor of the first order) made up of polarization vectors,

$$\sum_Q(-1)^Q Y_{1-Q}(\mathbf{n})\{\{\mathbf{e}'^*\otimes\mathbf{e}\}_k\otimes\{\mathbf{e}'\otimes\mathbf{e}^*\}_{k'}\}_{1Q}$$
$$= \sqrt{\frac{3}{4\pi}}(\mathbf{n}\cdot\{\{\mathbf{e}'^*\otimes\mathbf{e}\}_k\otimes\{\mathbf{e}'\otimes\mathbf{e}^*\}_{k'}\}_1). \tag{7}$$

For all possible sets of numbers k and k' ($k = 0$, $k' = 1$; $k = 1$, $k' = 0$; $k = k' = 1$; $k = 1$, $k' = 2$; $k = 2$, $k' = 1$; $k = k' = 2$), the scalars given by Eq. (7) may be expressed in terms of the scalar and vector products of the vectors entering these scalars, using formulas from the reference book [10]. After separating out all of the vector combinations, the expression for the second orientation term in Eq. (5) is written as

$$\frac{d\sigma^{(or)}}{d\Omega'} = \omega\omega'^3\alpha^4\rho^n_1\{a_+\eta_2\mathbf{n}\cdot\mathbf{k} + a_-\eta'_2\mathbf{n}\cdot\mathbf{k}'$$
$$- b_+\eta_2\mathrm{Re}[(\mathbf{n}\cdot\mathbf{e}')(\mathbf{k}\cdot\mathbf{e}'^*)] - b_-\eta'_2\mathrm{Re}[(\mathbf{n}\cdot\mathbf{e})(\mathbf{k}'\cdot\mathbf{e}^*)] \tag{8}$$
$$+ c\mathbf{n}\mathrm{Re}[[\mathbf{e}'^*\times\mathbf{e}](\mathbf{e}'\cdot\mathbf{e}^*)] + d\eta_2\eta'_2\mathbf{n}\cdot[\mathbf{k}'\times\mathbf{k}]\}.$$

Here,

$$\eta_2 = i\mathbf{k}\cdot[\mathbf{e}\times\mathbf{e}^*] \tag{9}$$

has the meaning of the degree of circular polarization of incident light. For right-hand (left-hand) polarization, $\eta_2 = \pm 1$. The parameter η'_2 is defined by the expression, similar to Eq. (9),

$$\eta'_2 = i\mathbf{k}'\cdot[\mathbf{e}'\times\mathbf{e}'^*].$$

To avoid misunderstanding, note that η'_2 is not the degree of circular polarization of the scattered photon. Scattered light, generally speaking, is in a mixed state as regards polarization (is partially polarized), and its Stokes parameters can readily be calculated using formulas (5), (6), and (8) for the scattering cross section. The coefficients $a_\pm$, $b_\pm$, c, and d are expressed in terms of 6j symbols and reduced matrix elements of the scattering tensor T_k,

$$a_\pm = \frac{1}{\sqrt{6}}R_{10} - \frac{1}{\sqrt{30}}R_{12} \pm \frac{1}{\sqrt{10}}R_{22},$$
$$b_\pm = \frac{1}{\sqrt{6}}R_{10} \pm \frac{1}{2\sqrt{2}}R_{11} + \sqrt{\frac{2}{15}}R_{12} \pm \frac{1}{2\sqrt{10}}R_{22},$$
$$c = \sqrt{\frac{2}{3}}I_{10} - \sqrt{\frac{2}{15}}I_{12}, \quad d = \sqrt{\frac{3}{10}}I_{12}, \tag{10}$$
$$R_{kk'} = (-1)^{j_1+j_2}\begin{Bmatrix} j_2 & j_1 & k' \\ 1 & k & j_1 \end{Bmatrix}\mathrm{Re}(T_kT^*_{k'}),$$
$$I_{kk'} = (-1)^{j_1+j_2}\begin{Bmatrix} j_2 & j_1 & k' \\ 1 & k & j_1 \end{Bmatrix}\mathrm{Im}(T_kT^*_{k'}).$$

The differential cross section given by Eq. (5), proportional to the probability of detection of a scattered photon in a state with certain polarization **e**', provides the most comprehensive data on scattering and enables one to predict the result of transmission of light scattered in the given direction through a polarization filter. If the polarization of scattered light is not recorded during the experiment, one must only know the angular distribution of scattered radiation. In order to find the angular distribution, the differential scattering cross section given by Eq. (5) is to be summed over two independent polarizations of the scattered photon. In all vector combinations containing the vector **e**' in formulas (6) and (8), the summation is readily performed using identity [9, Para. 45],

$$\sum_{\lambda}(\mathbf{a}\cdot\mathbf{e}'_{\lambda})(\mathbf{b}\cdot\mathbf{e}'^{*}_{\lambda}) = [\mathbf{k}'\times\mathbf{a}]\cdot[\mathbf{k}'\times\mathbf{b}], \qquad (11)$$

and the parameter η'_2 after summation vanishes. As a result, we derive the following expression for the angular distribution of scattered light:

$$\frac{d\sigma_s}{d\Omega'} = \frac{d\sigma_s^{(unp)}}{d\Omega'} + \frac{d\sigma_s^{(or)}}{d\Omega'},$$

where

$$\frac{d\sigma_s^{(unp)}}{d\Omega'} = \frac{\omega\omega'^3\alpha^4}{3(2j_1+1)}\Big[|T_0|^2 + \frac{1}{2}|T_1|^2 + \frac{7}{10}|T_2|^2 - \Big(|T_0|^2 - \frac{1}{2}|T_1|^2 + \frac{1}{10}|T_2|^2\Big)|\mathbf{k}'\cdot\mathbf{e}|^2\Big] \qquad (12)$$

is the angular distribution of light scattered by a nonpolarized atom, and

$$\frac{d\sigma_s^{(or)}}{d\Omega'} = \omega\omega'^3\alpha^4\rho_1^n\{(2a_+ - b_+)\eta_2\mathbf{n}\cdot\mathbf{k} + b_+\eta_2(\mathbf{k}\cdot\mathbf{k}')(\mathbf{n}\cdot\mathbf{k}') + c\mathbf{n}\cdot\mathrm{Re}([\mathbf{e}^*\times\mathbf{k}'](\mathbf{e}\cdot\mathbf{k}'))\} \qquad (13)$$

is an addition to the angular distribution caused by the atomic orientation.

We integrate the angular distribution using known identities

$$\int k'_i d\Omega' = 0, \quad \int k'_i k'_j d\Omega' = \frac{4\pi}{3}\delta_{ij}$$

over all directions of scattering to find the total cross section of light scattering by an oriented atom,

$$\sigma = 8\pi\omega\omega'^3\alpha^4\Big\{\frac{1}{9(2j_1+1)} \times \sum_{k=0}^{2}|T_k|^2 + \rho_1^n\Big(a_+ - \frac{1}{3}b_+\Big)\eta_2\mathbf{n}\cdot\mathbf{k}\Big\}. \qquad (14)$$

We will analyze the derived formulas from the general considerations of symmetry. The orientation ρ_1^n is a *T*-odd (changing sign upon time reversal) pseudoscalar, which follows directly from relation (1) relating this parameter to the mean momentum of an oriented atom. On the other hand, the scattering cross section must not vary either upon space inversion or upon time reversal. Therefore, the expressions in braces in Eqs. (8) and (13) must be *T*-odd pseudoscalars. The degree of circular polarization of incident light η_2 given by Eq. (9) and the parameter η'_2 are pseudoscalars, so that it is obvious that the above-identified expressions are scalar. At the same time, during the operation of time reversal, when the vectors **k** and **k**' defining the direction of light propagation change sign and each polarization vector is substituted by a complex-conjugate one, the pseudoscalar combinations of vectors in the last two terms of Eq. (8) and in the last term of Eq. (13) are not varied. Therefore, it is the coefficients *c* and *d* entering these terms that must be *T*-odd.

Physically, this *T*-oddness arises from the dissipation of light energy. Indeed, the coefficients *c* and *d* are proportional to the quantity $\mathrm{Im}(T_k T^*_{k'})$ (see Eq. (10)), which is strictly zero with the Hermitian scattering tensor given by Eq. (3). The anti-Hermitian part of the scattering tensor is other than zero only if a light energy dissipation channel is opened in the process of light scattering and this part is proportional to the *T*-odd dissipative parameter defining the rate of this dissipation. Manakov [11] was the first to pay attention to the part played by dissipative processes in the light scattering by a nonpolarized atom. However, in the case of a nonpolarized system, the dissipation shows up only if the nondipole effects are taken into account. On the other hand, under conditions of atomic orientation, the dissipatively induced effects show up even in the case of dipole scattering [7].

Note that these effects must be observed during above-threshold scattering (the photon energy is above the ionization threshold of an atom or molecule, or above the dissociation threshold of a molecule) and during resonance scattering.[2] In the case of above-threshold scattering, the anti-Hermitian part of the scattering tensor shows up owing to the fact that the $\hat{G}_{E_1+\omega}$ resolvent given by Eq. (4) at $E_1 + \omega > 0$ is anti-Hermitian, and it is the ionization (dissociation) width of the initial level that is the *T*-odd parameter. In the case of resonance scattering, the anti-Hermitian addition to the scattering tensor given by Eq. (3) arises when the width Γ of the resonance level (*T*-odd parameter) is introduced into the pole part of the resolvent. The resonance level must necessarily have a multiplet structure (with resonance on the singlet $\mathrm{Im}(T_k T^*_{k'}) = 0$), and the dissipatively induced effects prove to be of the order of Γ/Δ, where Δ has the order of fine splitting of resonance sublevels. Naturally, in the case of radiation broadening, when $\Gamma \sim \alpha^3$, these effects make a contribution of the order of α; however, their influence must increase with the width of the resonance level, for example, due to collisions. The importance of dissipatively induced effects in the scattering of light by an oriented atom is quite significant. In the absence of these effects, the angular distribution of scattered light (see Eq. (13))

---

[2] The dissipative effects associated with radiation corrections are negligibly small outside of the resonance region [11]; in the case of resonance scattering, the main part is taken into account by introducing the radiation width of the resonance level.

may depend on the atomic orientation only in the case of nonzero degree of circular polarization η_2 of incident light.

3. SEPARATION OF EXPLICIT DEPENDENCE ON STOKES PARAMETERS IN THE CROSS SECTION OF ARBITRARY PHOTOPROCESS

In the case of partial polarization of an electromagnetic wave, a photon is in a mixed (as regards polarization) quantum-mechanical state, which must be preassigned by a polarization density matrix [3, 9]. This is a second-order Hermitian matrix with a unit trace. Its elements may be expressed in terms of three real parameters which are usually provided by the Stokes parameters η_i, $i = 1, 2, 3$. The parameter η_2 defines the degree of circular polarization, the parameter η_3 preassigns the degree of linear polarization along the x and y axes (the z axis is directed along the wave propagation), and η_1 preassigns the degree of linear polarization along the p and q axes rotated in the xy plane through an angle of 45° relative to the x and y axes in the positive direction. In the general case, $\eta_1^2 + \eta_2^2 + \eta_3^2 \le 1$. With $\eta_1^2 + \eta_2^2 + \eta_3^2 = 1$, a photon is in the state of complete polarization (pure state) which may be preassigned by the vector $\mathbf{e}$. In the case of nonpolarized light, $\eta_1 = \eta_2 = \eta_3 = 0$.

The explicit form of the polarization density matrix depends naturally on the choice of basis. In a helicity basis, whose basis vectors are the vectors $\mathbf{e}_\pm$ of the right-hand and left-hand circular polarizations of a photon, the elements of this matrix are expressed in terms of the Stokes parameters as [3]

$$\begin{pmatrix} \rho_{++} & \rho_{+-} \\ \rho_{-+} & \rho_{--} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+\eta_2 & -\eta_3 + i\eta_1 \\ -\eta_3 - i\eta_1 & 1-\eta_2 \end{pmatrix}. \qquad (15)$$

The vectors $\mathbf{e}_\pm$ are related to the vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ of linear polarization along the x and y axes, which form the basis of plane Cartesian coordinates,

$$\mathbf{e}_\pm = \mp\frac{1}{\sqrt{2}}(\mathbf{e}_x \pm i\mathbf{e}_y). \qquad (16)$$

In this Cartesian basis, the density matrix is written as [9, 12]

$$\begin{pmatrix} \rho_{xx} & \rho_{xy} \\ \rho_{yx} & \rho_{yy} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1+\eta_3 & \eta_1 - i\eta_2 \\ \eta_1 + i\eta_2 & 1-\eta_3 \end{pmatrix}. \qquad (17)$$

We will demonstrate that the expression for the cross section of any photoprocess accompanied by the absorption or stimulated emission of a single partially polarized photon contains the sum of products of the Stokes parameters and the respective dichroisms of the process, namely, circular dichroism and two linear dichroisms.

If a photon is in the state of complete polarization with the polarization vector $\mathbf{e}$, then, in the first nonvanishing order of perturbation theory, the matrix element of transition is linear with respect to $\mathbf{e}$. Therefore, the cross section (total or differential) of the photoprocess has the following structure:

$$\sigma(\mathbf{e}) \propto |\mathbf{A}\cdot\mathbf{e}|^2, \qquad (18)$$

where the vector $\mathbf{A}$ is independent of polarization of light. The summation (averaging) over the atomic quantum numbers, for example, magnetic quantum numbers, which may be performed after transition to partial polarization of light as well, causes no changes in subsequent reasoning. The transition to partial polarization of a photon consists in averaging the cross section given by Eq. (18) over different realizations of polarization $\mathbf{e}$ and reduces to the formal substitution

$$e_j e^*_{j'} \longrightarrow \sum_{\lambda,\lambda'} \rho_{\lambda\lambda'} e_{\lambda j} e^*_{\lambda' j'}, \qquad (19)$$

where $\rho_{\lambda\lambda'}$ denotes the elements of the polarization density matrix of a photon in the basis $\{\mathbf{e}_\lambda\}$, and e_j denotes the Cartesian components of the polarization vectors.[3] As a result, the cross section of a photoprocess is represented in the well-known form [3, 9]

$$\sigma \propto \sum_{\lambda\lambda'} \mathbf{A}\cdot\mathbf{e}_\lambda \rho_{\lambda\lambda'} (\mathbf{A}\cdot\mathbf{e}_{\lambda'})^*. \qquad (20)$$

The explicit form of the polarization density matrix $\rho_{\lambda\lambda'}$ in the spiral basis given by Eq. (15) demonstrates that the Stokes parameter η_2 in the expression for cross section (20) is multiplied by the difference between the cross sections for the right-hand and left-hand circular polarizations of a photon (circular dichroism of the process), and it follows from Eq. (17) that η_3 is multiplied by linear dichroism. Therefore, the photoprocess cross section has the structure of

$$\sigma = \sigma_0 + \frac{1}{2}(\eta_1\sigma_{pq} + \eta_2\sigma_{+-} + \eta_3\sigma_{xy}), \qquad (21)$$

where

$$\sigma_0 = \frac{1}{2}\sum_\lambda \sigma(\mathbf{e}_\lambda) \qquad (22)$$

is the cross section of a photoprocess involving a nonpolarized photon,

$$\sigma_{+-} = \sigma(\mathbf{e}_+) - \sigma(\mathbf{e}_-)$$

---

[3] Note that the partial polarization of an electromagnetic wave cannot be preassigned by the polarization density matrix of photon if the cross section of a process with absorption (emission) of $N \ge 2$ photons is calculated. The number of necessary polarization parameters in these cases turns out to be more than three and depends on N [13].

is the difference between the cross sections for the right-hand and left-hand circular polarizations,

$$\sigma_{xy} = \sigma(\mathbf{e}_x) - \sigma(\mathbf{e}_y)$$

is the difference between the cross sections for the linear polarization along the *x* and *y* axes (linear *xy* dichroism of the process), and σ_{pq} is the linear *pq* dichroism of the process. We can easily reduce the cross section given by Eq. (20) to the form of (21) if we express the basis vectors $\mathbf{e}_x$ and $\mathbf{e}_y$ in terms of the basis of the *pq* coordinate system rotated through 45°,

$$\mathbf{e}_x = \frac{1}{\sqrt{2}}(\mathbf{e}_p - \mathbf{e}_q), \quad \mathbf{e}_y = \frac{1}{\sqrt{2}}(\mathbf{e}_p + \mathbf{e}_q),$$

and use Eq. (16) to write the sum given by Eq. (19) in the form

$$\sum_{\lambda, \lambda'} \rho_{\lambda\lambda'} e_{\lambda j} e_{\lambda' j'} = \frac{1}{2}\sum_{\lambda} e_{\lambda j} e^{*}_{\lambda j'}$$
$$+ \frac{1}{2}\eta_2(e_{+j}e^{*}_{+j'} - e_{-j}e^{*}_{-j'}) + \frac{1}{2}\eta_3(e_{xj}e_{xj'} - e_{yj}e_{yj'})$$
$$+ \frac{1}{2}\eta_1(e_{pj}e_{pj'} - e_{qj}e_{qj'}).$$

The symmetric part of the matrix given by Eq. (17),

$$\frac{1}{2}\begin{pmatrix} 1+\eta_3 & \eta_1 \\ \eta_1 & 1-\eta_3 \end{pmatrix},$$

is reduced to a diagonal form

$$\frac{1}{2}\begin{pmatrix} 1+\tilde{\eta}_3 & 0 \\ 0 & 1-\tilde{\eta}_3 \end{pmatrix}$$

by rotating the coordinate axes. Here, the parameter

$$\tilde{\eta}_3 = (\eta_1^2 + \eta_3^2)^{1/2},$$

referred to as the degree of linear polarization of light [12], has the meaning of the degree of linear polarization along the $\tilde{x}$ and $\tilde{y}$ axes rotated through an angle φ relative to the *x* and *y* axes,

$$\cos 2\varphi = \eta_3/\tilde{\eta}_3, \quad \sin 2\varphi = \eta_1/\tilde{\eta}_3.$$

The state of partial polarization of photon is preassigned by three parameters, namely, η_2, $\tilde{\eta}_3$, and φ. The Stokes parameter $\tilde{\eta}_1$ in rotated coordinates is apparently zero, so that the first term in brackets in Eq. (21) disappears. It was in this form that Schaphorst *et al.* [14] represented the dependence of the cross section of two-electron photoionization on the Stokes parameters.

It must be emphasized that formula (21) is universal and valid for the differential or total cross section of any photoprocess, as well as for the probability of transition induced by absorption or stimulated emission of a partially polarized photon.

4. SCATTERING OF PARTIALLY POLARIZED LIGHT

In accordance with Eq. (21), for transition from the complete to partial polarization of incident light, one must derive an expression for the scattering cross section of nonpolarized light, as well as for the circular and linear dichroisms of the process. The linear dichroism may be found by substituting the respective real vectors of linear polarization into the formulas for cross section. The transition to nonpolarized light consists in averaging expressions (6) and (8) over two orthogonal polarization vectors (see Eq. (22)). In averaging, the terms containing the parameter η_2 disappear, and the averaging of the remaining terms which contain the vector $\mathbf{e}$ is easily performed with the aid of an identity analogous to that given by Eq. (11).

When deriving a circular dichroism in the terms of Eq. (8) which contain the degree of circular polarization η_2 for completely polarized light, the parameter η_2 will be replaced by 2. As a result, these terms will enter Eq. (21) without alterations; however, η_2 will already have the meaning of the degree of circular polarization of partially polarized light. In determining the contribution made to the circular dichroism by other terms dependent on $\mathbf{e}$ in Eqs. (6) and (8), one can conveniently use the identity

$$(\mathbf{a}\cdot\mathbf{e}_{\pm})(\mathbf{b}\cdot\mathbf{e}_{\pm}) = \frac{1}{2}\{[\mathbf{k}\times\mathbf{a}]\cdot[\mathbf{k}\times\mathbf{b}] \pm i[\mathbf{b}\times\mathbf{a}]\cdot\mathbf{k}\}.$$

We will give the final expression for the cross section of scattering of partially polarized light by an oriented atom. The cross section is written in the form of Eq. (5), with the terms appearing as

$$\frac{d\sigma^{(unp)}}{d\Omega'} = \frac{\omega\omega'^3\alpha^4}{6(2j_1+1)}$$
$$\times\Big\{|T_0|^2 + \frac{1}{2}|T_1|^2 + \frac{7}{10}|T_2|^2 - T|\mathbf{k}\cdot\mathbf{e}'|^2$$
$$+ \eta_2\eta'_2\Big(|T_0|^2 + \frac{1}{2}|T_1|^2 - \frac{1}{2}|T_2|^2\Big)\mathbf{k}\cdot\mathbf{k}' \tag{23}$$
$$+ \eta_1\Delta_{pq}^{(unp)} + \eta_3\Delta_{xy}^{(unp)}\Big\},$$

where

$$T = |T_0|^2 - \frac{1}{2}|T_1|^2 + \frac{1}{10}|T_2|^2,$$

$$\Delta_{\beta\gamma}^{(unp)} = T(|\mathbf{e}'\cdot\mathbf{e}_{\beta}|^2 - |\mathbf{e}'\cdot\mathbf{e}_{\gamma}|^2);$$

$$\frac{d\sigma^{(or)}}{d\Omega'} = \frac{1}{2}\omega\omega'^3\alpha^4\rho_1^n\Big\{(2a_- - b_-)\eta'_2\mathbf{n}\cdot\mathbf{k}'$$

$$+ b_-\eta'_2(\mathbf{n}\cdot\mathbf{k})(\mathbf{k}'\cdot\mathbf{k}) + c\mathbf{n}\cdot\mathrm{Re}[[\mathbf{k}\times\mathbf{e}'^*](\mathbf{k}\cdot\mathbf{e}')]$$

$$+ 2\eta_2\Big(a_+\mathbf{n}\cdot\mathbf{k} - b_+\mathrm{Re}[(\mathbf{n}\cdot\mathbf{e}')(\mathbf{k}\cdot\mathbf{e}'^*)] \tag{24}$$

$$+\left(d + \frac{1}{4}c\right)\eta'_2\mathbf{n}\cdot[\mathbf{k}'\times\mathbf{k}]\Big) + \eta_1\Delta_{pq}^{(or)} + \eta_3\Delta_{xy}^{(or)}\Big\},$$

where

$$\Delta_{\beta\gamma}^{(or)} = c\mathbf{n}\cdot\mathrm{Re}([\mathbf{e}'\times\mathbf{e}_\beta](\mathbf{e}'^*\cdot\mathbf{e}_\beta) - [\mathbf{e}'\times\mathbf{e}_\gamma](\mathbf{e}'^*\cdot\mathbf{e}_\gamma))$$

$$- b_-\eta'_2\mathbf{n}\cdot(\mathbf{e}_\beta(\mathbf{k}'\cdot\mathbf{e}_\beta) - \mathbf{e}_\gamma(\mathbf{k}'\cdot\mathbf{e}_\gamma)).$$

The angular distribution of scattered light may be obtained by summing expressions (23) and (24) over its polarizations or by making the transition from the angular distribution obtained for complete polarization of incident light (see Eqs. (12) and (13)) to partial polarization using Eq. (21). As a result, the following expressions are derived for the angular distribution of radiation scattered by nonpolarized atoms and for the addition due to orientation:

$$\frac{d\sigma_s^{(unp)}}{d\Omega'} = \frac{\omega\omega'^3\alpha^4}{6(2j_1+1)}\Big\{|T_0|^2 + \frac{3}{2}|T_1|^2 + \frac{13}{10}|T_2|^2 \tag{25}$$

$$+ T(\mathbf{k}\cdot\mathbf{k}') + \eta_1\delta_{pq}^{(unp)} + \eta_3\delta_{xy}^{(unp)}\Big\},$$

where

$$\delta_{\beta\gamma}^{(unp)} = T[(\mathbf{k}'\cdot\mathbf{e}_\gamma)^2 - (\mathbf{k}'\cdot\mathbf{e}_\beta)^2];$$

$$\frac{d\sigma^{(or)}}{d\Omega'} = \frac{1}{2}\omega\omega'^3\alpha^4\rho_1^n\{c\mathbf{n}\cdot[\mathbf{k}'\times\mathbf{k}](\mathbf{k}'\cdot\mathbf{k})$$

$$+ 2\eta_2[(2a_+ - b_+)\mathbf{n}\cdot\mathbf{k} + b_+(\mathbf{k}\cdot\mathbf{k}')(\mathbf{n}\cdot\mathbf{k}')] \tag{26}$$

$$+ \eta_1\delta_{pq}^{(or)} + \eta_3\delta_{xy}^{(or)}\},$$

where

$$\delta_{\beta\gamma}^{(or)} = c\mathbf{n}\cdot([\mathbf{e}_\beta\times\mathbf{k}'](\mathbf{e}_\beta\cdot\mathbf{k}') - [\mathbf{e}_\gamma\times\mathbf{k}'](\mathbf{e}_\gamma\cdot\mathbf{k}')).$$

Formula (14) for the total cross section of light scattering by an oriented atom, derived for the case of complete polarization, retains its form upon transition to partial polarization of incident radiation.

5. DISCUSSION OF THE RESULTS

The formulas obtained in this study enable one to readily analyze any effects observed in the process of scattering of arbitrarily polarized light by oriented atoms. We will examine some of those effects.

First of all, note that, in the case of polarization of an atom in a state with $j_1 > 1/2$, both the orientation ρ_1^n and the subsequent multipoles of state up to the $2j_1$th order inclusive are other than zero. For example, if $j_1 = 1$, then, generally speaking, both the orientation and alignment are induced under conditions of polarization. Higher multipoles of state may not be excited because of certain selection rules for external effect that polarizes an atom. In the case of dipole absorption of light, only the orientation and alignment are induced. In the general case, the differential cross section of light scattering given by Eq. (2) receives contributions, in addition to the orientation, from three multipoles of state as well [6], so that the formulas derived above define only the orientation part of the cross section of light scattering by a polarized atom. Nevertheless, in the cases in which an atom is polarized axisymmetrically and only the orientation and alignment are induced, the orientation effects may be observed in the pure form, because the difference between the cross sections for two opposite directions of the vector **n** depends on the atomic orientation alone.

The atomic orientation affects considerably the polarization of scattered light. Formulas (23) and (24) enable one to calculate the Stokes parameters of scattered radiation and determine the state of its polarization. For example, the degree of circular polarization of scattered light (second Stokes parameter) is equal to the ratio between the difference between the cross sections for $\eta'_2 = \pm 1$ and the angular distribution. As follows from Eqs. (23) and (24), nonzero degree of circular polarization of scattered light at $\eta_2 = 0$ is induced only owing to the atomic orientation, when the system is characterized by the pseudoscalar parameter ρ_1^n. In addition, at $\eta_2 = 0$ and $\eta'_2 = 0$ (a linearly polarized photon is registered), the addition to the cross section due to orientation, as given by Eq. (24), proves to be other than zero because of the terms proportional to the T-odd coefficient c. This means that, with zero degree of circular polarization of incident light, the degree of linear polarization of scattered light depends on the atomic orientation due to dissipative effects which show up in the case of either above-threshold scattering or resonance scattering (see [7] and the discussion at the end of Section 2 of this paper).

The angular distribution of radiation scattered by a nonpolarized atom, given by Eq. (25), is independent of the degree of circular polarization η_2 of incident light, which is quite natural because η_2 is a pseudoscalar, and the vectors **k** and **k**' cannot make up a pseudoscalar combination. No less natural is the emergence of the dependence of angular distribution on η_2 under conditions of atomic orientation (see Eq. (26)). In the absence of dissipative effects (coefficient $c = 0$), the orientation part of the cross section given by Eq. (26) is proportional to the Stokes parameter η_2, so that the dependence of angular distribution of scattered radiation on atomic orientation arises only in the case of nonzero degree of circular polarization of incident light, as in the case of photon scattering by a free polarized electron [9, Para. 87]. This dependence is retained in the case of the total cross section given by Eq. (14). However, the effects of dissipation of the light energy, which show up during the scattering of light by an atom (scattering by bound electrons), result in the dependence of the angular distribution on the orientation at $\eta_2 = 0$ as well. These effects further bring about the dependence of the orientation addition to the angular distribution of scattered radiation, given by Eq. (26), on the degree of linear polarization of the initial radiation η_1 and η_3 and bring about the dependence of the angular distribution on the atomic orientation in the case of scattering of nonpolarized light [7].

The expressions for cross sections (23)–(26) and (14) derived in this study provide a complete solution to the problem on dipole scattering of arbitrarily polarized light by an oriented quantum system.

REFERENCES

1. G. V. Skrotskiĭ and T. G. Izyumova, Usp. Fiz. Nauk **73**, 423 (1961) [Sov. Phys. Usp. **4**, 177 (1961)].
2. W. Happer, Rev. Mod. Phys. **44**, 169 (1972).
3. K. Blum, *Density Matrix Theory and Its Applications* (Plenum, New York, 1981; Mir, Moscow, 1983).
4. H. Klar and H. Kleinpoppen, J. Phys. B **15**, 933 (1982).
5. N. A. Cherepkov and V. V. Kuznetsov, J. Phys. B **22**, L405 (1989).
6. M. Ya. Agre and L. P. Rapoport, Zh. Éksp. Teor. Fiz. **104**, 2975 (1993) [JETP **77**, 382 (1993)].
7. M. Ya. Agre and N. L. Manakov, J. Phys. B **29**, L7 (1996).
8. M. Ya. Agre, Zh. Éksp. Teor. Fiz. **110**, 2018 (1996) [JETP **83**, 1112 (1996)].
9. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics* (Nauka, Moscow, 1980; Pergamon, Oxford, 1982).
10. D. V. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (Nauka, Moscow, 1975; World Scientific, Singapore, 1988).
11. N. L. Manakov, Zh. Éksp. Teor. Fiz. **106**, 1286 (1994) [JETP **79**, 696 (1994)].
12. L. D. Landau and E. M. Lifshitz, in *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon, Oxford, 1975), Para. 50.
13. M. Ya. Agre, Opt. Spektrosk. **89**, 485 (2000) [Opt. Spectrosc. **89**, 445 (2000)].
14. S. J. Schaphorst, B. Krässig, O. Schwarzkopf, *et al.*, J. Electron Spectrosc. Relat. Phenom. **76**, 229 (1995).

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 499–509.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 570–582.
Original Russian Text Copyright © 2001 by Zheltikov, Alfimov, Fedotov, Ivanov, Syrchin, Tarasevitch, von der Linde.

ATOMS, SPECTRA, RADIATION

Controlled Light Localization and Nonlinear-Optical Interactions of Ultrashort Laser Pulses in Micro- and Nanostructured Fibers with a Tunable Photonic Band Gap

A. M. Zheltikov[a], M. V. Alfimov[b], A. B. Fedotov[a], A. A. Ivanov[b], M. S. Syrchin[a], A. P. Tarasevitch[c], and D. von der Linde[c]

[a]*International Laser Center, Physics Faculty, Moscow State University, Vorob'evy gory, Moscow, 119899 Russia*
e-mail: zheltikov@top.phys.msu.su
[b]*Center of Photochemistry, Russian Academy of Sciences, ul. Novatorov 7a, Moscow, 117421 Russia*
[c]*Institut für Laser- und Plasmaphysik, Universität Essen, D-45117 Essen, Germany*

Received April 27, 2001

Abstract—Physical principles behind the control of light localization and nonlinear-optical interactions in micro- and nanostructured fibers are demonstrated. Transmission measurements on the cladding of nanostructured fibers having a form of a two-dimensional periodic structure with a pitch less than 500 nm have revealed the existence of a photonic band gap tunable within the range from 930 to 1030 nm. The influence of the structure of the holey-fiber cladding on the effective area of the waveguide mode and the spectral broadening of Ti:sapphire and Cr:forsterite femtosecond laser pulses is experimentally studied. It is shown that the increase in the air-filling fraction of a holey-fiber cladding results in a considerable enhancement of spectral broadening of short laser pulses due to the increase in the light localization degree in the fiber core. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

One of the prominent tendencies of the development of natural sciences in recent years is the rapid growth of interdisciplinary areas, arising on borderlines of different fields of science and combining new ideas and achievements of physics, chemistry, biology, and other natural sciences. Whenever physics, chemistry, or biology reaches its edge, it searches for its natural extension, resorting to the tools of other natural sciences. This tendency has already resulted in many remarkable achievements and discoveries. In particular, a fusion of the achievements of modern laser physics and quantum chemistry gave birth to such a highly promising and rapidly growing area as femtochemistry [1], which opened the way to set a control over ultrafast chemical processes and to real-time monitoring of such processes [2, 3]. Progress in solid-state physics, optics of spatially nonuniform structures, and nanotechnologies based on a variety of physical and chemical processes has strongly stimulated and motivated research into the properties of photonic crystals [4–6] and resulted in the growth of applications of photonic band-gap (PBG) materials, i.e., artificially structured materials where optical parameters are periodically modulated in space with a period of a unit photonic-crystal cell on the order of the optical wavelength.

Presently, exciting prospects are open due to the advantageous integration of impressive achievements of photochemistry, including the use of ultrashort laser pulses for controlling the properties and the phase state of matter, with broad opportunities of advanced laser systems and nanostructure-fabrication technologies. We, in fact, witness a buildup of a new promising direction of research—femtonanophotonics. This interdisciplinary area of science opens the way of purposeful modifications of the properties of matter and a control of fundamental (often ultrafast) processes in matter on both physical and chemical levels by material nanostructuring. Investigations in this direction have already demonstrated the possibilities of luminescence control in photonic crystals [7, 8]; creation of nanostructures, including photonic crystals, by means of two-photon polymerization [9]; and the use of electrochemical and deep photoanodic etching [10–13] for the fabrication of nanostructures, nanocrystals, and photonic crystals, offering new avenues for controlling ultrashort pulses and nonlinear-optical interactions.

One of the promising methods to control fundamental physical and chemical processes and to solve many applied problems of ultrafast and nonlinear optics, optical technologies, high-precision optical measurements, photochemistry, and biomedicine is to use nanostructured fibers [14–24]. Fibers of this new type, also known as holey fibers (HFs) [21–24], are becoming more and more widespread in modern optics and photonics. The cladding of such fibers (Fig. 1a) has a form of a two-dimensional (often periodic) array of closely packed glass capillaries drawn at a high temperature. A fiber without a hole [14, 15] or holes corresponding to several missing capillaries [19] may serve as a core in a holey fiber. In the former case, waveguiding is achieved due to total internal reflection. In the latter case, waveguide modes are similar to the modes of hollow fibers, but, due to the presence of the photonic band gap in the transmission of the cladding, optical losses of hollow-core holey-fiber modes may be much lower than the optical losses characteristic of conventional hollow fibers.

One of the main advantages of holey fibers is that they support single-mode waveguiding within a remarkably broad spectral range [15, 16]. Fibers of this type provide exciting opportunities for the generation and control of ultrashort light pulses and for the control of luminescence of molecules within a broad spectral range [24]. Due to their properties, holey fibers offer new elegant solutions to many problems of fiber optics [14–24], nonlinear optics [21–26], the physics of photonic crystals and quantum electrodynamics [16, 27–31], atomic optics [30, 31], creation of optical frequency synthesizers and high-precision optical frequency measurements [32], biomedical optics [33], and optical data transmission [25].

In this paper, we develop and demonstrate the physical principles behind the control of light localization and nonlinear-optical interactions in micro- and nanostructured fibers. Transmission measurements on the cladding of nanostructured fibers having a form of a two-dimensional periodic structure with a pitch less than 500 nm have revealed the existence of a photonic band gap in transmission spectra of such claddings tunable within the range from 930 to 1030 nm. Below, we present the results of plane-wave simulations of optical properties of such fibers. We will also present the results of our studies devoted to the influence of the structure of the holey-fiber cladding on the effective area of the waveguide mode and the spectral broadening of ultrashort laser pulses in holey fibers with different structures of the core and the cladding and different air-filling fractions of the cladding. Our experimental results indicate that the increase in the air-filling fraction of a holey-fiber cladding results in a considerable enhancement of spectral broadening of short laser pulses due to the increase in the light localization degree in the fiber core.

2. TUNING THE PHOTONIC BAND GAP OF A PHOTONIC-CRYSTAL CLADDING OF NANOSTRUCTURED FIBERS

In the case of periodic arrangement of air holes in an HF cladding, the transmission spectrum of such a structure measured in the direction perpendicular to the direction of waveguiding displays photonic band gaps for certain directions of the wave vector. Within these frequency ranges, radiation cannot penetrate into the fiber cladding. Whenever a fiber without a hole at the center is used to produce a core in such a fiber, the fiber core can be considered as a defect in an otherwise perfect two-dimensional photonic-crystal lattice. Such structures are, in fact, one of the ways to solve the problem of fabrication of photonic crystals (see, e.g., [14, 24]), which can be employed to experimentally study the control of emission of atoms and molecules and light localization within the photonic band gap [4–6, 34–39].

Until recently, investigations of optical properties of holey fibers have been restricted to propagation regimes where the radiation wavelength is much less than the pitch of the photonic-crystal cladding and the core diameter. In this regime, the existence of a photonic band gap has, in fact, no influence on the propagation of light in a fiber. In this paper, we will demonstrate the physical principles of holey fibers with a photonic band gap (PBG) of the cladding tunable within the visible and near-IR spectral ranges. In Section 2.1, we will apply the method of plane-wave expansion to calculate the dispersion of a two-dimensional photonic-crystal structure in the cladding of a nanostructured fiber. The results of these calculations will then be used to determine the position of the photonic band gap and to specify the requirements to a photonic-crystal fiber guiding light due to the photonic band gap within the visible and near-IR ranges. Section 2.2 presents the results of experiments devoted to the observation of the photonic band gap in the transmission spectrum of the HF cladding and tuning of this photonic band gap.

2.1. Numerical Analysis of the Dispersion and Transmission Spectrum of Photonic-Crystal Fibers

To calculate the band structure of photonic energies for a photonic-crystal cladding of holey fibers, we employed a standard approach based on the plane-wave expansion (see [40]). As a model of an HF cladding, we considered an array of cylinders with a dielectric constant ε_2 and an infinite length arranged into a two-dimensional periodic triangular-lattice structure in a medium with a dielectric constant ε_1 (Fig. 1b).

The nodes of such a two-dimensional lattice with elementary translation vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ (see Fig. 1b) are defined by the following expression:

$$\mathbf{r}_i = i_1\mathbf{a}_1 + i_2\mathbf{a}_2, \quad (1)$$

where $i = \{i_1, i_2\}$, i_1 and i_2 being integers.

Let us also introduce a reciprocal lattice (Fig. 1c), defined by the vectors

$$\mathbf{g}_j = j_1\mathbf{b}_1 + j_2\mathbf{b}_2, \quad (2)$$

where $j = \{j_1, j_2\}$, j_1 and j_2 are integers, and translation vectors of the reciprocal lattice $\mathbf{b}_1$ and $\mathbf{b}_2$ can be found from the relations

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}, \quad i, j = 1, 2. \quad (3)$$

Since the dielectric constant of the system considered is a spatially periodic function, $\varepsilon(\mathbf{r}_{\parallel} + \mathbf{r}_i) = \varepsilon(\mathbf{r}_{\parallel})$,

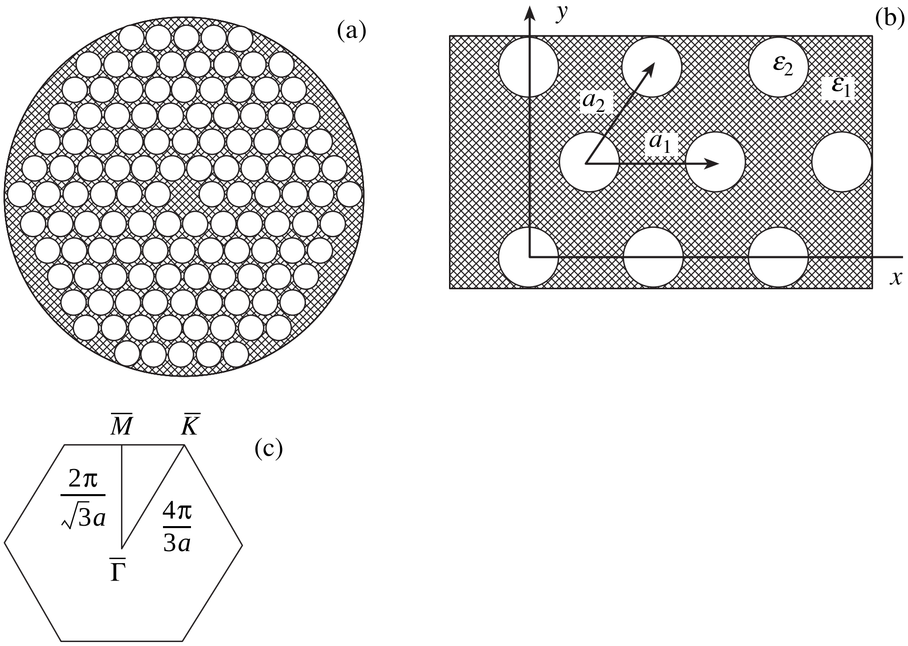


Fig. 1. (a) A holey fiber. Dark areas correspond to a material with a higher refractive index. (b) A two-dimensional periodic structure of the holey-fiber cladding. (c) The first Brillouin zone corresponding to a triangular lattice of air holes in a dielectric with characteristic directions determined by the points $\Gamma(0, 0)$, $M(0, 2\pi/\sqrt{3}a)$, and $K(2\pi/3a, 2\pi/\sqrt{3}a)$.

where $\mathbf{r}_\parallel$ is the radius vector in the XY plane, this quantity can be represented as a two-dimensional Fourier series:

$$\varepsilon(\mathbf{r}_\parallel) = \sum_{g_j}\varepsilon'(\mathbf{g}_j)\exp(i\mathbf{g}_j\cdot\mathbf{r}_\parallel). \tag{4}$$

In the case under study, expansion coefficients in Eq. (4) can be determined with the use of the following formula:

$$\varepsilon'(\mathbf{g}_j) = \begin{cases}\varepsilon_1 + (\varepsilon_2 - \varepsilon_1)f, & \mathbf{g}_j = 0\\ (\varepsilon_2 - \varepsilon_1)f\dfrac{2J_1(|\mathbf{g}_j|R)}{(|\mathbf{g}_j|R)}, & \mathbf{g}_j \neq 0,\end{cases} \tag{5}$$

where R is the radius of cylinders; J_1 is the first-order Bessel function; and f is the filling fraction, which is defined as the ratio of the cross-sectional area of a cylinder in the XY plane (πR^2) to the unit-cell area ($|\mathbf{a}_1 \times \mathbf{a}_2|$). Note that the quantity $\varepsilon^{-1}(\mathbf{r}_\parallel)$ is also periodic and can be expanded as a Fourier series:

$$\varepsilon^{-1}(\mathbf{r}_\parallel) = \sum_{g_j}\alpha'(\mathbf{g}_j)\exp(i\mathbf{g}_j\cdot\mathbf{r}_\parallel). \tag{6}$$

Expansion coefficients in Eq. (6) are given by formulas similar to Eq. (5) where the replacement $\varepsilon_i \longrightarrow \varepsilon_i^{-1}$ should be made.

Let us employ now the expressions derived above to find the dispersion relation for the two-dimensional periodic structure. We assume that a harmonic electric field $\mathbf{E}$ propagating in the XY plane is polarized along the axes of cylinders (E polarization) and is independent of z. The fields $\mathbf{E}$ and $\mathbf{H}$ in this case can be written as

$$\mathbf{E}(\mathbf{r}_\parallel, t) = \{0, 0, E_z(\mathbf{r}_\parallel)\}\exp(-i\omega t), \tag{7}$$

$$\mathbf{H}(\mathbf{r}_\parallel, t) = \{H_x(\mathbf{r}_\parallel), H_y(\mathbf{r}_\parallel), 0\}\exp(-i\omega t). \tag{8}$$

Substituting Eqs. (7) and (8) into Maxwell equations, we can derive the equation for $E_z(\mathbf{r}_\parallel)$:

$$\varepsilon^{-1}(\mathbf{r}_\parallel)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)E_z(\mathbf{r}_\parallel) + \frac{\omega^2}{c^2}E_z(\mathbf{r}_\parallel) = 0. \tag{9}$$

Applying the Bloch theorem, we can represent the field component E_z as a series

$$E_z(\mathbf{r}_\parallel) = \sum_{\mathbf{g}_j}B_{k_\parallel}(\mathbf{g}_j)\exp(-i(\mathbf{k}_\parallel + \mathbf{g}_j)\cdot\mathbf{r}_\parallel), \tag{10}$$

where $\mathbf{k}_\parallel$ is the component of the wave vector lying in the XY plane, $\mathbf{k}_\parallel = \{k_1, k_2, 0\}$. Substituting Eqs. (6) and (10) into Eq. (9), we arrive at the following equation for the function $B_{k_\parallel}(\mathbf{g}_j)$:

$$\sum_{\mathbf{g}_j}(\mathbf{k}_\parallel + \mathbf{g}_i)^2\alpha'(\mathbf{g}_j - \mathbf{g}_i)B_{\mathbf{k}_\parallel}(\mathbf{g}_j) = \frac{\omega^2}{c^2}B_{\mathbf{k}_\parallel}(\mathbf{g}_j). \tag{11}$$

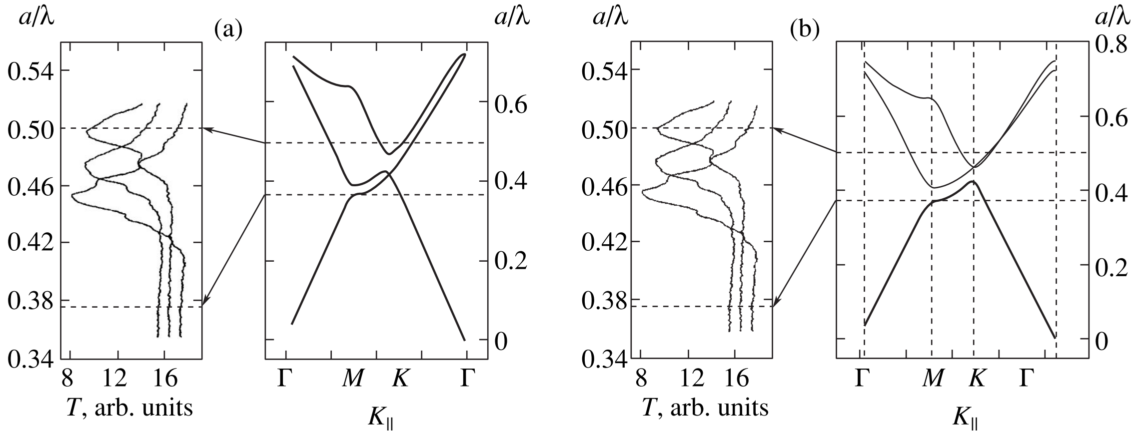


Fig. 2. The band structure of photon energies for (a) *E*- and (b) *H*-polarized radiation fields in a triangular-lattice cladding of a photonic-crystal fiber with the ratio of the hole radius to the pitch of the structure equal to 0.4 and the refractive index of glass equal to $n = 1.6$. The transmission *T* measured for nanostructured fibers is shown on the left.

Introducing the notation

$$A_{\mathbf{k}_\parallel} = \begin{bmatrix} (\mathbf{k}_\parallel + \mathbf{g}_0)^2\alpha'(\mathbf{g}_0 - \mathbf{g}_0) & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & (\mathbf{k}_\parallel + \mathbf{g}_i)^2\alpha'(\mathbf{g}_{i-1} - \mathbf{g}_i) & (\mathbf{k}_\parallel + \mathbf{g}_i)^2\alpha'(\mathbf{g}_i - \mathbf{g}_i) & \ldots \\ \ldots & \ldots & \ldots & \ldots \end{bmatrix}, \tag{12}$$

$$B_{\mathbf{k}_\parallel} = \begin{bmatrix} B_{\mathbf{k}_\parallel}(\mathbf{g}_0) \\ \ldots \\ B_{\mathbf{k}_\parallel}(\mathbf{g}_j) \\ \ldots \end{bmatrix}, \tag{13}$$

and

$$\mu^2 = \frac{\omega^2}{c^2} \tag{14}$$

and rewriting Eq. (11) as

$$A_{\mathbf{k}_\parallel} B_{\mathbf{k}_\parallel} = \mu B_{\mathbf{k}_\parallel}, \tag{15}$$

we can find the relation between $\mathbf{k}_\parallel$ and ω and the field distribution by numerically solving the eigenvalue and eigenfunction problems with the matrix $A_{\mathbf{k}_\parallel}$ (15). Note that, although formally the solution of this problem requires an infinite number of matrix elements, an accuracy of calculations no lower than 1% can be achieved with matrices consisting of approximately 400 × 400 elements. In the case of an *H*-polarized field, the problem can be solved in a similar way.

The band structure of photonic energies calculated with the use of the above-described procedure for *E*- and *H*-polarized radiation fields is presented in Figs. 2a and 2b, respectively. The results of our simulations show that the creation of holey fibers with a photonic band gap lying in the wavelength range accessible with the available standard femtosecond Ti:sapphire and forsterite lasers requires the fabrication of a two-dimensional periodic fiber structure with a pitch less than 500 nm. The results of these simulations qualitatively agree with the data of our experiments (presented on the left of Figs. 2a and 2b), which will be discussed in greater detail in the following section.

The results of simulations presented in Fig. 3 illustrate the possibility of tuning the photonic band gap by filling air holes in an HF cladding with material whose refractive index differs from the refractive index of the air. Calculations were performed for the case when the holes are filled with ethanol. Theoretical predictions concerning the possibility of tuning the photonic band gap are confirmed by our experimental data (see Section 2.2).

Figures 4a–4c present the band structures of photonic energies calculated for HF claddings with different air-filling fractions *f*. As can be seen from the data presented in these figures, the photonic band gap of an HF cladding can also be tuned by changing the content of air and glass in the fiber cladding. The experimental data presented in Section 3 of this paper indicate that the growth

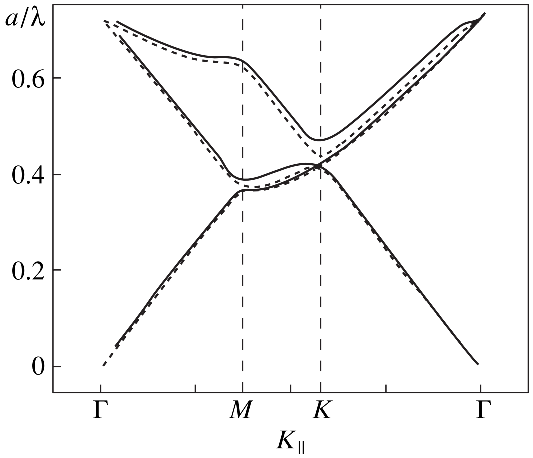


Fig. 3. The band structure of photon energies for an *E*-polarized radiation field in a triangular-lattice cladding of a photonic-crystal fiber with (solid curve) air- and (dashed curve) ethanol-filled holes, the ratio of the hole radius to the pitch of the structure equal to 0.4, and the refractive index of glass equal to $n = 1.6$.

in *f* also increases the light localization degree in the core of micro- and nanostructured fibers, which allows the efficiency of nonlinear-optical processes in a fiber to be noticeably improved.

2.2. Transmission Measurements on the Cladding of a Photonic-Crystal Fiber

Nanostructured fibers with a pitch of the cladding less than 500 nm were fabricated at the Institute of Technology and Processing of Glass Structures. The technology employed to fabricate such fibers was similar to the technology described in [14, 26–29] and was based on the following procedure. Identical glass capillaries were stacked into a periodic structure, which was then fused at a high temperature, in order to eliminate air gaps between the capillaries, and drawn. The resulting structure was cut into segments. These segments were also stacked into a periodic array, which underwent the drawing process again.

The above-described procedure allowed a fabrication of fibers with different cladding geometries (Fig. 5), a pitch ranging from 400 nm up to 32 μm (see also [27–29]), and the ratio of the hole diameter to the pitch of the structure variable within a broad range (Fig. 5). The central fiber in the stack had no hole in it and served as a core in the holey fiber.

The periodic structure in the cladding of a holey fiber gives rise to photonic band gaps in transmission spectra of the structure measured in the direction perpendicular to the direction of waveguiding (Figs. 2–4). Sub-500-nm-pitch two-dimensional periodic structures allowed us to observe photonic band gaps within the wavelength range characteristic of widespread lasers. The experimental setup for measuring the spectra of holey fibers was based on a Hitachi-333 spectrophotometer and included the signal and reference channels. A 5-cm-focal-length quartz lens was used in the signal channel to ensure the predominant illumination of the central part of the sample having a photonic-crystal structure. This lens focused the light beam on a slit diaphragm with an aperture $d = 250$ μm. A holey-fiber sample was placed behind the diaphragm. Radiation transmitted through the sample was collimated with a quartz lens, which was identical to the focusing lens. Transmission spectra were measured within the range of wavelengths from 400 to 1400 nm. A mercury lamp in the spectrometer was replaced by a tungsten lamp around 870 nm. To be able to measure transmission

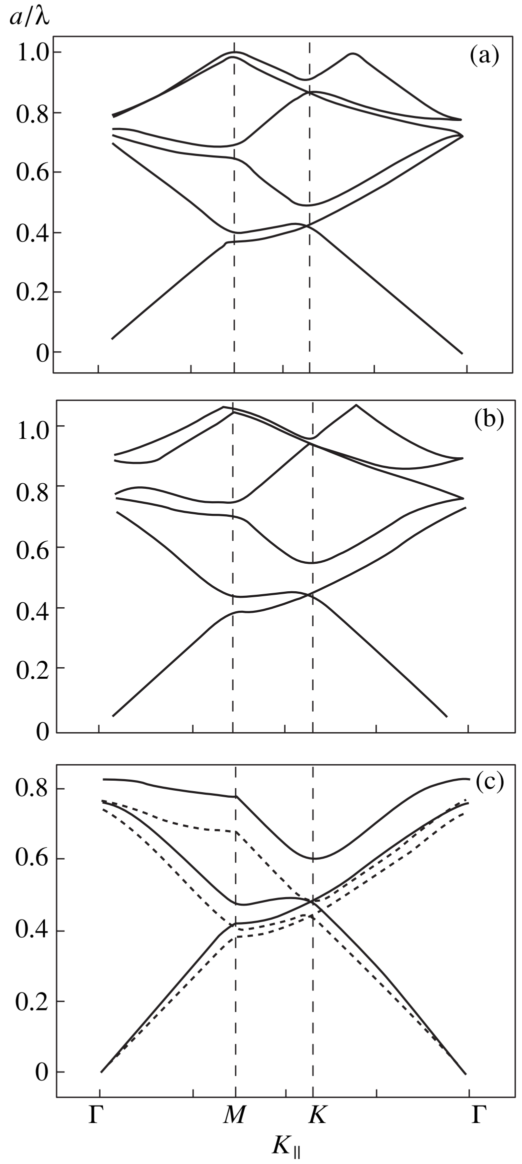


Fig. 4. The band structure of photon energies calculated for an *E*-polarized radiation field in a triangular-lattice cladding of a photonic-crystal fiber with the air-filling fraction (a) $f = 0.2$, (b) 0.4, and (c) 0.58. The holes of the HF cladding are filled with (solid curve) air and (dashed curve) ethanol.

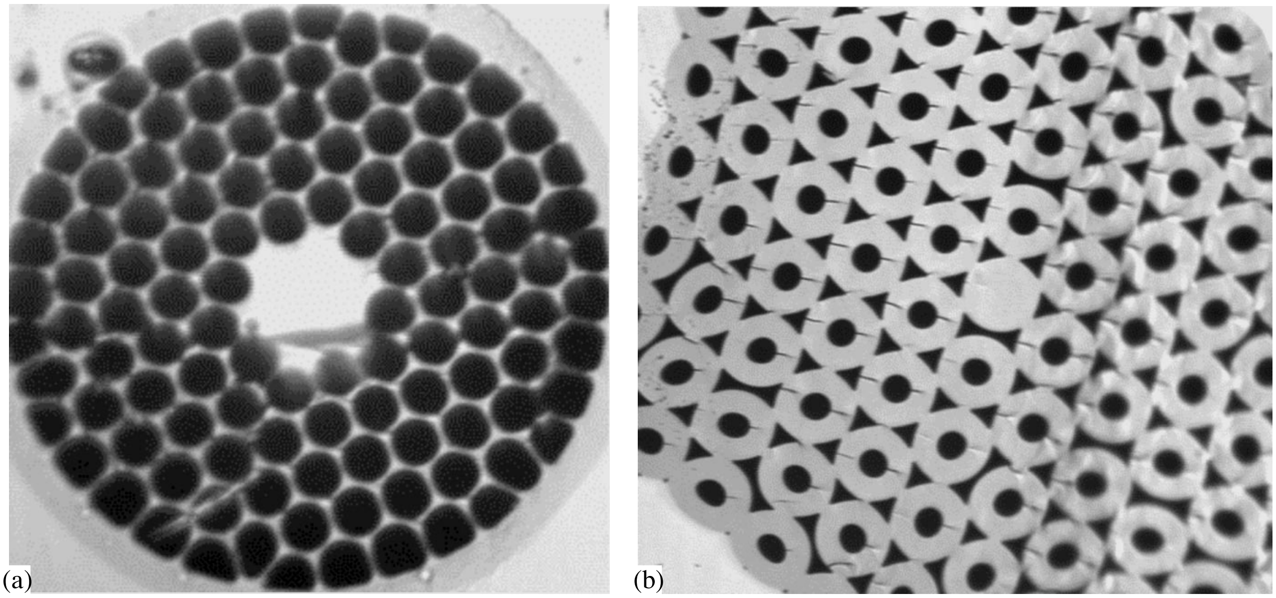


Fig. 5. Cross-sectional microscope images of holey fibers with a photonic-crystal cladding.

spectra of a photonic-crystal sample for different directions of the wave vector in the first Brillouin zone of the photonic-crystal lattice, we rotated the sample around its axis corresponding to the direction of waveguiding.

Transmission spectra measured for a holey fiber with a period of the PBG structure less than 500 nm are presented in Fig. 6. The photonic band gap in the transmission spectra of such samples was observed within the range of wavelengths from 930 to 1030 nm. Since air holes periodically arranged in the fiber cladding form a hexagonal lattice, the position of the photonic band gap in the transmission spectrum changes, depending on the rotation angle of the structure with respect to incident radiation. Comparison of the results of simulations and experimental data (Figs. 2a and 2b) shows that the position of the photonic band gap can be satisfactorily described within the framework of the plane-wave-expansion approach.

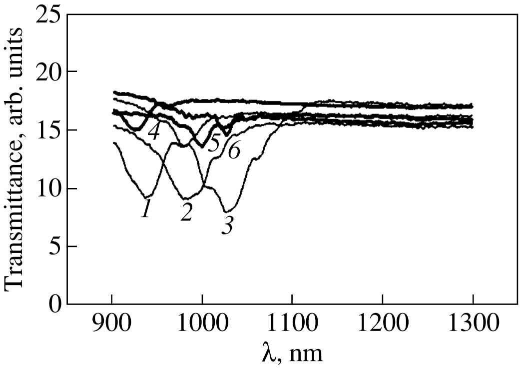


Fig. 6. Transmission spectra of a photonic-crystal fiber measured for different directions of the wave vector in the first Brillouin zone of the photonic-crystal lattice of the PBG cladding for (thin curves *1–3*) silica–air and (thick curves *4–6*) silica–ethanol PBG structures.

The structure of the photonic band gap in the transmission spectrum of a photonic-crystal structure depends on the ratio of the refractive indices of materials forming the structure (Fig. 3). In the case of holey fibers, this circumstance opens up the opportunity of tuning the photonic band gap by filling the air holes in the PBG structure with various materials whose refractive indices differ from unity.

We have studied changes in the photonic band gap of holey-fiber samples arising when the air holes of the structure were filled with ethanol. Both the position and the width of the photonic band gap noticeably changed in this case (curves *4–6* in Fig. 6). The experimentally demonstrated possibility of tuning the photonic band gap of a photonic-crystal cladding in holey fibers by using various materials to fill air holes in the fiber cladding offers broad opportunities for tuning the dispersion of holey fibers and controlling the luminescence of molecules within a broad spectral range.

3. NONLINEAR-OPTICAL INTERACTIONS OF ULTRASHORT PULSES IN MICROSTRUCTURED FIBERS

3.1. The Waveguide Mode Area and the Efficiency of Self-Phase Modulation in a Holey Fiber

Micro- and nanostructured fibers offer vast opportunities for enhancing nonlinear-optical interactions and for using the methods of nonlinear optics to produce ultrashort light pulses. Recent experiments [25, 26] have demonstrated that holey fibers can be employed to control the spectrum of ultrashort laser pulses and to generate a supercontinuum starting with very low energies of laser radiation.

The efficiency of nonlinear-optical processes in HFs, including the processes resulting in spectral broadening of femtosecond pulses, can be controlled by changing the localization degree of the light field in the fiber core. The calculation of the light-field distribution in the cross section of a holey fiber is a rather complicated problem. Several numerical methods have been recently developed for such simulations (e.g., see [21, 41–43]). Below, we will illustrate our idea of controlling field localization in an HF core using a simple qualitative approach. A microstructured cladding of an HF will be characterized by the effective refractive index [15]

$$n_{cl} = \beta_{cl}/k, \tag{16}$$

where β_{cl} is the propagation constant of the fundamental space-filling mode, i.e., the fundamental mode of an infinite structure obtained by periodically translating a unit cell of the HF cladding; $k = 2\pi/\lambda$; and λ is the radiation wavelength. Representing the effective refractive index of the cladding in the form of Eq. (16), we take into consideration the real spatial distribution of the light field in the fiber cladding. The profile of such a distribution can be estimated from two-dimensional images of radiation intensity distribution in the cross section of an HF.

To provide a rough estimate for the effective radius r of the waveguide mode in the HF core, we will employ the following formula for the radius of the waveguide mode in a conventional step-index fiber with a solid cladding [44]:

$$r = w + 1/p, \tag{17}$$

where w is understood as the HF core radius;

$$p^2 = \beta_c^2 - \beta_{cl}^2$$

is the transverse component of the wave vector in the fiber core;

$$\beta_c = n_c k \cos\varphi \tag{18}$$

is the propagation constant of the waveguide mode in the fiber core, which meets the conditions

$$n_c k \geq \beta_c \geq \beta_{cl}; \tag{19}$$

n_c is the refractive index of the fiber core; and φ is the incidence angle characteristic of the considered waveguide mode.

Using Eqs. (16)–(18), we arrive at the following estimate for the radius of the waveguide mode:

$$r = w + \frac{\lambda}{2\pi\sqrt{n_c^2\cos\varphi - n_{cl}^2}}. \tag{20}$$

As can be seen from Eq. (20), the localization degree of light field in a fiber core can be increased and, consequently, the efficiency of nonlinear-optical processes can be improved by increasing the difference between the refractive index of the core n_c and the effective refractive index of the cladding n_{eff}. Physically, the field localization degree can be increased in fibers with large differences between n_c and n_{eff} due to the fact that modes with large differences of propagation constants may exist in the core and the cladding of such fibers. The transverse component p of the wave vector of the mode propagating in the fiber core decreases under these conditions, which implies higher localization degrees of light field in the fiber core.

The relative deviation w of the frequency of a laser pulse induced by self-phase modulation (SPM) due to the nonlinear additive to the refractive index $\Delta n = n_2 I$ can be estimated with the use of the formula [45]

$$\frac{\Delta\omega}{\omega} = \frac{n_2}{c}\frac{P_0}{S\tau}L, \tag{21}$$

where c is the speed of light, P_0 is the peak power of the laser pulse, $S = \pi r^2$ is the effective waveguide mode area, τ is the pulse duration, and L is the fiber length.

As can be seen from Eq. (21), a decrease in S due to the increase in the air-filling fraction in a fiber core enhances the SPM-induced spectral broadening of a laser pulse. The difference between the refractive indices of the core and the cladding in HFs can be increased by changing the structure of the cladding and making the air holes in the cladding larger (Fig. 7). Note that the results of simulations presented in Section 2.1 indicate that the growth in the air-filling fraction f of the HF cladding results also in a shift of the photonic band gap of a two-dimensional periodic structure of the HF cladding (see Figs. 4a–4c). In our study, we employed this method to control light localization in the fiber core and to enhance nonlinear-optical processes in holey fibers.

3.2. Experimental

Titanium–sapphire and forsterite laser systems were employed in our experiments as sources of ultrashort light pulses. A Ti:sapphire laser consisted of an oscillator and a multipass amplifier pumped by the second harmonic of a pulsed Nd:YAG laser, operating at a repetition rate of 1 kHz. Laser pulses produced by this system had a duration ~70 fs and an energy up to 1 mJ.

Experiments were also performed with an all-solid-state self-starting Cr^{4+}:forsterite laser [33, 46], which allowed light pulses with durations less than 40 fs and radiation wavelength tunable within the range from 1.21 to 1.29 μm to be produced. A nonlinear crystal was used to double the frequency of this radiation. The master oscillator of this laser system included a Nd:YAG-pumped 19-mm Cr^{4+}:forsterite crystal, mirrors with a radius of curvature equal to 100 mm, and a 4.5% output

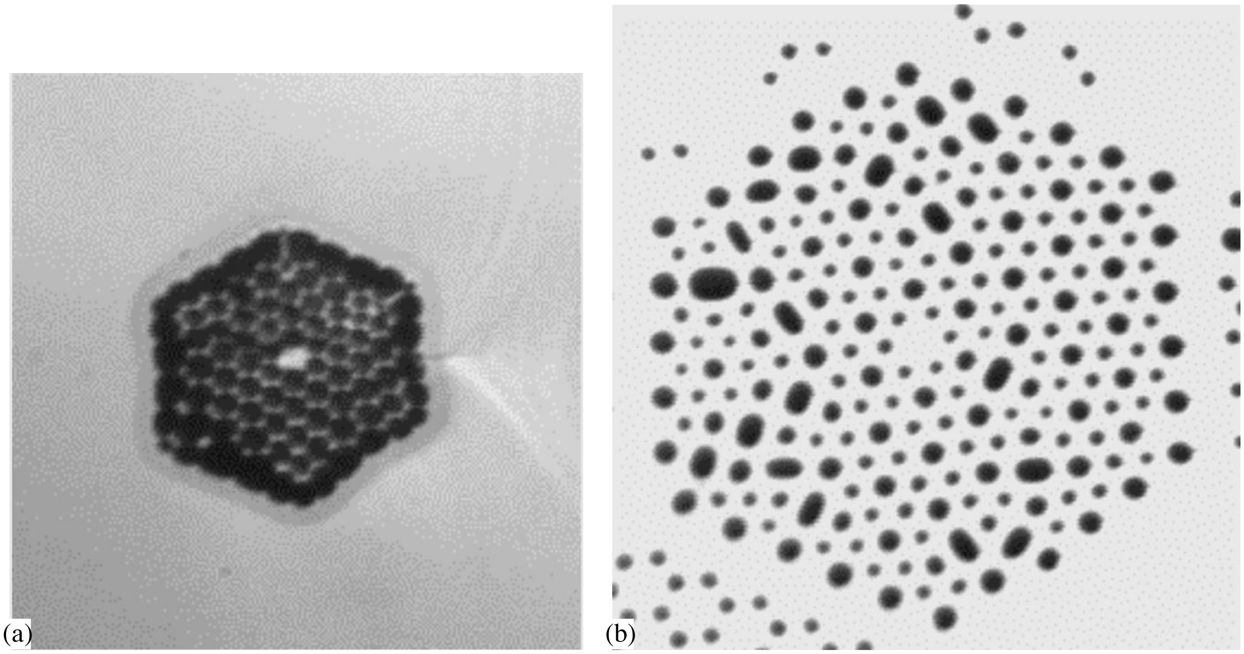


Fig. 7. Cross-sectional microscope images of holey fibers with a pitch of the photonic-crystal cladding equal to 2 μm. The air-filling fraction is (a) 65% and (b) 16%.

coupler. As an option, a semiconductor-saturable-absorber reflector could be used as a rear cavity mirror. Self-starting mode locking in the created system can be implemented both with and without semiconductor saturable-absorber mirrors.

The energy of laser radiation in our experiments was varied with the use of a half-wave plate and a Glan prism. A microobjective used to couple laser radiation into a holey fiber provided a coupling efficiency of 10–25%, depending on the size of the fiber core. The spectra of light pulses emerging from the fiber were analyzed with the use of a spectrometer and a CCD camera.

3.3. Results and Discussion

To study the possibilities of controlling nonlinear-optical processes in holey fibers by changing the ratio of refractive indices of the core and the microstructured cladding, we employed short HF samples with a length of 3–4 cm. Spectral broadening was investigated for 70-fs Ti:sapphire laser pulses coupled into HF samples within the range of radiation energies where no supercontinuum generation was observed. The values of pulse energy and power presented in figures and in the text are corrected for the input losses estimated using the light pulse energy measured of the holey fiber output. In particular, the appearance of the anti-Stokes component in the spectrum of a pulse emerging from a holey fiber, indicating the initial phase of supercontinuum generation, was observed with laser pulse powers on the order of 10 kW (Fig. 8). The length of fibers was also chosen sufficiently small in order to avoid effects related to group-velocity dispersion, which would otherwise have a considerable influence on short pulses.

The bandwidth of laser pulses transmitted through a holey fiber increased with the growth in the energy of laser pulses (Fig. 9). At the initial phase, this process can be approximately described by Eq. (21), which allows an estimation of the influence of the cladding structure and the air-filling fraction of the cladding on the efficiency of SPM, resulting in the spectral broadening of laser pulses.

Figure 9 presents the spectral broadening of Ti:sapphire laser pulses with a wavelength of 800 nm and a pulse duration of 70 fs at the output of a 2-μm-pitch HF sample with a length of 3 cm and different air-filling fractions of the cladding as a function of radiation energy coupled into the fiber. The results of these measurements show that the increase in the air-filling fraction of the HF cladding leads to a considerable enhancement of self-phase modulation. In particular, increasing the air-filling fraction of the cladding f from 16 (curve *2*) up to 65% (curve *1*), we observed the enhancement of spectral broadening of Ti:sapphire laser pulses by a factor of about 1.5.

The ratio of the effective mode areas in holey fibers with different structures can be estimated, in accordance with Eq. (21), as the ratio of the slopes of the spectral broadening of laser pulses at the output of a fiber as a function of radiation energy coupled into the fiber (Fig. 9). Such estimates for the contraction ratio of the effective waveguide mode area in a holey fiber agree well with the results of measurements based on the imaging of the output fiber end. This circumstance

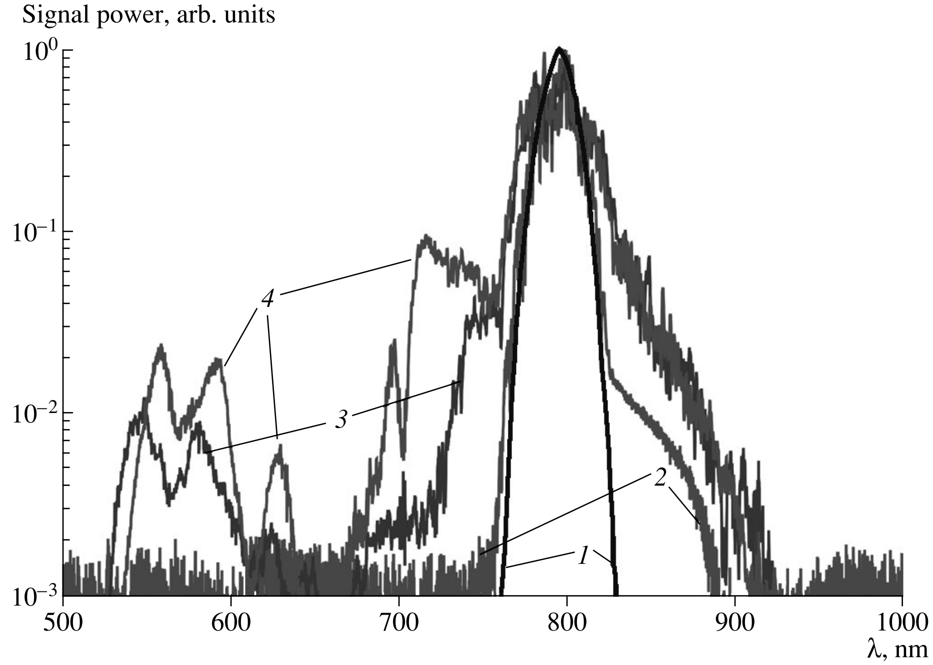


Fig. 8. The spectra of Ti:sapphire laser pulses with τ = 70 fs (*1*) at the input and (*2–4*) at the output of a holey fiber with a length of 3 cm and the pitch of the cladding equal to 3 μm for pulse energies of (*2*) 0.7, (*3*) 15, and (*4*) 35 nJ.

indicates that the enhancement of spectral broadening of laser pulses in our experiments is mainly due to the decrease in the effective mode area of radiation propagating in a holey fiber.

4. CONCLUSION

Thus, the analysis of optical properties of micro- and nanostructured fibers performed in this paper demonstrates the possibility of creating holey fibers with a photonic-crystal cladding whose photonic band gap lies within the frequency range characteristic of the available convenient femtosecond Ti:sapphire and forsterite lasers if a two-dimensional periodic structure with a period less than 500 nm is employed as a cladding in such fibers. The fabrication of holey fibers meeting this requirement allowed us to experimentally demonstrate the existence of a photonic band gap for such structures within the range of 930–1030 nm by measuring the transmission spectra of holey fibers in the direction perpendicular to the direction of waveguiding. This photonic band gap is satisfactorily described within the framework of the numerical approach based on the plane-wave expansion. The experimentally demonstrated possibility of tuning the photonic band gap of a photonic-crystal cladding in holey fibers by using various materials to fill air holes in the fiber cladding offers broad opportunities for tuning the dispersion of holey fibers and controlling radiative and photochemical processes within a broad spectral range.

The results of our experiments also demonstrate the possibility of controlling light localization and the efficiency of nonlinear-optical interactions of ultrashort laser pulses in holey fibers by changing the fiber structure. By increasing the air-filling fraction of HF cladding from 16 up to 65%, we were able to improve the efficiency of spectral broadening of 70-fs pulses of a Ti:sapphire laser by a factor of about 1.5 due to the increase in the light-localization degree in the core of the fiber.

The possibility of controlling the properties of waveguide modes in holey fibers by changing the structure of the HF cladding, which was demonstrated in this

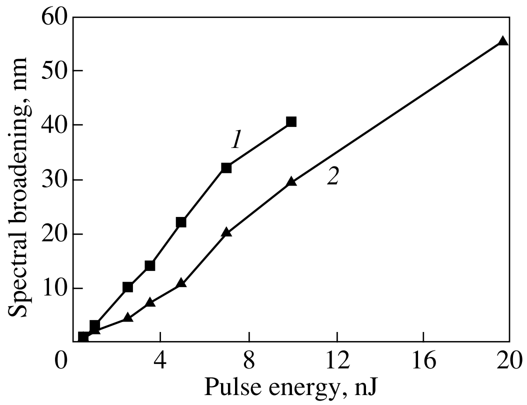


Fig. 9. The spectral broadening of 70-fs pulses of 800-nm Ti:sapphire laser radiation transmitted through a 2-μm-pitch HF sample with a length of 3 cm and the air-filling fraction equal to (*1*) 65% and (*2*) 16% as a function of the radiation energy coupled into the fiber.

paper, also seems to offer much promise as a method of tuning the dispersion of holey fibers. This dispersion tunability makes holey and photonic-crystal fibers very useful for the solution of many problems of fiber optics, information technologies, telecommunications, optics of ultrashort pulses, and nonlinear optics and spectroscopy.

ACKNOWLEDGMENTS

We are grateful to V.I. Beloglazov, N.B. Skibina, and Yu.S. Skibina for fabricating micro- and nanostructured fiber samples and to L.A. Mel'nikov for useful discussions. We would also like to express our gratitude to M. Aeschlimann and members of his group, M. Bauer, M. Scharte, and M. Wessendorf, for the help with the Ti:Sapphire laser.

This study was supported in part by the President of Russian Federation Grant no. 00-15-99304, the Russian Foundation for Basic Research project no. 00-02-17567, the Volkswagen Foundation (project I/76 869), Awards nos. RP2-2266 and RP2-2275 of the US Civilian Research and Development Foundation for the Independent States of the former Soviet Union (CRDF), and "Fundamental Metrology," "Fundamental Spectroscopy," and "Optics:Laser Physics" Federal Programs of the Ministry of Industry, Science, and Technologies of Russian Federation.

REFERENCES

1. A. H. Zewail, in *Femtochemistry and Femtobiology*, Ed. by V. Sundström (World Scientific, Singapore, 1998).
2. S. Pedersen, J. L. Herek, and A. H. Zewail, Science **266**, 1359 (1994).
3. P. Y. Cheng, D. Zhong, and A. H. Zewail, J. Chem. Phys. **105**, 6216 (1996).
4. E. Yablonovitch, J. Opt. Soc. Am. B **10**, 283 (1993).
5. J. Joannopoulos, R. Meade, and J. Winn, *Photonic Crystals* (Princeton Univ. Press, Princeton, 1995).
6. *Photonic Band Gaps and Localization*, Ed. by C. M. Soukoulis (Plenum, New York, 1993).
7. E. P. Petrov, V. N. Bogomolov, I. I. Kalosha, and S. V. Gaponenko, Phys. Rev. Lett. **81**, 77 (1998).
8. M. Megens, H. P. Schriemer, A. Lagendijk, and W. L. Vos, Phys. Rev. Lett. **83**, 5401 (1999); E. P. Petrov, V. N. Bogomolov, I. I. Kalosha, and S. V. Gaponenko, Phys. Rev. Lett. **83**, 5402 (1999).
9. R. A. Borisov, G. N. Dorojkina, N. I. Koroteev, *et al.*, Appl. Phys. B: Lasers Opt. **B67**, 765 (1998).
10. U. Gruening, V. Lehmann, S. Ottow, and K. Busch, Appl. Phys. Lett. **68**, 747 (1996).
11. V. V. Aristov, S. A. Magnitskii, V. V. Starkov, *et al.*, Laser Phys. **9**, 1260 (1999).
12. V. V. Aristov, S. A. Magnitskii, V. V. Starkov, *et al.*, Laser Phys. **10**, 946 (2000).
13. L. A. Golovan', A. M. Zheltikov, P. K. Kashkarov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 274 (1999) [JETP Lett. **69**, 300 (1999)].
14. J. C. Knight, T. A. Birks, P. St. J. Russell, and D. M. Atkin, Opt. Lett. **21**, 1547 (1996).
15. T. A. Birks, J. C. Knight, and P. St. J. Russell, Opt. Lett. **22**, 961 (1997).
16. J. C. Knight, J. Broeng, T. A. Birks, and P. St. J. Russell, Science **282**, 1476 (1998).
17. J. C. Knight, T. A. Birks, R. F. Cregan, *et al.*, Electron. Lett. **34**, 1347 (1998).
18. J. C. Knight, T. A. Birks, R. F. Cregan, *et al.*, Opt. Mater. **11**, 143 (1999).
19. R. F. Cregan, B. J. Mangan, J. C. Knight, *et al.*, Science **285**, 1537 (1999).
20. A. B. Fedotov, A. M. Zheltikov, A. A. Ivanov, *et al.*, Laser Phys. **10**, 723 (2000).
21. T. M. Monro, D. J. Richardson, N. G. R. Broderick, and P. J. Bennett, J. Lightwave Technol. **17**, 1093 (1999).
22. T. M. Monro, P. J. Bennett, N. G. R. Broderick, and D. J. Richardson, Opt. Lett. **25**, 206 (2000).
23. N. G. R. Broderick, T. M. Monro, P. J. Bennett, and D. J. Richardson, Opt. Lett. **24**, 1395 (1999).
24. A. M. Zheltikov, Usp. Fiz. Nauk **170**, 1203 (2000).
25. J. K. Ranka, R. S. Windeler, and A. J. Stentz, Opt. Lett. **25**, 25 (2000).
26. A. B. Fedotov, A. M. Zheltikov, L. A. Mel'nikov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 407 (2000) [JETP Lett. **71**, 281 (2000)].
27. A. B. Fedotov, A. A. Ivanov, M. V. Alfimov, *et al.*, Laser Phys. **10**, 1086 (2000).
28. M. V. Alfimov, A. M. Zheltikov, A. A. Ivanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 714 (2000) [JETP Lett. **71**, 489 (2000)].
29. A. B. Fedotov, M. V. Alfimov, A. A. Ivanov, *et al.*, Laser Phys. **11**, 138 (2001).
30. A. V. Tarasishin, A. M. Zheltikov, S. A. Magnitskiĭ, and V. A. Shuvaev, Kvantovaya Élektron. (Moscow) **30**, 843 (2000).
31. A. V. Tarasishin, S. A. Magnitskiĭ, V. A. Shuvaev, and A. M. Zheltikov, Opt. Commun. **184**, 391 (2000).
32. R. Holzwarth, Th. Udem, T. W. Hansch, *et al.*, Phys. Rev. Lett. **85**, 2264 (2000).
33. A. A. Ivanov, M. V. Alfimov, A. B. Fedotov, *et al.*, Laser Phys. **11**, 158 (2001).
34. P. Russell, *Two-Dimensional Photonic Crystals* (Ettore Majorana Center, Erice, 2000).
35. N. I. Koroteev, S. A. Magnitskii, A. V. Tarasishin, and A. M. Zheltikov, Opt. Commun. **159**, 191 (1999).
36. A. M. Zheltikov, A. V. Tarasishin, and S. A. Magnitskiĭ, Zh. Éksp. Teor. Fiz. **118**, 340 (2000) [JETP **91**, 298 (2000)].
37. T. Monro, *Exploring the Optical Properties of Holey Fibres* (Ettore Majorana Center, Erice, 2000).
38. S. John, *Photonic Band Gap Materials: A New Frontier in Quantum and Nonlinear Optics* (Ettore Majorana Center, Erice, 2000).

39. A. M. Zheltikov, *Controlling Light Pulses and Light Beams with Photonic Band-Gap Structures* (Ettore Majorana Center, Erice, 2000).
40. V. Kuzmiak, A. A. Maradudin, and A. R. McGurn, Phys. Rev. B **55**, 4298 (1997).
41. J. Broeng, S. E. Barkou, T. Söndergaard, and A. Bjarklev, Opt. Lett. **25**, 96 (2000).
42. A. Ferrando, E. Silvestre, J. J. Miret, *et al.*, Opt. Lett. **24**, 276 (1999).
43. A. Ferrando, E. Silvestre, J. J. Miret, *et al.*, Opt. Lett. **25**, 1328 (2000).
44. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman and Hall, New York, 1983; Radio i Svyaz', Moscow, 1987).
45. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984; Nauka, Moscow, 1989).
46. A. A. Ivanov, M. V. Alfimov, and A. M. Zheltikov, Laser Phys. **10**, 796 (2000).

Translated by A. Zheltikov

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 510–518.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 583–592.
Original Russian Text Copyright © 2001 by Gorbunov, Frolov.

PLASMA,
GASES

Collision of Two Short Laser Pulses in Plasma and the Generation of Short-Lived Bragg Mirrors

L. M. Gorbunov[a], * and A. A. Frolov[b]

[a]*Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 119991 Russia*
[b]*High Energy Density Research Center, Joint Institute for High Temperatures, Russian Academy of Sciences, ul. Izhorskaya 13/19, Moscow, 127412 Russia*
**e-mail: gorbun@sci.lebedev.ru*

Received December 15, 2000

Abstract—We consider the nonlinear excitation of small-scale electron-density perturbations when two identical short laser pulses propagating toward each other collide in plasma. Pulses with duration τ of the order of the plasma-oscillation period ($\omega_p\tau \le 1$, ω_p is the plasma frequency) are shown to excite long-lived localized plasma oscillations in the collision region. The energy conservation laws for the nonlinear mixing of short laser pulses in plasma are analyzed. We investigate the scattering of a sounding wave by the electron-density perturbations produced in the pulse collision region (short-lived Bragg mirror).

INTRODUCTION

The excitation of plasma oscillations during the nonlinear mixing of two high-frequency electromagnetic waves [1, 2] is of considerable interest in diagnosing laboratory [3, 4] and ionospheric [5] plasma and electron acceleration [6], as well as in maintaining a quasi-steady current in controlled thermonuclear fusion (CTF) studies [7]. When these issues are discussed, the amplitudes of reference electromagnetic waves are commonly assumed to change only slightly over the period of plasma oscillations; the latter are excited by the resonant force that arises because the reference-wave frequencies differ by the plasma frequency. Recently, interest in studying the various nonlinear effects that emerge when short (subpicosecond) laser pulses, whose duration is often shorter than the plasma-oscillation period, propagate in plasma has increased considerably [8]. Clearly, for the nonlinear mixing of such pulses, the concept of resonant plasma-wave excitation loses its meaning.

In this paper, we consider the nonlinear mixing that arises when two identical short laser pulses collide in plasma. The standing electromagnetic wave generated when the pulses overlap produces small-scale ponderomotive forces whose action gives rise to electron-density modulation. If the pulse duration τ is much longer than the plasma period ($\omega_p\tau \gg 1$, where ω_p is the plasma frequency), then the electron-density perturbations in the collision region persist only during the collision and disappear as the pulses move apart. The interaction causes the pulse shape to change only slightly, while the pulse energy is essentially constant. If, alternatively, the duration of the collision is shorter than the plasma-oscillation period ($\omega_p\tau \le 1$), then, after its completion and pulse divergence, long-lived localized plasma oscillations remain in the collision region, while the pulse energy decreases by the energy of these oscillations. In this case, the excitation of plasma oscillations by counterpropagating pulses with equal frequencies can be said to be an impact one. Such an impact excitation of natural oscillations and waves during the nonlinear mixing of short pulses with equal frequencies can also take place in other material media and in the case of other physical mechanisms responsible for nonlinear mixing. The impact excitation of acoustic waves during the nonlinear mixing of picosecond laser pulses, which produce spatially periodic matter heating, seems to have been first observed in molecular crystals [9, 10].

The excitation of plasma waves when two laser pulses collide is considered in Section 1. The collision effect on the pulse shape and energy is analyzed in Section 2. The scattering of a sounding wave by the short-lived Bragg mirror that emerges in the pulse collision region is studied in Section 3. In Conclusion, we discuss the possible applications of the effects considered.

1. THE EXCITATION OF SMALL-SCALE ELECTRON-DENSITY PERTURBATIONS AS LASER PULSES COLLIDE

Consider two identical short laser pulses with a carrier frequency ω_0 that propagate along the z axis toward each other in rarefied plasma with an electron density N_{0e} (Fig. 1a). We write the electric field of the laser pulses as

$$\mathbf{E}_L(\mathbf{r}, t) \doteq \frac{1}{2}\{\exp(-i\omega_0 t)[\mathbf{E}_+(\mathbf{r}, t)\exp(ik_0 z) + \mathbf{E}_-(\mathbf{r}, t)\exp(-ik_0 z)] + \text{c.c.}\}, \quad (1.1)$$

where

$$k_0 = \frac{\omega_0}{c}\sqrt{\varepsilon(\omega_0)}, \quad \varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega^2},$$

$$\omega_p = \sqrt{4\pi e^2 N_{0e}/m}$$

is the plasma frequency; e and m are the electron charge and mass; c is the speed of light; and $\mathbf{E}_\pm(\mathbf{r}, t)$ are the complex amplitudes of the laser pulses propagating from left to right and from right to left, respectively. $\mathbf{E}_\pm(\mathbf{r}, t)$ are assumed to vary slowly on scales k_0^{-1} in space and ω_0^{-1} in time. We assume the pulse lengths to be small compared to the diffraction length and write the pulse shape in the interaction region as

$$\mathbf{E}_+(\mathbf{r}, t) = \mathbf{E}_\perp(\rho)\exp\left(-\frac{\xi^2}{2L^2}\right),$$
$$\mathbf{E}_-(\mathbf{r}, t) = \mathbf{E}_\perp(\rho)\exp\left(-\frac{\eta^2}{2L^2}\right), \tag{1.2}$$

where

$$\xi = z - V_g t, \quad \eta = z + V_g t,$$

$V_g = (k_0/\omega_0)c^2$ is the group velocity of the pulses, L is their length, and the vector $\mathbf{E}_\perp(\rho)$ characterizes the pulse polarization and radial ($\rho = \sqrt{x^2 + y^2}$) shape. The pulses are assumed to be linearly polarized (e is the polarization vector) and axially symmetric. We chose the coordinate system and the initial time in such a way that, in the absence of interaction at $t = 0$, the pulses exactly overlap and the maxima of functions $\mathbf{E}_+$ and $\mathbf{E}_-$ are at $z = 0$ (Fig. 1b).

Note that, while propagating in plasma, the laser pulses also generate low-frequency and large-scale plasma fields, including wake waves [11]. The role of such fields in the pulse collision will be studied separately.

As the pulses approach each other, a small-scale (with wave number $2k_0$) ponderomotive force arises when they begin to overlap. This force produces the corresponding electron-density perturbations (indicated in Fig. 1b by the heavy solid curve),

$$\delta n = n_2\exp(2ik_0z) + n_2^*\exp(-2ik_0z), \tag{1.3}$$

with the perturbation amplitude n_2 slowly varying in time and space on scales ω_0^{-1} and $(2k_0)^{-1}$, respectively.

In the linear approximation, which holds for $|n_2| < N_{0e}$, and disregarding the thermal electron motion, the equation for n_2 is (see, e.g., [1])

$$\left(\frac{\partial^2}{\partial t^2} + \omega_p^2\right)\frac{n_2}{N_{0e}} = -\frac{e^2}{m^2\omega_0^2}\left(k_0^2 - ik_0\frac{\partial}{\partial z} - \frac{1}{4}\Delta_\perp\right)$$
$$\times\left\{\mathbf{E}_+\cdot\mathbf{E}_-^* + \frac{i}{\omega_0}\left(\mathbf{E}_+\cdot\frac{\partial\mathbf{E}_-^*}{\partial t} - \mathbf{E}_-^*\cdot\frac{\partial\mathbf{E}_+}{\partial t}\right)\right\}, \tag{1.4}$$

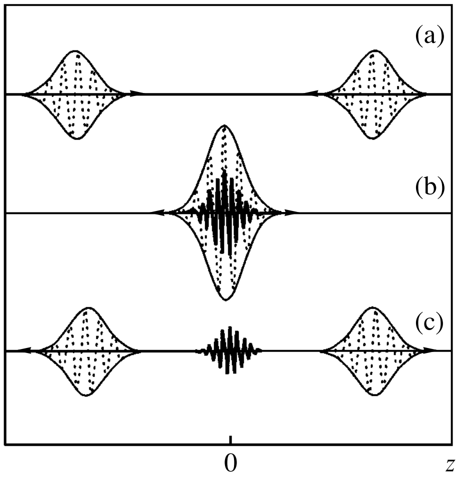


Fig. 1. A diagram for the collision of two laser pulses. (a) The pulses before the collision move toward each other; the laser-pulse electric-field envelope and the high-frequency filling are indicated by a thin solid line and dots, respectively. (b) The time of complete pulse overlapping; the small-scale ($2k_0$) electron-density perturbations are represented by a heavy solid curve. (c) The pulses after the collision; the small-scale plasma oscillations (heavy solid curve) remain in the collision region.

where

$$\Delta_\perp = \frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial}{\partial\rho}\right).$$

We assume the pulses to be changed only slightly by the collision and ignore this change when calculating δn (the validity condition is given below). Substituting expressions (1.2) into (1.4) yields the solution of Eq. (1.4) that satisfies the condition for the absence of electron-density perturbations before the pulse arrival to the interaction region:

$$\frac{n_2}{N_{0e}} = -\frac{e^2}{m^2\omega_0^2\omega_p^2}\exp\left(-\frac{z^2}{L^2}\right)$$
$$\times\left(k_0^2 + 2ik_0\frac{z}{L^2}\frac{\omega_p^2}{\omega_0^2} - \frac{1}{4}\Delta_\perp\right)|\mathbf{E}_\perp(\rho)|^2\Phi\left(\frac{t}{\tau}, \omega_p\tau\right). \tag{1.5}$$

Here, the time dependence of n_2 is described by the function

$$\Phi(x, a) = a\int_{-\infty}^{x}\sin[a(x-y)]\exp(-y^2)dy, \tag{1.6}$$

where $\tau = L/V_g$ is the pulse duration.

The time evolution of the plasma-electron-density perturbations is described by function (1.6). This function includes both the plasma oscillations that remain after the pulse divergence and the induced density perturbations that persist only during the pulse collision.

Bearing in mind the discussion of the energy conservation law for the laser-pulse collision, we write the equation for the energy density w_p of small-scale plasma perturbations as

$$\frac{\partial w_p}{\partial t} = -\frac{\omega_p^2}{16\pi\omega_0^2} \times\left\{\mathbf{E}_+\mathbf{E}_-^*\frac{\partial}{\partial t}\left(\frac{n_2^*}{N_{0e}}\right) + \mathbf{E}_-\mathbf{E}_+^*\frac{\partial}{\partial t}\left(\frac{n_2}{N_{0e}}\right)\right\}. \tag{1.7}$$

Using Eq. (1.5), we derive from (1.7)

$$w_p = \frac{e^2k_0^2|\mathbf{E}_\perp(\rho)|^4}{16\pi m^2\omega_0^4}\exp\left(-\frac{2z^2}{L^2}\right) \times\left\{\Phi^2\left(\frac{t}{\tau}, \omega_p\tau\right) + \frac{1}{\omega_p^2}\left[\frac{\partial}{\partial t}\Phi\left(\frac{t}{\tau}, \omega_p\tau\right)\right]^2\right\}. \tag{1.8}$$

To calculate the total energy of the plasma oscillations, we assume the pulses to have a Gaussian distribution in the transverse direction:

$$\mathbf{E}_\perp(\rho) = \mathbf{e}E_0\exp(-\rho^2/2R^2), \tag{1.9}$$

where E_0 is the maximum electric-field amplitude and R is the characteristic pulse width. Integrating expression (1.8) over volume then yields the total energy

$$W_p(t) = \frac{k_0^2V_E^2}{4\sqrt{2}\omega_0^2}W_L \times\left\{\Phi^2\left(\frac{t}{\tau}, \omega_p\tau\right) + \frac{1}{\omega_p^2}\left[\frac{\partial}{\partial t}\Phi\left(\frac{t}{\tau}, \omega_p\tau\right)\right]^2\right\}, \tag{1.10}$$

where W_L is the laser-pulse energy,

$$W_L = \frac{\pi^{3/2}E_0^2R^2L}{8\pi}. \tag{1.11}$$

As follows from (1.10), our assumption that the electron-density perturbations have a weak effect on the laser pulses holds if $V_E/c \ll 1$, where $V_E = eE_0/m\omega_0$ is the electron oscillation velocity in the laser field.

To analyze expressions (1.5) and (1.10), we use the standard function in plasma theory (see, e.g., [12]):

$$J_+(\beta) = \beta\exp\left(-\frac{\beta^2}{2}\right)\int_{i\infty}^{\beta}\exp\frac{x^2}{2}dx.$$

Using this function, we represent (1.6) as

$$\Phi(x, a) = \frac{a}{2\sqrt{2}}\exp(-x^2) \times\left[\frac{1}{\beta_+}J_+(\beta_+) - \frac{1}{\beta_-}J_+(\beta_-)\right], \tag{1.12}$$

where

$$\beta_\pm = -i\sqrt{2}x \pm \frac{a}{\sqrt{2}} = -i\beta'' \pm \beta'. \tag{1.13}$$

Since $x = t/\tau$ and $a = \omega_p\tau$ in expressions (1.5) and (1.10), the real parts of $\beta_\pm$ characterize the ratio of the pulse duration to the plasma-oscillation period, while the imaginary parts of $\beta_\pm$ give the time, in units of the pulse duration. Values of $|\beta''| \le 1$ correspond to the interaction (collision) time.

When analyzing the evolution of the perturbations in electron density (1.5) and energy (1.10), it can be broken down into several stages. The function $J_+(\beta)$ becomes simplified after the collision, when the pulses have moved far apart ($t \gg \tau$ and $t \gg \omega_p\tau^2/2$) [12]. According to (1.12), we have

$$\Phi\left(\frac{t}{\tau}, \omega_p\tau\right) = \sqrt{\pi}\omega_p\tau\exp\left(-\frac{\omega_p^2\tau^2}{4}\right)\sin(\omega_pt). \tag{1.14}$$

As we see from (1.14), the plasma oscillations excited by pulses with duration $\tau = \sqrt{2}/\omega_p$ have the largest amplitude.

For long laser pulses ($\omega_p\tau > \sqrt{2}$), both during ($t < \tau$) and after ($t > \tau$) their interaction on condition that $t < \omega_p\tau^2/2$, the evolution of the density perturbations is described by the function

$$\Phi\left(\frac{t}{\tau}, \omega_p\tau\right) = \omega_p\tau\left\{\frac{\omega_p\tau}{\omega_p^2\tau^2 + 4(t/\tau)^2}\exp\left(-\frac{t^2}{\tau^2}\right) + \frac{\sqrt{\pi}}{2}\exp\left(-\frac{\omega_p^2\tau^2}{4}\right)\sin(\omega_pt)\right\}. \tag{1.15}$$

As follows from (1.15), the first term, which describes quasi-static density perturbations, becomes exponentially small after the pulse collision ($t > \tau$); the second term, which describes the excitation of plasma oscillations, is exponentially small for all times.

Let us now turn to expression (1.10) for the total energy of small-scale plasma perturbations. After the collision ($t/\tau > 1$), we find using (1.14) that

$$W_p = \frac{\pi\omega_p^2\tau^2}{\sqrt{2}}\frac{k_0^2V_E^2}{4\omega_0^2}W_L\exp\left(-\frac{\omega_p^2\tau^2}{2}\right). \tag{1.16}$$

In the general case, time variations in the total energy of plasma perturbations for various $\omega_p\tau$ are shown in Fig. 2. For long pulses ($\omega_p\tau > \sqrt{2}$), the energy is seen to reach a maximum during the collision and to decrease after this process. This is because, at the initial collision stage, the laser pulses spend part of their energy on the excitation of quasi-static electron-density perturbations. Subsequently, since the phase difference between the driving force and the oscillations changes, this energy almost completely returns to the laser pulses at the final collision stage. By contrast, when short pulses ($\omega_p\tau \le \sqrt{2}$) collide, part of their energy remains in the interaction region as small-scale plasma oscillations.

2. CHANGES IN THE SHAPE AND ENERGY OF LASER PULSES AS THEY COLLIDE

In our approach, the change in the shape and energy of the laser pulses caused by their collision is small. However, we consider this issue in more detail, bearing in mind an additional possibility for detecting the pulse-excited plasma perturbations. Substituting (1.1) and (1.3) into the equations for the electric field yields expressions for the $\mathbf{E}_+$ and $\mathbf{E}_-$ amplitudes:

$$\left\{2i\omega_0\frac{\partial}{\partial t} - \frac{\partial^2}{\partial t^2} + c^2\left(\Delta_\perp + \frac{\partial^2}{\partial z^2} + 2ik_0\frac{\partial}{\partial z}\right)\right\}\mathbf{E}_+ = \omega_p^2\frac{n_2}{N_{0e}}\mathbf{E}_- + i\frac{\omega_p^2}{\omega_0}\mathbf{E}_-\frac{\partial}{\partial t}\left(\frac{n_2}{N_{0e}}\right), \tag{2.1}$$

$$\left\{2i\omega_0\frac{\partial}{\partial t} - \frac{\partial^2}{\partial t^2} + c^2\left(\Delta_\perp + \frac{\partial^2}{\partial z^2} - 2ik_0\frac{\partial}{\partial z}\right)\right\}\mathbf{E}_- = \omega_p^2\frac{n_2^*}{N_{0e}}\mathbf{E}_+ + i\frac{\omega_p^2}{\omega_0}\mathbf{E}_+\frac{\partial}{\partial t}\left(\frac{n_2^*}{N_{0e}}\right). \tag{2.2}$$

Here, we included small terms with the second derivatives of the slowly varying amplitudes, because they determine the change in pulse energy [13]. Small terms with the time derivatives of the density-perturbation amplitude were also included in the right-hand parts of Eqs. (2.1) and (2.2). They describe the energy transfer from laser emission to plasma oscillations.

To derive equations for the laser-pulse energy from Eqs. (2.1) and (2.2), we must define the time-averaged energy density and energy flux density $\mathbf{S}$ to within terms of the first order of smallness in the slow spatial and time derivatives. Let us first consider the definitions of these quantities:

$$w = \frac{1}{8\pi}\langle\mathbf{E}^2 + \mathbf{B}^2\rangle + \frac{1}{2}mN_{0e}\langle\mathbf{V}^2\rangle, \qquad \mathbf{S} = \frac{c}{4\pi}\langle\mathbf{E}\times\mathbf{B}\rangle, \tag{2.3}$$

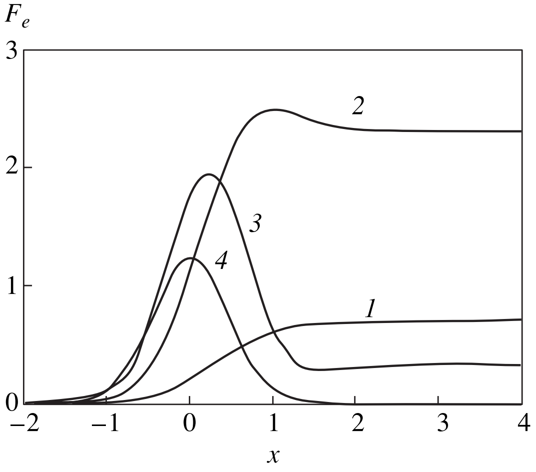


Fig. 2. Dimensionless plasma-perturbation energy $F_e = (W_p/W_L)(4\sqrt{2}\omega_0^2/k_0^2V_E^2)$ versus time ($x = t/\tau$). Curves *1–4* correspond to $\omega_p\tau = 0.5, \sqrt{2}$, 3, and 5.

where $\mathbf{V}$ is the electron velocity and the angular brackets denote averaging over the rapidly varying time dependence. Substituting $\mathbf{E}$, $\mathbf{B}$, and $\mathbf{V}$, which are expressed in terms of the complex amplitude of electric field $\mathscr{E}$ slowly varying on scales ω^{-1} and k^{-1},

$$\mathbf{E}(\mathbf{r}, t) = \frac{1}{2}\{\mathscr{E}(\mathbf{r}, t)\exp(-i\omega t + ikz) + \mathscr{E}^*(\mathbf{r}, t)\exp(i\omega t - ikz)\},$$

into (2.3), we obtain, to within terms of the first order of smallness in the slow derivatives,

$$w = \frac{1}{8\pi}\left\{|\mathscr{E}|^2 + \frac{i}{2\omega}\left(\mathscr{E}\cdot\frac{\partial\mathscr{E}^*}{\partial t} - \mathscr{E}^*\cdot\frac{\partial\mathscr{E}}{\partial t}\right) + \frac{ikc^2}{2\omega^2}\left(\mathscr{E}\cdot\frac{\partial\mathscr{E}^*}{\partial z} - \mathscr{E}^*\cdot\frac{\partial\mathscr{E}}{\partial z}\right)\right\}, \tag{2.4}$$

$$S_z = \frac{V_g}{8\pi}\left\{|\mathscr{E}|^2 + \frac{i}{2\omega}\left(\mathscr{E}\cdot\frac{\partial\mathscr{E}^*}{\partial t} - \mathscr{E}^*\cdot\frac{\partial\mathscr{E}}{\partial t}\right) + \frac{i}{2k}\left(\mathscr{E}\cdot\frac{\partial\mathscr{E}^*}{\partial z} - \mathscr{E}^*\cdot\frac{\partial\mathscr{E}}{\partial z}\right)\right\}, \tag{2.5}$$

$$\mathbf{S}_\perp = \frac{ic^2}{16\pi\omega}\{\mathrm{rot}[\mathscr{E}\times\mathscr{E}^*] + \mathcal{E}_k\nabla_\perp\mathcal{E}_k^* - \mathcal{E}_k^*\nabla_\perp\mathcal{E}_k\}, \tag{2.6}$$

where $\nabla_\perp$ is the transverse part of the nabla operator.

Multiplying Eqs. (2.1) and (2.2) by $\mathbf{E}_+^*$ and $\mathbf{E}_-^*$, respectively, and using definitions (2.4)–(2.6), we

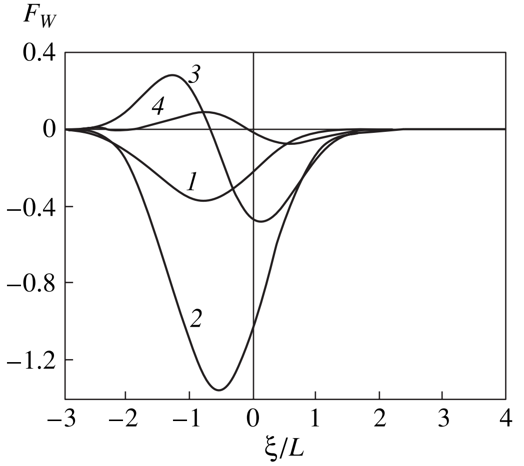


Fig. 3. Dimensionless pulse energy density $F_W = 32\pi m^2\omega_0^4\delta w_+/\sqrt{\pi}k_0^2e^2|\mathbf{E}_\perp|^4$ after the collision versus dimensionless accompanying variable ξ/L. The pulse moves from left to right. Curves *1–4* correspond to the same values of $\omega_p\tau$ as in Fig. 2.

derive equations for the change in laser-pulse energy density w_+ and w_-:

$$\frac{\partial w_+}{\partial t} + V_g\frac{\partial w_+}{\partial z} + \nabla_\perp \mathbf{S}_{+\perp}$$
$$= -\frac{\omega_p^2}{16\pi\omega_0^2}\left\{\frac{n_2}{N_{0e}}\mathbf{E}_+^*\cdot\left(i\omega_0 + \frac{\partial}{\partial t}\right)\mathbf{E}_- \right. \tag{2.7}$$
$$\left. + \frac{n_2^*}{N_{0e}}\mathbf{E}_+\cdot\left(-i\omega_0 + \frac{\partial}{\partial t}\right)\mathbf{E}_-^*\right\},$$

$$\frac{\partial w_-}{\partial t} - V_g\frac{\partial w_-}{\partial z} + \nabla_\perp \mathbf{S}_{-\perp}$$
$$= -\frac{\omega_p^2}{16\pi\omega_0^2}\left\{\frac{n_2^*}{N_{0e}}\mathbf{E}_-^*\cdot\left(i\omega_0 + \frac{\partial}{\partial t}\right)\mathbf{E}_+ \right. \tag{2.8}$$
$$\left. + \frac{n_2}{N_{0e}}\mathbf{E}_-\cdot\left(-i\omega_0 + \frac{\partial}{\partial t}\right)\mathbf{E}_+^*\right\}.$$

Adding up expressions (2.7) and (2.8) yields an equation for the change in the total energy of the two pulses:

$$\frac{\partial}{\partial t}(w_+ + w_-) + V_g\frac{\partial}{\partial z}(w_+ - w_-) + \nabla_\perp(\mathbf{S}_{+\perp} + \mathbf{S}_{-\perp})$$
$$= -\frac{\omega_p^2}{16\pi\omega_0^2}\left\{\frac{n_2}{N_{0e}}\frac{\partial}{\partial t}(\mathbf{E}_+^*\cdot\mathbf{E}_-) + \frac{n_2^*}{N_{0e}}\frac{\partial}{\partial t}(\mathbf{E}_-^*\cdot\mathbf{E}_+)\right\}. \tag{2.9}$$

The right-hand side of Eq. (2.9) describes the change in the energy of the laser pulses because of their interaction, which gives rise to plasma perturbations. Clearly, the total energy of the pulses and plasma oscillations must be conserved in this case. This directly follows from Eqs. (2.9) and (1.7), whose addition yields

$$\frac{\partial}{\partial t}(w_+ + w_- + w_p + w_{\text{int}})$$
$$+ V_g\frac{\partial}{\partial z}(w_+ - w_-) + \nabla_\perp(\mathbf{S}_{+\perp} + \mathbf{S}_{-\perp}) = 0, \tag{2.10}$$

where the interaction energy density is

$$w_{\text{int}} = \frac{\omega_p^2}{16\pi\omega_0^2}\left\{\frac{n_2}{N_{0e}}\mathbf{E}_+^*\cdot\mathbf{E}_- + \frac{n_2^*}{N_{0e}}\mathbf{E}_-^*\cdot\mathbf{E}_+\right\}. \tag{2.11}$$

To determine the change in the energy and shape of the pulses after their collision, we turn to Eqs. (2.7) and (2.8). Let us consider the times t that, on the one hand, are large compared to the pulse duration ($t \gg \tau$) but, on the other hand, are small compared to $k_0R^2/2c$. In this case, the change in the shape of the pulses because of their diffraction may be disregarded.

Bearing in mind that the collision causes the initial energy density distributions in the pulses, $w_{0\pm}(z \mp V_gt)$, to change by small values $\delta w_\pm(\rho, z, t)$, we find from (2.7) that

$$\left(\frac{\partial}{\partial t} + V_g\frac{\partial}{\partial z}\right)\delta w_+ = -\frac{\omega_p^2}{16\pi\omega_0^2}$$
$$\times\left\{\frac{n_2}{N_{0e}}\mathbf{E}_+^*\cdot\left(i\omega_0 + \frac{\partial}{\partial t}\right)\mathbf{E}_- + \frac{n_2^*}{N_{0e}}\mathbf{E}_+\cdot\left(-i\omega_0 + \frac{\partial}{\partial t}\right)\mathbf{E}_-^*\right\}. \tag{2.12}$$

A similar equation for δw_- follows from (2.8).

Using Eq. (1.4) to determine n_2 in (2.12), after some mathematical operations, we obtain for the change in pulse energy after the collision ($t \gg \tau$)

$$\delta w_+ = \frac{\sqrt{\pi}k_0^2e^2|\mathbf{E}_\perp(\rho)|^4}{32\pi m^2\omega_0^4}\left[\frac{\omega_0^2}{c^2k_0^2}L\frac{\partial}{\partial\xi} + \frac{2\xi}{L}\right]$$
$$\times\exp\left(-\frac{2\xi^2}{3L^2}\right)\Phi\left(-\frac{\xi}{\sqrt{3}L}, \frac{2\omega_p\tau}{\sqrt{3}}\right). \tag{2.13}$$

Figure 3 shows a plot of pulse energy density after the collision against accompanying variable ξ. The pattern of energy-density variations is seen to depend significantly on pulse duration. Short pulses ($\omega_p\tau \le \sqrt{2}$) lose their energy over the entire pulse length, more effectively in its second half (curves *1* and *2*), through the excitation of plasma oscillations, which remain in the collision region. For longer pulses (or for a denser plasma), quasi-static small-scale electron-density perturbations appear during the collision and disappear after its completion. During the first half of the collision, the pulses lose their energy, while, during its second half, the energy from the density perturbations returns to the pulses (curves *3* and *4*). As a result, the total energy of the pulses is virtually constant, while their shape changes.

To determine the change in the total energy of the pulse through its collision, let us integrate expression (2.13) over volume by using formula (1.9) for $\mathbf{E}_\perp(\rho)$. As might be expected, we then obtain an energy that is half the plasma-oscillation energy (1.16). The second half of the energy is transferred to the plasma oscillations by the oppositely moving pulse.

Note that, while propagating in plasma, the laser pulses also lose their energy and change their shape through other linear and nonlinear processes (Coulomb collisions, dispersion, stimulated Raman scattering, etc.). As the pulses collide, an additional mechanism for the change in their shape and energy arises.

3. THE SCATTERING OF A SOUNDING ELECTROMAGNETIC WAVE BY SMALL-SCALE LOCALIZED PLASMA PERTURBATIONS

In this section, we investigate the reflection (scattering) of a sounding electromagnetic wave from the electron-density perturbations excited when two counterpropagating laser pulses collide in plasma. Let a plane electromagnetic wave with frequency ω_1 ($\omega_1 \gg \omega_p$) and amplitude $\mathbf{E}_1$, whose electric field is

$$\mathbf{E}_p = \mathbf{E}_1 \cos(\omega_1 t - \mathbf{k}_1 \cdot \mathbf{r}), \tag{3.1}$$

propagate at angle α to the z axis along the wave vector

$$\mathbf{k}_1 = k_1(\mathbf{e}_x \sin\alpha + \mathbf{e}_z \cos\alpha),$$

where

$$k_1 = \frac{\omega_1}{c}\sqrt{\varepsilon(\omega_1)},$$

$\mathbf{e}_x$ and $\mathbf{e}_z$ are the unit vectors of a Cartesian coordinate system. The interaction of field (3.1) with the density perturbations results in the excitation of a scattered wave, whose electric field $\mathbf{E}_S(\mathbf{r}, t)$ satisfies the standard equation in the scattering theory [14]:

$$\frac{\partial^2}{\partial t^2}\mathbf{E}_S + \omega_p^2 \mathbf{E}_S + c^2 \operatorname{rot}\operatorname{rot}\mathbf{E}_S = -\omega_p^2 \frac{\delta n}{N_{0e}}\mathbf{E}_p. \tag{3.2}$$

By performing a Fourier transformation in time and coordinates in (3.2) and, subsequently, an inverse Fourier transformation in coordinates, we derive the following expression for $\mathbf{E}_S(\omega, \mathbf{r})$ in the wave zone ($r \gg R, L$) far from the collision region using formulas (1.3) and (1.5):

$$\begin{aligned} \mathbf{E}_S(\omega, \mathbf{r}) &= \pi\omega_p^2 (k_0 R)(k_0 L)\tau \\ &\times \frac{R}{2r}\frac{V_E^2}{4c^2}[\mathbf{e}_r \times [\mathbf{e}_r \times \mathbf{E}_1]]\exp(i\chi r)(\Psi_+ + \Psi_-), \end{aligned} \tag{3.3}$$

where

$$\begin{aligned} &\Psi_\pm \\ &= \frac{\exp\left(-\dfrac{(\omega \pm \omega_1)^2\tau^2}{4} - \dfrac{(\mathbf{e}_\rho \chi \sin\theta \pm \mathbf{k}_{1\perp})^2 R^2}{4}\right)}{(\omega \pm \omega_1)^2 - \omega_p^2 + 2i\gamma_L(\omega \pm \omega_1)} \\ &\times \left\{ \exp\left[-\frac{(\chi\cos\theta \pm k_{1z} - 2k_0)^2 L^2}{4}\right] \right. \\ &\left. + \exp\left[-\frac{(\chi\cos\theta \pm k_{1z} + 2k_0)^2 L^2}{4}\right] \right\}. \end{aligned} \tag{3.4}$$

Here,

$$\mathbf{k}_{1\perp} = \mathbf{e}_x k_1 \sin\alpha, \quad k_{1z} = k_1 \cos\alpha;$$
$$\mathbf{e}_r = \mathbf{e}_z \cos\theta + \mathbf{e}_\rho \sin\theta$$

is the unit vector along the radius vector $\mathbf{r}$;

$$\mathbf{e}_\rho = \mathbf{e}_x \cos\phi + \mathbf{e}_y \sin\phi$$

is the projection of vector $\mathbf{e}_r$ onto the xy plane; θ and ϕ are the azimuthal and polar angles, respectively; and

$$\chi = \frac{\omega}{c}\sqrt{\varepsilon(\omega)}.$$

To avoid singularities, formula (3.4) includes the damping of small-scale plasma oscillations with a decrement $\gamma_L \ll \omega_p$.

One of the main characteristics of scattered emission is the spectral energy flux density vector $\mathbf{p}(\omega, \mathbf{r})$. This vector can be expressed in terms of Fourier components of the electric and magnetic fields for positive frequencies $\omega > 0$ as follows [15]:

$$\mathbf{p}(\omega, \mathbf{r}) = \frac{c}{8\pi^2}\{\mathbf{E}_S(\omega, \mathbf{r}) \times \mathbf{B}_S^*(\omega, \mathbf{r}) + \text{c.c.}\}. \tag{3.5}$$

Vector (3.5) gives the energy emitted into unit solid angle $dO = \sin\theta d\theta d\phi$ in frequency interval $d\omega$:

$$\frac{dW_S(\omega)}{d\omega dO} = r^2 \mathbf{e}_r \cdot \mathbf{p}(\omega, \mathbf{r}). \tag{3.6}$$

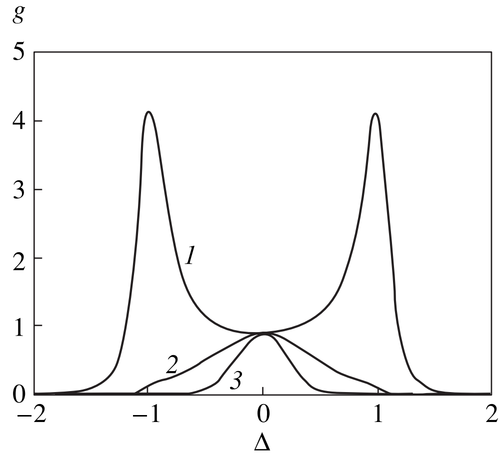


Fig. 4. Spectral density of the scattered radiation energy $g(\omega)$ (3.8) versus dimensionless frequency $\Delta = (\omega - \omega_1)/\omega_p$ for $\gamma_L/\omega_p = 0.15$ and various $\omega_p\tau$. Curves *1–3* correspond to $\omega_p\tau = \sqrt{2}$, 3, and 5, respectively.

When $\omega_1\tau \gg 1$, $k_1R \gg 1$, and $k_1L \gg 1$, taking into account relations (3.3)–(3.5), the scattered energy (3.6) can be written in the form

$$\begin{aligned}\frac{dW_S(\omega)}{d\omega dO} &= c\sqrt{\varepsilon(\omega_1)}[\mathbf{e}_1 \times \mathbf{e}_r]^2 E_1^2 R^2 \\ &\times \tau^2 (k_0R)^2 (k_0L)^2 \left(\frac{V_E}{4c}\right)^4 g(\omega) F^2(\theta, \phi),\end{aligned} \tag{3.7}$$

where $\mathbf{e}_1$ is the polarization vector of the incident wave, and the functions $g(\omega)$ and $F(\theta, \phi)$ describe the frequency and angular dependences, respectively:

$$\begin{aligned}g(\omega) &= \frac{\omega_p^4}{\{(\omega - \omega_1)^2 - \omega_p^2\}^2 + 4\gamma_L^2\omega_p^2} \\ &\times \exp\left\{-\frac{(\omega - \omega_1)^2\tau^2}{2}\right\},\end{aligned} \tag{3.8}$$

$$\begin{aligned}F(\theta, \phi) &= \exp\left\{-\frac{k_1^2R^2}{4}\right. \\ &\times (\sin^2\theta + \sin^2\alpha - 2\sin\theta\sin\alpha\cos\phi) \\ &\left. - \frac{k_1^2L^2}{4}\left(\cos\theta - \cos\alpha + \frac{2k_0}{k_1}\right)^2\right\}.\end{aligned} \tag{3.9}$$

Note that $W_S(\omega)$ is the spectral density of the energy scattered during the entire lifetime of the density perturbations. If the damping of plasma oscillations is ignored, they live infinitely long and $W_S(\omega)$ at $\omega = \omega_1 \pm \omega_p$ has a singularity. Taking into account the dissipation, the scattering time is limited by the lifetime of the plasma oscillations, and the spectral density characterized by function (3.8) is finite at $\omega = \omega_1 \pm \omega_p$.

The scattered spectrum [function (3.8)] is shown in Fig. 4 at various $\omega_p\tau$. For long pulses (or for a dense plasma), when $\omega_p\tau \gg 1$, the radiation at frequency $\omega \approx \omega_1$ dominates the spectrum. When short pulses ($\omega_p\tau \le 1$) collide, the scattered spectrum exhibits two narrow lines with frequencies $\omega = \omega_1 \pm \omega_p$.

The angular distribution of the scattered radiation is described by function (3.9). The direction for which the scattered energy is at a maximum is characterized by the angles $\phi = 0$ and $\theta = \pi - \alpha$, which corresponds to the satisfaction of the Bragg condition:

$$k_0 = k_1\cos\alpha. \tag{3.10}$$

The scattered radiation has a narrow directivity diagram, and if

$$\cos\alpha \gg \frac{1}{k_1R}, \frac{1}{k_1L},$$

the function $F(\theta, \phi)$ takes a Gaussian shape with the angular widths

$$\Delta\theta = \max\left(\frac{2}{k_1R\cos\alpha}, \frac{2}{k_1L\sin\alpha}\right),$$

$$\Delta\phi = \frac{2}{k_1R\sin\alpha}.$$

Integrating expression (3.7) over frequency yields the total energy scattered into an element of solid angle dO,

$$\begin{aligned}\frac{dW_S}{dO} &= c\sqrt{\varepsilon(\omega_1)}[\mathbf{e}_1 \times \mathbf{e}_r]^2 E_1^2 R^2 \\ &\times \tau^2 (k_0R)^2 (k_0L)^2 \left(\frac{V_E}{4c}\right)^4 G F^2(\theta, \phi),\end{aligned} \tag{3.11}$$

where

$$G = \int_0^\infty g(\omega)d\omega$$

is generally complex in form. It is significantly simplified in the limit of short ($\omega_p\tau \ll 1$) and long ($\omega_p\tau \gg 1$) pulses. In the former case, $G \approx \pi\omega_p^2/2\gamma_L$ and is determined by the damping of plasma oscillations γ_L. In the latter case, $G \approx \sqrt{2\pi}/\tau$ and depends on the pulse collision duration.

So far, we have considered the spectral properties of the radiation scattered during the entire lifetime of the electron-density perturbations. Let us now analyze the spatiotemporal structure of the electric field for the scattered radiation in a far zone. Using an inverse Fourier transformation in time in formula (3.3), we obtain

$$\mathbf{E}_S(\mathbf{r}, t) = \sqrt{\pi}(k_0R)(k_0L)\frac{R}{r}\frac{V_E^2}{4c^2}[\mathbf{e}_r \times [\mathbf{E}_1 \times \mathbf{e}_r]] \times \Phi_S\left(\frac{t}{\tau} - \frac{r}{V_1\tau}, \omega_p\tau\right)F(\theta, \phi)\cos(\omega_1 t - k_1 r). \tag{3.12}$$

Here, the function Φ_S is a generalization of formula (1.6) to damped plasma oscillations, and it can be represented in a form similar to (1.12):

$$\Phi_S\left(\frac{t}{\tau} - \frac{r}{V_1\tau}, \omega_p\tau\right) = \frac{\omega_p\tau}{2\sqrt{2}} \times \exp\left\{-\frac{1}{\tau^2}\left(t - \frac{r}{V_1}\right)^2\right\}\left\{\frac{J_+(\gamma_+)}{\gamma_+} - \frac{J_+(\gamma_-)}{\gamma_-}\right\}, \tag{3.13}$$

where

$$\gamma_\pm = -i\frac{\sqrt{2}}{\tau}\left(t - \frac{r}{V_1}\right) \pm \frac{\omega_p \pm i\gamma_L}{\sqrt{2}}\tau,$$

and $V_1 = c\sqrt{\varepsilon(\omega_1)}$ is the group velocity of the scattered wave in plasma. Using an asymptotic expansion for $J_\pm(\gamma_\pm)$, we find that the electric field (3.12) for short pulses ($\omega_p\tau < 1$) takes the form

$$\mathbf{E}_S(\mathbf{r}, t) = \pi\omega_p\tau(k_0R)(k_0L)\frac{R}{2r}\frac{V_E^2}{4c^2}[\mathbf{e}_r \times [\mathbf{E}_1 \times \mathbf{e}_r]] \times \exp\left\{-\frac{\omega_p^2\tau^2}{4} - \gamma_L\left(t - \frac{r}{V_1}\right)\right\} \times F(\theta, \phi)\left\{\sin\left[(\omega_1 + \omega_p)t - \left(k_1 + \frac{\omega_p}{V_1}\right)r\right] - \sin\left[(\omega_1 - \omega_p)t - \left(k_1 - \frac{\omega_p}{V_1}\right)r\right]\right\}. \tag{3.14}$$

Expression (3.14) describes two damped (with time) electromagnetic waves with frequencies ($\omega_1 \pm \omega_p$), which result from the scattering of the sounding radiation by excited plasma oscillations in the laser-pulse collision region. By contrast, the electric field of the scattered wave when longer pulses ($\omega_p\tau > 1$) collide has the shape of a pulse with duration τ and, according to (3.12) and (3.13), is given by

$$\mathbf{E}_S(\mathbf{r}, t) = \sqrt{\pi}(k_0R)(k_0L)\frac{R}{r}\frac{V_E^2}{4c^2}[\mathbf{e}_r \times [\mathbf{E}_1 \times \mathbf{e}_r]] \times \left\{\frac{\omega_p^2\tau^2}{\omega_p^2\tau^2 + \frac{4}{\tau^2}\left(t - \frac{r}{V_1}\right)^2}\exp\left[-\frac{1}{\tau^2}\left(t - \frac{r}{V_1}\right)^2\right] + \frac{\sqrt{\pi}\omega_p\tau}{2}\exp\left[-\frac{\omega_p^2\tau^2}{4} - \gamma_L\left(1 - \frac{r}{V_1}\right)\right] \times \sin\left[\omega_p\left(t - \frac{r}{V_1}\right)\right]\right\}F(\theta, \phi)\cos(\omega_1 t - k_1 r). \tag{3.15}$$

In this case, the sounding electromagnetic wave is reflected from the quasi-static electron-density perturbations only during the pulse collision time, and the radiation frequency is virtually constant.

The concept of effective cross section, which is defined as the ratio of the average radiation intensity to the energy flux density of the incident wave $c\sqrt{\varepsilon(\omega_1)}E_1^2(8\pi)^{-1}$, is commonly used to consider scattering by localized structures. Using formula (3.12) to determine the radiation intensity, we obtain the differential cross section

$$\frac{d\sigma}{dO} = 16\pi R^2[\mathbf{e}_1 \times \mathbf{e}_r]^2(k_0R)^2(k_0L)^2 \times \left(\frac{V_E}{4c}\right)^4\Phi_S^2\left(\frac{t}{\tau} - \frac{r}{V_1\tau}, \omega_p\tau\right)F^2(\theta, \phi), \tag{3.16}$$

where functions (3.13) and (3.9) introduced above describe the time and angular dependences. For relatively long laser pulses ($\omega_p\tau > 1$), the efficiency of the reflection from a Bragg mirror can be characterized by the total time-averaged cross section, which can be calculated from formula (3.16) by integration over solid angle and time and by dividing the result by the reference-pulse duration τ:

$$\sigma = 8\pi R^2\frac{k_0^2}{k_1^2}\frac{(2\pi)^{3/2}\beta R(k_0L)^2}{\sqrt{R^2\cos^2\alpha + L^2\sin^2\alpha}}\left(\frac{V_E}{4c}\right)^4 \times \exp\left\{-\frac{2(k_0 - k_1\cos\alpha)^2R^2L^2\cos^2\alpha}{R^2\cos^2\alpha + L^2\sin^2\alpha}\right\}, \tag{3.17}$$

where the coefficient β is equal to unity for a sounding wave polarized along the y axis (s polarization) and $\beta = \cos^2(2\alpha)$ if the polarization vector lies in the xz plane (p polarization). The total scattered energy W_S can be easily determined either from expression (3.17) by multiplying it by energy flux density in the sounding wave and by the scattering time τ or from formula (3.11) by integration over solid angle.

CONCLUSION

Below, we make several remarks concerning the possible applications of the effects considered above.

Let us estimate the reflectivity of a Bragg mirror for relatively long laser pulses ($\omega_p\tau > 1$), when no shift of the reflected radiation frequency occurs. Let two laser pulses with an energy of 1 J, duration $\tau = 400$ fs ($L = 120\ \mu$m), frequency $\omega_0 = 2.4 \times 10^{15}\ \text{s}^{-1}$ ($\lambda_0 = 0.8\ \mu$m), and width $R = 100\ \mu$m collide in plasma with electron density $N_{0e} = 1.3 \times 10^{19}\ \text{cm}^{-3}$. The corresponding intensities are 4.5×10^{15} W/cm^2, and the amplitude of the small-scale electron-density perturbations is $(n_2/N_{0e}) \approx 0.3$. For these plasma and pulse parameters, $\omega_p\tau \approx 80 \gg 1$, and quasi-static electron-density perturbations are excited in the collision region. If the sounding wave with s polarization is the second harmonic of the reference waves ($k_1 = 2k_0$), and if the angle of incidence α is 60°, then the total scattered energy at the angle $\theta = \pi - \alpha = 120°$ is

$$W_S = 0.25\frac{E_1^2}{8\pi}V_L,$$

where $V_L = \pi^{3/2}R^2L$ is the volume occupied by the reference pulse; the total scattering cross section is $\sigma = 0.44\pi R^2$. It can be said that 25% of the sounding-wave energy contained in the volume where the scattering occurs is reflected from the Bragg mirror. The frequency of the scattered wave can be tuned by changing the angle of incidence of the sounding wave. This opens up possibilities for using a short-lived Bragg mirror to “cut out” short pulses from longer laser pulses with different frequencies.

Note that, since we assume the electron-density perturbations to be linear [Eq. (1.4)], our analysis is restricted to reference-wave intensities of 10^{15}–10^{16} W/cm^2 at $\lambda_0 \approx 1\ \mu$m. The intensities of currently available lasers can be much higher when the corresponding electron-density perturbations become nonlinear. By measuring the dependence of spectral and angular characteristics of the scattered sounding wave on the intensity of reference laser pulses, we can investigate the transition from linear to nonlinear excited plasma oscillations and their subsequent evolution.

In particular, the small-scale localized coherent plasma oscillations excited by short laser pulses ($\omega_p\tau \le 1$) in the collision region can be an appealing object for studying the destruction of coherence and the growth of strong Langmuir turbulence in laser experiments.

As was already pointed out in the Introduction, the impact excitation of oscillations and waves when short laser pulses with equal frequencies collide is possible not only in plasma but also in other material media. Our analysis leads us to conclude that the impact excitation mechanism is most efficient for oscillations and waves whose period is close to the pulse collision duration.

ACKNOWLEDGMENTS

This study was supported in part by the Russian Foundation for Basic Research (project no. 01-02-16723).

REFERENCES

1. A. G. Litvak, Izv. Vyssh. Uchebn. Zaved., Radiofiz., No. 7, 562 (1964).
2. N. M. Kroll, A. Ron, and N. Rostoker, Phys. Rev. Lett. **13**, 83 (1964).
3. J. Meyer, Phys. Rev. A **6**, 2291 (1972).
4. B. L. Stansfield, R. Nodwell, and J. Meyer, Phys. Rev. Lett. **26**, 1219 (1971); L. A. Godfrey, R. Nodwell, and F. L. Curzon, Phys. Rev. A **20**, 567 (1979); V. G. Zhukovskiĭ, Fiz. Plazmy **3**, 1142 (1977) [Sov. J. Plasma Phys. **3**, 633 (1977)].
5. J. Lavergnat, P. Bauer, J. V. Delahaye, and R. Ney, Geophys. Res. Lett. **4**, 417 (1977); K. B. Dysthe, E. Mjolhus, and J. Trulsen, J. Geophys. Res. **83**, 1985 (1978); L. M. Gorbunov and A. B. Romanov, Fiz. Plazmy **15**, 83 (1989) [Sov. J. Plasma Phys. **15**, 49 (1989)]; P. A. Bernhardt, L. S. Wagner, J. A. Goldstein, *et al.*, Phys. Rev. Lett. **72**, 2879 (1994).
6. C. Joschi, W. B. Mori, T. Katsouleas, *et al.*, Nature **311**, 525 (1984); B. Amini, Phys. Rev. Lett. **54**, 1163 (1985).
7. J. A. Heikkinen, S. J. Karttunen, and R. R. Salomaa, Nucl. Fusion **28**, 1845 (1988).
8. E. Esarey, P. Sprangle, J. Krall, and A. Ting, IEEE Trans. Plasma Sci. **24**, 252 (1996).
9. J. Salcedo and A. E. Siegman, Phys. Rev. Lett. **41**, 131 (1978).
10. S. A. Akhmanov and N. I. Koroteev, in *Methods of Nonlinear Optics in Light Scattering Spectroscopy* (Nauka, Moscow, 1981), p. 470.
11. L. M. Gorbunov and V. I. Kirsanov, Zh. Éksp. Teor. Fiz. **93**, 509 (1987) [Sov. Phys. JETP **66**, 290 (1987)].
12. V. P. Silin and A. A. Rukhadze, in *Electromagnetic Properties of Plasma and Plasmalike Media* (Gosatomizdat, Moscow, 1961), p. 90.
13. N. E. Andreev, L. M. Gorbunov, V. I. Kirsanov, *et al.*, Phys. Scr. **49**, 101 (1994).
14. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
15. L. M. Gorbunov and A. A. Frolov, Zh. Éksp. Teor. Fiz. **110**, 1757 (1996) [JETP **83**, 967 (1996)].

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 519–532.
Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 593–608.
Original Russian Text Copyright © 2001 by Kuznetsov, Davydova.

Detonation of Explosives Containing Heavy Inert Particles

N. M. Kuznetsov[a, *] **and O. N. Davydova**[b]

[a]*Semenov Institute of Chemical Physics, Russian Academy of Sciences, Moscow, 117334 Russia*
[b]*Institute for Chemical Physics Problems, Russian Academy of Sciences,*
p/o Chernogolovka, Moscow oblast, 142423 Russia
**e-mail: kuznetzv@chph.ras.ru*

Received March 6, 2001

Abstract—The physical and mathematical aspects of the theory of a detonation wave containing heavy inert particles are considered. The detonation wave intensity and structure are determined by the relaxation of velocities of both the reactive explosive and the inert admixture. The generalized Jouguet condition is formulated for the velocity of a self-sustained detonation wave. The results of analytical treatment and the model numerical solutions of the problem of the detonation wave velocity selection and the wave structure determination are presented as a function of the ratio of the characteristic times of the heat evolution and the two-component flow velocity relaxation. A limiting case of the fast particle drag is represented by the shock wave structure determined by relaxation of the two-component flow velocity. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

According to the classical detonation theory of Chapman–Jouguet and Zeldovich–Neiman–Dering (ZND), the parameters of a self-sustained detonation regime are determined on the pressure–specific volume (p, V) plane by the coordinates of a tangency point of the Michelson–Rayleigh line (MRL) and a detonation adiabat of the final reaction products (the Jouguet condition) [1, 2]. This regime is frequently referred to as “normal detonation,” which is explained by the fact that a usual adiabat of the final reaction products is lying on the (p, V) plane above all intermediate adiabats corresponding to incomplete heat evolution. The above tangency condition also refers to such a case. If the position of a detonation adiabat in the course of the heat evolution changes in a nonmonotonic manner (which may be caused, for example, by a nonmonotonic heat evolution), the tangency condition still remains valid but refers (instead of the equilibrium adiabat) to an intermediate detonation adiabat lying on the (p, V) plane above all other adiabats (Fig. 1) [2–4]. In this case, unlike the normal detonation regime, the terminal point of the wave relaxation zone does not coincide with the MRL–adiabat tangency point, but is situated on the equilibrium detonation adiabat at a lower pressure and smaller compression (Fig. 1, point *4*). Detonation with such a reduced compression is naturally referred to as undercompressed detonation [4]. For the same final heat evolution (corresponding to a transition to the thermodynamically equilibrium state), the velocity of an undercompressed detonation is higher than that of the normal detonation.

The case of a self-sustained undercompressed detonation, which was originally studied in connection to the possible nonmonotonic heat evolution during chemical reactions [3], can be also controlled by some other factors of nonmonotonic heat evolution, for example, by sufficiently slow (compared to exothermal chemical reactions) endothermal relaxation processes such as heating and fusion of inert additives [5–9].

Another possible reason for an undercompressed detonation development is a slow relaxation of the velocity of heavy particles of the inert admixtures in blend systems (for brevity, this case will be hereinafter referred to as detonation with relaxing velocities) [7–10]. The self-sustained detonation velocity in such systems depends on the ratios of the characteristic times of relaxation of the temperature (τ_T) and of the impurity particle velocity (τ_U) to the characteristic time of the heat evolution (τ_q). For $\tau_T/\tau_q \ll 1$ and $\tau_U/\tau_q \ll 1$, the system features a normal detonation at a velocity d_J determined by the Jouguet condition. The combination of inequalities $\tau_T/\tau_q > 1$, $\tau_U/\tau_q \ll 1$ correspond to an undercompressed detonation controlled by nonmonotonic heat evolution. A more complicated variant of the undercompressed detonation is represented by the condition $\tau_U/\tau_q > 1$. In this case, the component velocities in a relaxation zone of the detonation wave (more exactly, in that part of the zone where maximum heat evolution takes place) are not equal and a change in the state of the relaxing substance in the detonation wave is no longer described by the MRL. As a result, the above simple and illustrative interpretation of the detonation velocity selection using the MRL becomes inapplicable.

Approximate calculations of the relationship between the particle velocity relaxation and the observed detonation velocity for the experiments [7] with an explosive blend containing a variable amount of a tungsten powder were performed in [8, 11]. Al’tshuler *et al.* [11] estimated the velocity of impurity (tungsten) particles for each experimental variant [7] at a certain conditional time instant corresponding to a virtually complete heat evolution. Davydova *et al.* [10] performed calculations of the detonation wave velocity depending on the τ_U/τ_q ratio using the principle of minimum detonation velocity.

For consistently solving the problem of determination of the parameters of an undercompressed detonation with relaxing velocities, it is necessary (i) to establish the physical and gasdynamic factors determining the selection of a certain velocity of the self-sustained detonation from a set of possible velocities and (ii) to find a mathematical expression for this selection, that is, to formulate an analog of the Jouguet condition in the case under consideration.

This study aimed at answering the above questions and solving the problems by carrying out a mathematical analysis of the relationship between the velocity and structure of the detonation wave and the τ_U/τ_q ratio. On this mathematical basis, we developed a method for the numerical calculation of the velocity and structure of the detonation wave and applied this formalism to the model equations of state and heat evolution kinetics. Our first purpose is to establish the qualitative laws independent (or weakly dependent) on special features of the equation of state of the explosive and the detonation products. For this formulation of the problem, we may use a simple model equation of state for a condensed explosive and the corresponding products, in particular, a common equation of state of the ideal gas with a large adiabatic exponent γ and a large initial density (~1 g/cm^3). The problem parameters, such as the mass concentration of impurity particles and the ratio $K \equiv \tau_U/\tau_q$, can be varied in a broad range without restrictions corresponding to certain experimental conditions.

In Section 2, we will consider the features of the problem of determining the structure of a self-sustained detonation wave and formulate a generalized principle of minimum detonation velocity, known in detonation theory, for the velocity of a self-sustained detonation with relaxing velocities. Sections 3 and 4 are devoted to mathematical formulation of the problem and qualitative analysis of the solution. Sections 5 and 6 present the results of an analytical treatment and the model numerical calculations of the velocity and structure of a detonation wave as functions of the parameter *K*. In Section 7, we will consider the structure of a shock wave in the initial explosive (or in the chemically inert substance) containing an admixture of heavy particles. This section supplements Section 5 by providing an illustrative interpretation of the structure of a detonation wave with small *K*. In addition, Section 7 may present independent interest from the standpoint of the shock wave structure. A part of mathematical transformations and a study of the relationship between the boundary of real solutions of a differential equation for the impurity particle velocity relaxation and the saddle-type singular point are presented in the Appendices.

2. QUALITATIVE FEATURES OF A SELF-SUSTAINED DETONATION WAVE AND THE PRINCIPLE OF MINIMUM DETONATION VELOCITY

In this section we will qualitatively analyze the properties and structure of a detonation wave based on the well-known laws of the theory of shock waves and detonation. Upon a sufficiently powerful initiation, the detonation wave occurs at the first instant in an overcompressed state, while a flow behind the wave is subsonic.[1] The front edge of a wave of rarefaction linking the shock wave to the flow far behind (where the density of substance under the conditions of particular problems is several times or several orders of magnitude lower than that in the detonation wave) coincides with the shock wave front of the detonation complex and leads gradually to a decrease in its intensity. In order to exclude this mechanism leading to a decrease in the wave velocity *d*, it is necessary to provide that the stationary relaxation wave zone would propagate in the substance at a velocity not smaller than the local equilibrium sound velocity *c* at the rear boundary of the zone:

$$d - U \geq c, \tag{1}$$

where *U* is the velocity of substance at the rear boundary of the zone in the laboratory frame.

A phase trajectory describing the structure of the stationary relaxation wave zone in the system with relaxing velocities does not coincide with the MRL, but the end point of the trajectory (corresponding to the boundary of the relaxation zone) occurs at the intersection of the detonation adiabat with the MRL. For $d > d_J$, there are two such intersections for each MRL (see Fig. 1). However, a physical problem must possess a single solution. A principal qualitative difference between points *3*, *3*' and *4*, *4*' (Fig. 1) is that the latter pair obeys inequality (1), while the former pair satisfies the reverse condition. In other words, only points *4*, *4*' are supersonic, and, hence, the relaxation zone of a nondecaying (self-sustained) wave must terminate at point *4* [2, 4]. Since condition (1) is quantitatively uncertain, this relationship cannot be used to determine the velocity of the self-sustained detonation. For this purpose, an additional physical relationship and the corresponding mathematical expression are necessary. In the case of a normal detonation, this is provided by the Jouguet condition (detonation adiabat–MRL tangency).

For determining an analogous condition, let us turn to a mental experiment with an overcompressed detonation wave maintained by a piston. Starting with a

---

[1] The leading front of the detonation wave, where the heat evolution is initially absent or relatively small, represents a shock wave followed (as is well known) by a subsonic flow.

strong overcompression, we will sufficiently slowly (so as not to violate stationary or quasi-stationary character of the flow) decrease the piston velocity. As a result, the degree of overcompression decreases and the relaxation zone boundary (Fig. 1, point *3'*) moves down along the detonation adiabat. In the course of this process, however, the detonation velocity cannot become lower than the velocity d^* of the self-sustained detonation wave (which can exist without any piston). Obviously, the effect of the piston on the wave velocity would cease at a piston velocity U_p^* corresponding to the detonation wave velocity decreased to $d = d^*$. From the standpoint of gasdynamics, termination of the effect of the piston on the wave velocity implies that a sound point appears in the relaxation zone, through which perturbations originating from the piston cannot pass to the shock wave front. At a piston velocity U_p^*, the stationary relaxation zone terminates at the sound point and the subsequent transition to thermodynamic equilibrium (including accomplished velocity relaxation) proceeds via a nonstationary gasdynamic process relating the sound point to the boundary condition at the piston (preset velocity). When the piston velocity decreases further, the detonation wave velocity and the structure of the stationary relaxation zone between the shock wave front and sound point remain unchanged, but the length of this zone grows as the boundary shifts away from the wave front behind the sound point (Fig. 2). Finally, when the piston velocity decreases down to the value corresponding to the equilibrium stationary flow at point *4* (Fig. 1), the boundary of the stationary relaxation zone coincides with the equilibrium point *4*. Thus, the relaxation in the stationary zone is completed. The region between sound point and the piston (with the length uniformly increasing with time) represents constant equilibrium flow with the same parameters (velocity, pressure, density) as in point *4* [9]. Further decrease in the piston velocity (even stopping or reverse motion) affects neither the detonation wave velocity nor the relaxation zone (this behavior is similar to the case of a normal detonation initiated at a tube edge, whereby the detonation velocity far from the edge does not depend on whether the tube is open or closed [2, 4]).

For a given detonation velocity (e.g., $d = d^*$), the problem of determining the structure of the stationary wave zone represents, from the mathematical standpoint, a Cauchy problem for a system of equations with the initial conditions on the wave front. A solution to this problem is represented by an integral curve independent of the piston velocity ($U_p < U_p^*$). This velocity only determines whether the curve reaches the equilibrium point *4* (Fig. 1) or the stationary zone terminates before this point. Anyhow, the curve does not lead to point *3* (in contrast to the integral curves corresponding to $d > d^*$). Therefore, the relaxation wave structure considered as a function of the wave velocity d exhibits a qualitative change at the point $d = d^*$. Indeed, the boundary of the relaxation zone terminates on the upper branch of the detonation adiabat for $d > d^*$ and on the lower branch (at the lower intersection of the detonation adiabat and MRL) for $d = d^*$.

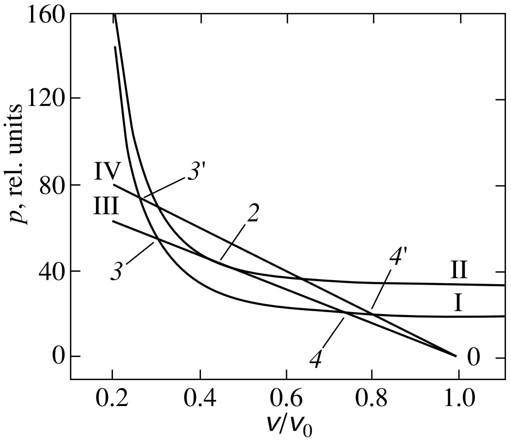


Fig. 1. Detonation adiabats: (I) equilibrium detonation adiabat; (II) detonation adiabat of maximum heat evolution; (III, IV) Michelson–Rayleigh lines (MRLs); 0 is the point corresponding to the state of substance in front of the wave; (*2*) tangency point; (*3*, *3*' and *4*, *4*') upper and lower intersection points between MRL and the equilibrium detonation adiabat.

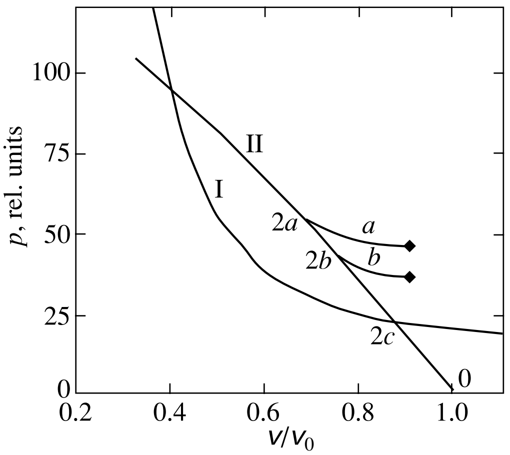


Fig. 2. A qualitative relationship between the parameters of a nonstationary flow between boundaries of the stationary zone (points *2a*, *2b*, and *2c* in the generalized Michelson–Rayleigh line II) and the corresponding boundaries corresponding to termination of the relaxation processes (points ◆) for three values of the piston velocity $U = U_a > U_b > U_c$ ($U_a = U^*$, $U_c = U_{\min}$). The piston pressures p are the same as at points ◆; I is the equilibrium detonation adiabat; 0 is the point corresponding to the state of substance in front of the wave.

As is known from the theory of differential equations, this qualitative change takes place when the integral curve becomes a separatrix of the saddle-type singular point. Thus, the problem of determining the velocity of a self-sustained detonation can be reduced to the problem of finding a wave velocity $d = d^*$ such that the wave structure is described by a separatrix converting the initial subsonic flow into a supersonic one in a stationary relaxation zone. This concept, established in the classical detonation theory for single-component systems (see, e.g. [4]), is equally applicable to the case of detonation with relaxing velocities. The quantity d^* is related to the undercompressed detonation velocity by the relationship

$$d^* < d. \tag{2}$$

Besides this method of determining the velocity of a self-sustained detonation, it is interesting from both the theoretical and the practical standpoint to approach the problem in a different way. An alternative is suggested by the condition of existence of a stationary wave structure in which the explosive passes from the initial nonequilibrium state to the state of detonation products in complete thermodynamic equilibrium, including equal velocities of the directed motion of all components. For a fixed finite heat of explosion q, there exists a certain minimum value of the wave velocity $d_{\min 1}$ (exceeding the sound velocity in the initial explosive) below which no real solutions exist for the stationary wave structure. This is well known from the classical ZND theory and can be proved for any relationship between the characteristic times τ_T, τ_q, and τ_U.

According to the definition of $d_{\min 1}$, detonation waves with a stable one-dimensional structure satisfy the condition

$$d^* \geq d_{\min 1}.$$

The fact that the velocities d^* and $d_{\min 1}$ possess certain minimum properties (see condition (2)) suggests that, in the case of a unique self-sustained detonation regime, these velocities are equal (so that we only deal with different notations of the velocity of the same detonation wave). This suggestion is known to be valid both for the normal Jouguet detonation and for the detonation with nonmonotonic heat evolution. Validity of this suggestion in the general case of undercompressed detonation is confirmed in Appendix 1, where it is proved that for $d = d_{\min 1}$ the integral curve describing a stationary structure of the wave with relaxation of the component velocity is a separatrix passing through a saddle-type singular point.

The like properties of the detonation regimes and the corresponding solutions of the problem of the stationary wave structure description for $d = d^*$ and $d = d_{\min 1}$ indicate that, under the condition of uniqueness of the $d_{\min 1}$ value,[2]

$$d^* = d_{\min 1}. \tag{3}$$

In the space of complex variables, a solution to the problem of the stationary wave structure does not vanish at $d < d_{\min 1}$, but becomes a complex quantity, not satisfying a physical formulation of the problem. Accordingly, the value of $d_{\min 1}$ can be found based on the condition of tangency between the integral curve (describing a stationary detonation wave structure) and the boundary of the domain of real solutions. This tangency condition will be used below (Section 4) in formulating the necessary generalized Jouguet condition.

3. MATHEMATICAL FORMULATION OF THE PROBLEM

The detonation wave structure, calculated with an allowance for both the heat evolution and the relaxation of velocities of the impurity particles, is described by a system of five equations of a stationary two-component flow (including two equations reflecting conservation of the mass of components, the equations of conservation of momentum and energy, and a differential equation for the velocity relaxation of impurity particles) with five unknowns (density and velocity of each component and the pressure).

To write these equations, we introduce the following notation: d, the detonation wave velocity in the laboratory frame of reference; x and y, velocities of the gas and impurity particles behind the shock wave front of the detonation complex in the frame of the front; α and β, mass concentrations of the gas and impurity in the initial blend ($\beta = 1 - \alpha$); R_{ex} and r_p, densities of the explosive component and impurity in the initial state; R_0 and r_0, densities (mass to volume ratios) of the gas and impurity in the initial mixture) defined as

$$R_0 = \frac{\alpha}{v_0}, \quad r_0 = \frac{\beta}{v_0};$$

R and r, densities of the gas and impurity behind the shock wave front; v_0, specific volume of the initial blend related to the densities of individual components as

$$v_0 = \frac{\alpha}{R_{ex}} + \frac{\beta}{r_p};$$

C_1 and C_2, gas and impurity fluxes; p and v, pressure and specific volume of the blend behind the shock wave front; p_0, pressure in front of the shock wave; H_1 and H_2, specific enthalpies (without latent heat of explosion) of the gas and impurity behind the shock wave front; $H_{1,0}$ and $H_{2,0}$, specific enthalpies (without latent heat of explosion) of the gas and impurity before the shock wave front; q_0, specific heat of explosion of the initial explosive; q, current evolved heat (i.e., the amount of heat evolved by a given time instant).

---

[2] A nonunique detonation regime in an infinite medium is possible, but this poses some additional conditions on the relaxation parameters (concerning the kinetics of the nonmonotonic heat evolution [5] or a special type of dependence of the coefficient of friction between the system components during the velocity relaxation on the detonation wave velocity).

The first four equations, expressing the conservation laws for a stationary two-component flow, can be written as follows:

$$xR = C_1, \tag{4}$$

$$yr = C_2, \tag{5}$$

$$p + C_1 x + C_2 y = C_3, \tag{6}$$

$$C_1(H_1 - q) + C_2 H_2 + C_1\frac{x^2}{2} + C_2\frac{y^2}{2} = C_4, \tag{7}$$

where C_1, C_2, C_3, and C_4 are constants depending on the detonation wave velocity and on the initial blend composition and state:

$$\begin{gathered} C_1 = R_0 d, \quad C_2 = r_0 d, \quad C_3 = p_0 + (C_1 + C_2)d, \\ C_4 = C_1 H_{1,0} + C_2 H_{2,0} + (C_1 + C_2)\frac{d^2}{2}. \end{gathered} \tag{8}$$

Below, we will consider only strong detonation waves, for which the terms involving p_0, $H_{1,0}$, and $H_{2,0}$ in Eqs. (4)–(8) can be omitted. In addition, we make the following simplifying assumptions not essentially affecting the solution of the problem under consideration. We assume that the impurity particles are incompressible and their thermal relaxation times are either infinitely large or negligibly small (in these limiting cases, the kinetics of particle heating can be ignored and the difference between the cases reduces to a relatively small heat change q_0). Using these simplifying assumptions, substituting an expression for the enthalpy of the ideal gas

$$H_1 = \frac{Sp}{R}, \quad S \equiv \frac{\gamma}{\gamma - 1}, \tag{9}$$

into Eq. (7), expressing p according to Eq. (6), and accomplishing simple transformations, we arrive at the following equation:

$$\frac{d^2}{p} = [d - (\alpha x + \beta y)]Sx - \alpha q + \alpha\frac{x^2}{2} + \beta\frac{y^2}{2}. \tag{10}$$

This equation determines, for a given wave velocity d and the heat evolution law $q = q(t)$, a time-dependent relationship between the component velocities in the relaxation zone of the detonation wave. The system is closed, allowing the component velocities x, y and the other unknowns (R, r, p) to be determined by a relaxation equation for the particle acceleration:

$$m\frac{dy}{dt} = F, \tag{11}$$

where F is a force depending on the difference of component velocities and m is the particle mass. The form of Eq. (11) may vary depending on the Reynolds number Re. For Re $\le$ 10, the system features a Stokes streamlining, whereby the force is a linear function of the relative velocity: $F \propto x - y$; for Re $\gg$ 10, the relaxation equation becomes more complicated. This behavior can be approximately described by a sum of linear and nonlinear terms [12, 13]

$$\begin{gathered} F \approx \frac{1}{8}\pi A \rho \nu d_p (x - y)\left[1 + \frac{B}{A}(\mathrm{Re})^n\right], \\ \frac{1}{3} \le n \le \frac{2}{3}, \quad \mathrm{Re} \equiv |x - y|\frac{d_p}{\nu}, \end{gathered} \tag{12}$$

where ρ and ν are the density and kinematic viscosity of the detonation products, respectively, and d_p is the particle diameter. In comparison to the velocity difference, the density can be considered as a constant and set equal to the value in the beginning or at the end of the relaxation zone. The coefficients A and B in Eq. (12) depend on n: for $n = 1/3$, $A = 24$ and $B = 5.9$. The nonlinear term in Eq. (12) does not qualitatively change the integral curves of the system of equations under consideration (this was previously established for a particular case, where numerical calculations were performed for an analogous system comprising a water-saturated hexogen blend with a tungsten powder [10]). This term should be taken into account mostly in the quantitative calculations of the detonation parameters for particular explosive blends using sufficiently precise equations of state of the explosive and detonation products. Our analysis will be restricted to a linear approximation with respect to the force F.

Equation (11) has to be solved with the initial condition

$$y = d|_{t=0}, \tag{13}$$

which implies that the velocity of particles exhibits no break at the shock wave.

The heat evolution law, that is, the quantity q as a function of time, is set in the following form:

$$q = q_0\left[1 - \exp\left(-\frac{t}{\tau_q}\right)\right], \tag{14}$$

where q_0 is the total evolved heat (heat of explosion). For the convenient comparison of the characteristic times of the heat evolution and the velocity relaxation, the relaxation equation obtained upon substituting an expression for F from (12) into (11) in the linear approximation ($B = 0$) can be transformed as follows:

$$\begin{gathered} \frac{dy}{dT} = \frac{x - y}{K}, \\ T \equiv \frac{t}{\tau_q}, \quad K \equiv \frac{\tau_U}{\tau_q}, \quad \frac{1}{\tau_U} \equiv \frac{\pi A \rho \nu d_p}{8m}. \end{gathered} \tag{15}$$

For the exponential heat evolution (14), we obtain

$$\frac{dq}{dT} = q_0 - q.$$

Using this equation and formula (14) to exclude the dimensionless time T, we arrive at the differential equation describing the dependence of y on q:

$$\frac{dy}{dq} = \frac{x-y}{K(q_0-q)}. \tag{16}$$

The initial condition (13) is expressed as

$$y = d|_{q=0}. \tag{17}$$

4. THE GENERAL SCHEME AND A PRACTICAL METHOD FOR SOLVING THE PROBLEM OF DETERMINING THE PARAMETERS AND STRUCTURE OF A SELF-SUSTAINED DETONATION WAVE

Depending on the parameter d (detonation wave velocity), the system of Eqs. (10) and (16) describes both the self-sustained detonation and the overcompressed detonation waves which, similar to the shock waves, can propagate with a constant amplitude only in the presence of an appropriate external action (e.g., under the action of a piston). Otherwise, the wave velocity decreases to a level of the self-sustained detonation d^*. The calculation of d^* according to Eq. (3) reduces to determining the value $d_{\min1}$ below which no real solutions exist for the stationary detonation wave structure. This value depends on the parameter K of the relaxation equation (16). Determining the velocity of a self-sustained undercompressed detonation wave is a kind of eigenvalue problem: among the manifold of solutions describing the wave structure, we have to find one solution corresponding to a preset value of the parameter K, such that $d = d_{\min1}(K)$, which is described in the variables (q, x) by a separatrix passing through a saddle-type singular point (see Section 2). For brevity, $d_{\min1}(K)$ will be written below as $d_{\min1}$.

For the real initial conditions, the boundary of the real solutions (integral curves) for the system of Eqs. (10) and (16), (in which the detonation velocity d is a parameter) satisfying the initial condition (17) and leading to a local thermodynamic equilibrium

$$q = q_0, \quad y = x, \tag{18}$$

is determined by two surfaces in a three-dimensional space (y, q, d). One of these surfaces corresponds to the requirement for the integral curves to be real, while the other surface reflects the condition of existence of the points of equilibrium (18) for a given value of the parameter d. Let us consider these surfaces.

1. The first surface is a boundary of the real values of the function $x = x(y, q, d)$ determined by Eq. (10). This function (a solution to Eq. (10) with respect to x) is considered in Appendix 2. Equating the determinant of the solution to zero, we obtain an equation for this boundary:

$$Z(y, q, d) = 0, \tag{19}$$

which is referred to below as the boundary surface. An explicit expression for the function $Z(y, q, d)$ and the equation for the boundary surface in the form

$$d = F(y, q) \tag{20}$$

are also presented in Appendix 2, where it is shown that the domain of real solutions satisfies the condition

$$d - F(y, q) \geq 0. \tag{21}$$

This domain contains the real integral curves of the system of Eqs. (10) and (16), which correspond to various values of the parameters d and K such that

$$f(y, q)_{d,K} = 0. \tag{22}$$

Consider a three-dimensional space with the Cartesian coordinates (y, q, d), where the d axis is vertical. Then condition (21) implies that the integral curves of the system (10) and (16), which by definition are lying in the planes d = const, must not fall below the boundary surface (20). A minimum value of the velocity d (denoted by $d_{\min1}$) among all values satisfying condition (21) corresponds to tangency between the integral curve and the boundary surface. The problem of determining $d_{\min1}$ will be considered below.

2. The second surface is the plane $d = d_{\min2}$ dividing the (y, q, d) space (i.e., the manifold of planes d = const) into two parts. In one of these, each d value corresponds to at least one point (y, q_0) such that equilibrium (18) is a solution to Eq. (10). In the other part, no such points exist for any d value. According to the classical detonation theory, the manifold of points (18) corresponds to an equilibrium detonation adiabat on the (p, V) plane and there are two points (18) satisfying Eq. (10) for each value of $d > d_J$. These points are determined by two intersections of the detonation adiabat with the MRL. For $d = d_J$, the two intersection points merge at a tangency point. For $d < d_J$, the MRL has no common points with the detonation adiabat; that is, Eq. (10) has no real solutions (18). From this and the above definition of the plane $d = d_{\min2}$, we infer that

$$d_{\min2} \equiv d_J. \tag{23}$$

The $d_{\min2}$ value is determined from the equation

$$\mathrm{Det}(d, q_0) = 0, \tag{24}$$

where $\mathrm{Det}(d, q_0)$ is the determinant of Eq. (10) with respect to x under condition (18), that is, of the equation

$$\frac{d^2}{2} - (d-x)Sx + q_0\alpha x - \frac{x^2}{2} = 0.$$

Solving this equation with respect to d yields a well-known formula of detonation theory:

$$d_{\min2} \equiv d_J = \sqrt{2(\gamma^2-1)\alpha q_0}, \tag{25}$$

where α is the factor taking into account a decrease in the specific heat of explosion related to the presence of an inert admixture.

Note that the variable y disappears from Eq. (20) for $\beta = 0$ (see the explicit form of $F(y, q)$ in Appendix 2) and substituting $q = q_0$ into this equation yields, as expected, formula (25). Figure 3 gives a qualitative illustration of the relative arrangement of the surface (20) and the plane (23), bounding the domain of real solutions to Eqs. (10) and (16), with the integral curves containing equilibrium points (18). Figure 4 shows sections of the boundary surface (20) by the planes d = const (curves y_1–y_4 are discussed in Section 5).

Since the detonation velocity cannot be lower than $d_{\min 1}$ and d_J, a condition of minimum for the detonation velocity $d_{\min}$ can be written as

$$d_{\min} = \max\{d_{\min 1}, d_J\}. \tag{26}$$

This is the generalized Jouguet condition. The first quantity compared in (26) depends on (and is a continuous function of) the parameter K in Eq. (16), while the second quantity d_J is independent of K. In the limit of $K \longrightarrow 0$, which corresponds to the classical Jouguet detonation, we have $d_{\min} = d_J$. Therefore, according to condition (26), this limit corresponds to $d_{\min 1} < d_J$ (the other possibility $d_{\min 1} = d_J$ admitted by this condition is realized only in the limit of $\beta = 0$). For $K \longrightarrow \infty$, which corresponds to the most pronounced case of undercompressed detonation, we have $d_{\min} > d_J$ and, according to (26), $d_{\min 1} > d_J$. The difference between $d_{\min 1}$ and d_J in the two limits implies that the sense of the inequality between $d_{\min 1}$ and d_J changes when the parameter K passes through a certain intermediate finite value K^*. In other words, the normal Jouguet detonation under the exact condition $d_{\min 1} = d_J$ corresponds to the limit $K = 0$ and to a certain finite interval

$$0 \le K \le K^*. \tag{27}$$

A physical meaning of this relationship is considered in Section 4. For finding the detonation velocity $d_{\min}$ corresponding to (26) in the entire range of K, we must first calculate d_J and K^* and then determine the dependence of $d_{\min 1}$ on K for $K > K^*$.

Now let us derive an equation determining $d_{\min 1}$ and consider a method for the practical calculation of K^* and $d_{\min 1}$. As pointed out above, the integral curve for $d = d_{\min 1}$ touches the boundary surface (20), that is, has a single common point with this surface (below, such curves will be referred to as "intrinsic"). This point is the point of tangency between two curves lying in the plane $d = d_{\min 1}(K)$. One is the intrinsic integral curve

$$f(y, q)_{d_{\min 1}} = 0, \tag{28}$$

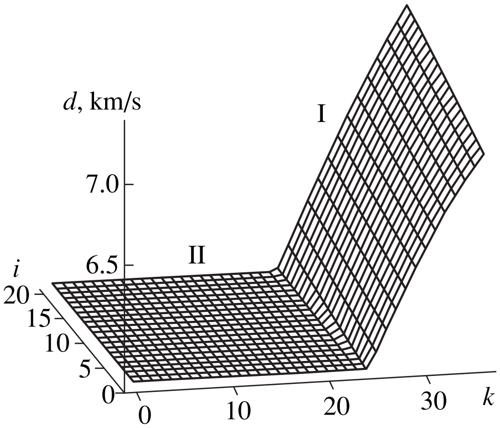


Fig. 3. Mutual arrangement of (I) the boundary surface (20) and (II) the plane $d = d_J = 5.759$ km/s ($q = 3 + 0.027i$ kJ/g; $y = 0.2k$ km/s).

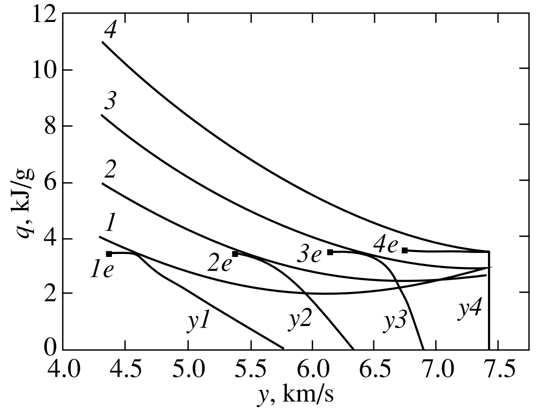


Fig. 4. Cross sections of the boundary surface (20) by the planes d = const = 5.759 (*1*), 6.33 (*2*), 6.9 (*3*), 7.413 km/s (*4*). Points *1e*–*4e* denote the boundaries of the corresponding stationary zones in the self-sustained detonation wave structure (point *4* in Fig. 1); $q_0 = 3.455$ kJ/g; $\beta = 0.4$; $\gamma = 3$; *y1*–*y4* are the corresponding intrinsic integral curves.

while the other is the section of the boundary surface (20) by the plane $d = d_{\min 1}$:

$$F(y, q) = d_{\min 1}. \tag{29}$$

The condition of tangency is expressed by the equality of two derivatives dy/dq at the same point with the coordinates (y, q, $d_{\min 1}$). One of these derivatives is taken along the integral curve (28), and the other, along the line (29). The equality is written as

$$\left(\frac{\partial f}{\partial q}\right)_y \Big/ \left(\frac{\partial f}{\partial y}\right)_q = \left(\frac{\partial F}{\partial q}\right)_y \Big/ \left(\frac{\partial F}{\partial y}\right)_q \Bigg|_{d = d_{\min 1}}. \tag{30}$$

The three equations (28)–(30) determine three unknowns, the y and q values at the tangency point and $d_{\min 1}$, for a given value of the parameter K. Once $d_{\min 1}$ is known, the value of $d_{\min}$ is determined from (26).

Formally, the procedure of determining the intrinsic value of $d_{\min 1}$ and the corresponding intrinsic curve (28) consists in selecting a function f in the manifold (22) with a d value satisfying the condition of tangency (30). When the integral curve (28) is found (or in the course of the calculation), we determine x as a function of q (or y) from Eqs. (28) and (10), which represents the detonation wave structure in these variables. Then the wave structure can be expressed in the (p, V) variables using the formulas (see Eqs. (4)–(9))

$$p = C_3 - C_1 x - C_2 y, \quad V = \frac{xy}{C_1 y + C_2 x}.$$

The above scheme for determining the detonation wave velocity $d(K)$ and the intrinsic integral curve (28) gives a general notion of the mathematical structure of the solution and offers a geometric interpretation. However, this pathway is rather difficult for practical implementation. In practice, numerical solution of the problem for a given $K > K^*$ is significantly facilitated if, instead of searching for the tangency point according to Eq. (30), one performs the equivalent search for a minimum value of the parameter d in Eq. (10), such that a smallest (to within a preset accuracy) decrease in this value transfers a solution to Eqs. (10) and (16) into the complex plane. The point $q = q^*$ corresponding to the appearance (disappearance) of this complex character for a smallest decrease (increase) in the parameter d coincides (to within a preset accuracy) with the required tangency point. By way of such calculations for a series of $K = K_i$ values, we determine the function $d_{\min 1}(K_i)$. Owing to a monotonic character of this function and a weak dependence on the argument (when K grows from K^* to ∞, the velocity $d_{\min 1}$ increases by a factor of $1/\sqrt{\alpha}$ for an ideal gas and approximately to the same extent for any other equation of state), it is sufficient to perform these calculations for several (about ten) K_i values for restoring the continuous function $d_{\min 1}(K)$ with a satisfactory accuracy. In practice, the calculations are conveniently carried out by exchanging argument and function, that is, by setting the values $d = d_i$ at a constant step in the entire range (from d_J to $\approx d_J/\sqrt{\alpha}$) and determining the corresponding maximum values of $K = K_i$ from Eq. (16) such that a solution to Eqs. (10) and (16) becomes complex after even a smallest (to within a preset accuracy) increase in K. The K^* value is determined as a maximum of K for $d = d_J$. The results of numerical calculations are presented in Section 6.

5. DEPENDENCE OF THE TANGENCY POINT COORDINATES ON THE RATIO (K) OF THE CHARACTERISTIC TIMES OF RELAXATION OF THE PARTICLE VELOCITY AND HEAT EVOLUTION

Since the system of Eqs. (10), (16), and (30) cannot be solved analytically, dependence of the tangency point coordinates on the parameter K cannot be studied analytically either. However, a qualitative notion about this dependence can be obtained based on the analysis of asymptotics for small and large K values.

a. The case of $K \longrightarrow 0$ corresponds to the normal (classical) detonation, whereby the velocity relaxation terminates at $q = 0$ and the heat evolution stage corresponds to $x = y$. In this limit, the detonation velocity is given by formula (25).

In Section 4, where a qualitative analysis of the velocity of a self-sustained detonation as a function of the parameter K was performed and the limiting values were calculated, it was demonstrated that the classical expression for the detonation velocity is valid not only in the limit $K \longrightarrow 0$, but within a certain finite interval (27) of small K values as well. Let us consider the physical reasons for this behavior of the detonation wave.

At a finite but small value of the parameter K, the relaxation of the component velocities is comparatively fast and the main stage of the heat evolution proceeds under the conditions of approximately equal velocities of the explosive blend components, which can be expressed by the relationships

$$|\Delta| \ll x, y, \quad \Delta \equiv y - x.$$

In addition,

$$\frac{dy}{dq} \approx \frac{dy_e}{dq} \equiv \frac{dx_e}{dq},$$

where x_e (and y_e) is the flow velocity for the equal component velocities and the quantity Δ is determined by the relaxation equation (16) with dy/dq in the left-hand part replaced by dx_e/dq:

$$\Delta = -K(q_0 - q)\frac{dx_e}{dq}. \tag{31}$$

The dependence of x_e on q for a given wave velocity d, determined by Eqs. (10) and the condition $x = y$, can be expressed as

$$x_e = \frac{1}{\gamma + 1}[\gamma d - \sqrt{d^2 - 2(\gamma^2 - 1)\alpha q}].$$

Substituting $d = d_J$ (see Eq. (25)), we obtain

$$x_e = \frac{1}{\gamma + 1}\sqrt{2(\gamma^2 - 1)\alpha q_0}\left(\gamma - \sqrt{1 - \frac{q}{q_0}}\right). \tag{32}$$

Differentiating (32) with respect to q and substituting the result into (31), we find

$$\Delta = -K\sqrt{\frac{\gamma - 1}{\gamma + 1}}\sqrt{2\alpha(q_0 - q)}. \tag{33}$$

In considering a dependence of the self-sustained detonation velocity on the small parameter K, a key point is that the difference of the component velocities Δ is negative in the main and final stage of heat evolution, which is qualitatively evident in (31) and is quantitatively expressed by formula (33) (see also the results of numerical calculations in Section 6 and in Fig. 6a). The negative value of Δ indicates that the impurity particles move faster than the detonation products in the laboratory frame of reference. Therefore, the subsequent relaxation of velocities is accompanied by redistribution of the kinetic energy from impurity particles to detonation products, so that the latter obtain an additional energy besides the evolved heat. Since the signs of the energy fluxes supplied to the detonation products due to the heat evolution and the relative retardation of particles are the same (positive), the transferred energy reaches a maximum upon completion of the heat evolution and the velocity relaxation. In the (p, V) plane, the corresponding point is situated on the equilibrium detonation adiabat. Appendix 3 provides a proof of the fact that this point is determined by the classical MRL tangency condition. In a finite interval (27) of small K values, the self-sustained detonation velocity is independent of K and determined by formula (25).

b. For $K = \infty$, the velocity of impurity particles during the heat evolution does not change, remaining equal to the initial value $y = d$. In this limit (corresponding to a zero viscosity ν), the $d_{\min 1}$ value is determined, according to Eq. (20), from an equation $d = F(d, q_0)$. A solution to this equation has the form

$$d_{\min 1} = \sqrt{2(\gamma^2 - 1)q_0}, \tag{34}$$

coinciding with the known formula for the velocity of a detonation wave from an individual explosive [2]. The absence of the factor α under the root sign in (34) is explained by the fact that the inert impurity particles, not changing momentum in the course of the heat evolution, do not affect the detonation wave parameters (to within the effects related to the thermal relaxation of particles) and the wave propagates as in the individual explosive.

For $K \gg 1$, the velocity relaxation takes place after virtually complete heat evolution. For this reason, we may leave only maximum terms in the equations determining the tangency point, expanded in powers of $Q^* - 1$ ($Q^* \equiv q^*/q_0$), and find an explicit expression for the tangency condition (30). The corresponding calculations, performed in Appendix 4, lead to the following final result:

$$Q^* - 1 = \frac{\text{const}}{K},$$
$$\text{const} = -\frac{2\beta[\gamma^2 + 1 + \beta(\gamma^2 - 1)]}{(1-\beta)(\gamma+1)[1 + \beta(\gamma^2 - 1)]}. \tag{35}$$

The dependence of Q^* on K in the entire range of this parameter is considered in Section 6.

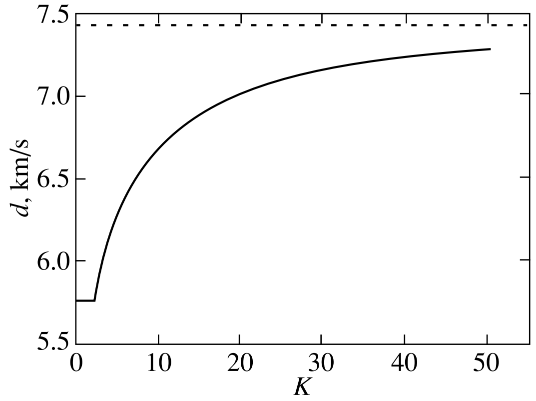


Fig. 5. Dependence of the detonation wave velocity on the ratio K of the characteristic times of the velocity relaxation and heat evolution calculated for q_0 = 3.455 kJ/g, β = 0.4, and γ = 3. In the asymptotics of large K, the wave velocity approaches d = 7.43 km/s (dashed line).

6. CALCULATI0N OF THE DETONATION VELOCITY AND DETONATION WAVE STRUCTURE AS FUNCTIONS OF THE PARAMETER K

The problem of determining the minimum detonation velocity and the detonation wave structure as functions of K, formulated in Section 3, was solved by numerical methods using the following initial data: explosion heat q_0 = 3.455 kJ/g; individual explosive density R_{ex} = 1.45 g/cm^3 (approximately equal to the values for water-saturated hexogen [10, 14]; impurity particle (tungsten) density r_p = 19.35 g/cm^3; mass concentration of impurity (tungsten) in the initial blend, $0 \le \beta \le 0.5$. The caloric equation of state of the explosive was assumed not to depend on the degree of conversion and was modeled in two variants: variant 1, formula (9) (ideal gas); variant 2, equation of state of the detonation products of water-saturated hexogen [14]. For example, Figs. 4–6 show the results of these mode calculations: (i) detonation velocity versus K (Fig. 5); (ii) detonation wave structure in the variables (x, y, q) for several values of K and the corresponding intrinsic detonation velocities d_K (Figs. 4 and 6). In addition, Fig. 4 shows the boundary lines (29) corresponding to these d_K values and the equilibrium points (16) (each equilibrium point corresponds to point *4* in Fig. 1). These illustrations correspond to the following variant: equation of state (9) with adiabatic exponent $\gamma = 3$ for a blend with the tungsten concentration β = 0.4. For this tungsten content, the specific volume of the initial blend is V_0 = 0.435 cm^3/g. Note that the saddle-point coordinate q^* in the case of undercompressed detonation coincides with q_0 only for $K = \infty$. The variation of $q^*(K)$ is nonmonotonic: as K increases in the interval (17), the value of q^* initially decreases and then begins to grow asymptotically approaching q_0 in accordance with (35).

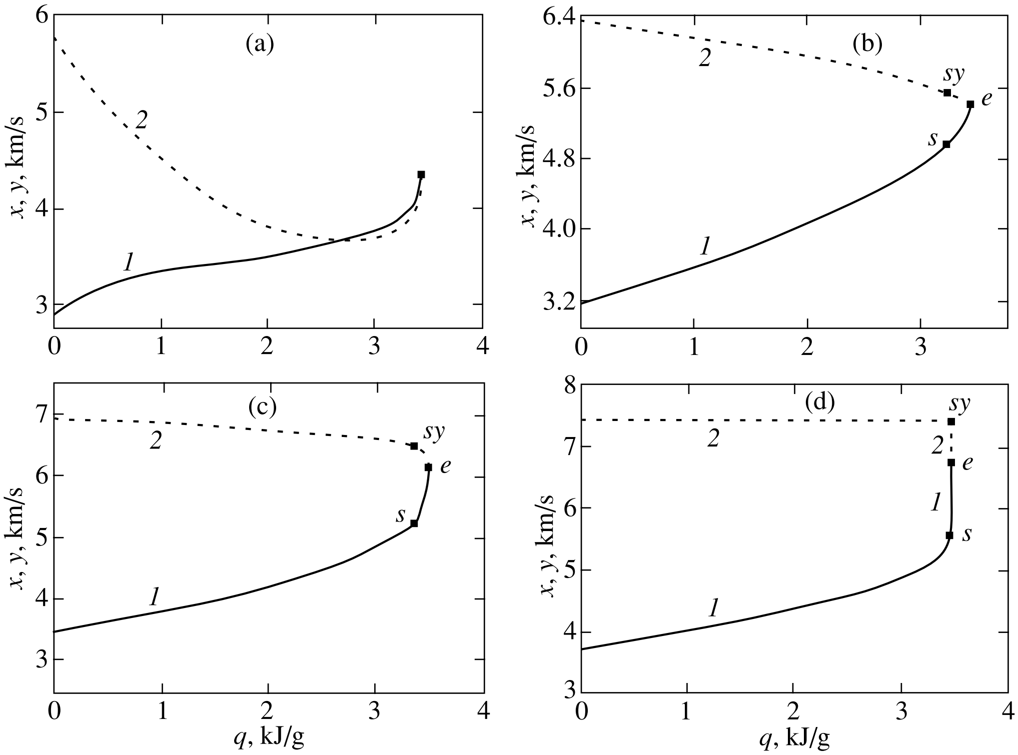


Fig. 6. Plots of the (*1*) gas and (*2*) particle velocity versus the heat evolution q calculated for $q_0 = 3.455$ kJ/g, $\beta = 0.4$, $\gamma = 3$, and (a) $d = 5.759$ km/s, $K = 0.5$; (b) $d = 6.328$ km/s, $K = 5.5$; (c) $d = 6.92$ km/s, $K = 16$; (d) $d = 7.435$ km/s, $K = \infty$. Point e is the boundary of the stationary zone ($q = q_0$, $x = y = 4.392$ (a), 5.403 (b), 6.15 (c), 6.75 km/s (d)); s is the saddle singularity on the (q, x) plane; sy is the point in the (y, q) plane corresponding to saddle s (tangency point, see lines in Fig. 4: *2* and $y2$ (b), *3* and $y3$ (c), *4* and $y4$ (d)).

7. THE STRUCTURE OF STRONG SHOCK WAVES RELATED TO COMPONENT VELOCITY RELAXATION

For $K \ll 1$, the component velocity relaxation begins and terminates in the stage where the heat evolution is negligibly small. (This statement does not apply to very special cases when a considerable part (~10%) of the heat is liberated in the stage of the shock wave compression [15].) The zones of relaxation of the component velocities and the heat evolution in the detonation wave structure are spatially separated, and the former belongs to the shock weave of the detonation complex. Investigation of the structure of this zone presents independent interest as a problem of the theory of shock waves in disperse media (gas or liquid containing solid particles). The shock wave structure in such systems was studied predominantly in the case of weak shock waves with an extended viscous profile [16–18]. In contrast, we are considering a strong shock wave in which the viscous jump is formed before the velocity relaxation and can be considered as a break, followed by the velocity relaxation. It is important to note that a small specific heat of the heavy impurity particles allows the heat exchange kinetics to be ignored and the effect of velocity relaxation on the shock wave structure to be studied separately (not masked by other processes).

In a strong shock wave, the gas component (ideal gas) is compressed by factor of

$$\theta \equiv \frac{\gamma + 1}{\gamma - 1}.$$

Since the velocity and density of impurity particles (mass per unit blend volume) in this stage remain unchanged, the total degree of compression is $\alpha\theta + \beta$. After leveling of the component velocities, the total degree of compression is θ. Important characteristic features of the velocity relaxation zone are as follows: (i) a nonmonotonic dependence of the gas velocity on the particle velocity and the time of particle residence behind the shock wave front (and on the distance traveled by particles behind the shock wave); the nonmonotonic character is weakly pronounced, but exists for any value of the adiabatic exponent γ and the particle concentration β; (ii) the generalized Michelson–Rayleigh line $V(p)$ is a convex curve asymptotically (for $t/\tau_v \gg 1$) tending to a straight line. In the vicinity of the onset of the velocity relaxation zone, the $V(p)$ line exhibits a maximum (compression preceded by a small expansion of the blend); an example of the calculated wave structure is given in Figs. 7 and 8.

In the other limit, $K \gg 1$, the stages of velocity relaxation and heat evolution are also separated in time, but the velocity relaxation takes place after the heat evolu-

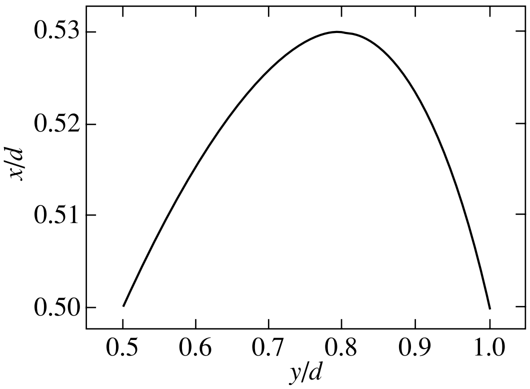


Fig. 7. Structure of the zone of velocity relaxation in the (x, y) plane for a two-component flow in a strong shock wave ($\gamma = 3$).

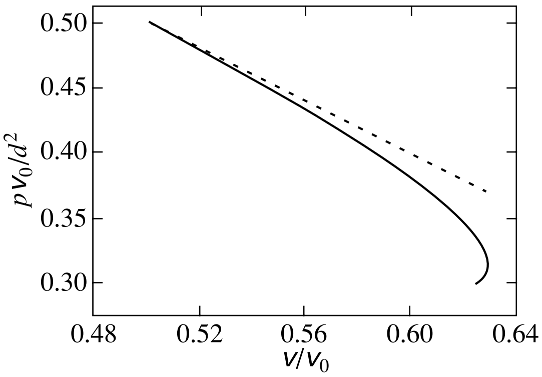


Fig. 8. Structure of the zone of velocity relaxation in the (v, p) plane for a two-component flow in a strong shock wave ($\gamma = 3$). Dashed line shows the Michelson–Rayleigh line.

tion and forms the final part of the stationary zone of the detonation wave (in a charge of unlimited diameter). Here, a relationship between the component velocities x and y is determined by Eq. (10) with $q = q_0$. In contrast to the case of $K \ll 1$, the functions $x(y)$ and $p(v)$ exhibit no anomalies in the intrinsic integral curves and the x and y values approach one another varying in the opposite directions: $dx/dt \geq 0$, $dy/dt \leq 0$; $dp/dt \leq 0$, $dv/dt \geq 0$ (the sign of equality corresponds to the point on the equilibrium detonation adiabat).

8. CONCLUSIONS

The physical and mathematical interpretation was provided for the principle of minimum for the velocity of self-sustained detonation with relaxing component velocities in a two-component flow. It was proved that the formula of the classical detonation theory expressing the detonation wave velocity as a function of the heat of explosion is valid not only in the asymptotics of small ratio K of the characteristic relaxation times of the component velocities and heat evolution, but in a certain finite interval ΔK adjacent to this asymptotics as well. For $K \sim 1$, the characteristic time of leveling of the component velocities (i.e., the time of attaining the point of intersection of the curves $x(q)$ and $y(q)$ (Fig. 6a) is smaller than the characteristic time τ_U entering into the relaxation equation (15), because the component velocities vary during relaxation at approximately the same rate in the opposite directions. For this reason, a change in the detonation regime from normal to undercompressed takes place at $K \approx 2$, rather than at $K \approx 1$ (see Fig. 5). The transition to undercompressed detonation is manifested by the appearance of a saddle-type singular point in the integral curve describing the detonation wave structure on the (x, q) plane. In the (y, q) plane, this transition corresponds to the point of tangency of the integral curve $y(q)$ and the boundary of real x values (velocity of the explosive) as a function of the variables y, q and the parameter d (detonation wave velocity) (Fig. 4). This correspondence of the saddle point and the tangency point takes place in the entire domain of existence of the undercompressed detonation. For this detonation regime, the saddle point coordinate q^* coincides with q_0 only for $K = \infty$. The function $q^*(K)$ is nonmonotonic: as K increases in the interval (27), q^* initially decreases and then begins to grow asymptotically, approaching q_0 in accordance with (35).

For $K \ll 1$, we have determined some characteristic features of the wave structures related to relaxation of the component velocities in a two-component flow.

APPENDIX 1

Let us prove that a real integral curve $x = x(q)$ at the boundary of the region of complex solution to Eqs. (10) and (16) is a separatrix of a saddle-type singular point.

Consider a segment

$$0 \leq q \leq q_0 \qquad \text{(A.1)}$$

and an integral curve satisfying certain initial conditions (e.g. $y = d|_{q=0}$), which is a solution to the differential equation

$$\frac{dy}{dq} = f_K(x, y, q), \qquad \text{(A.2)}$$

in which x is a root of the equation[3]

$$\varphi_d(x, y, q) = 0, \qquad \text{(A.3)}$$

where $\varphi_d(x, y\ q)$ is a differentiable function with real coefficients (parameters) and d is a parameter. In this paper, Eq. (A.2) was represented by Eq. (16). For the subsequent analysis, a particular form of the function $f_k(x, y, q)$ is insignificant. It is only assumed that this is a real continuous function of the parameter K, which allows the derivative dy/dq to be infinitely varied at a fixed value of the arguments. Of the two variables y and

---

[3] If there are several such roots, one has to be selected in accordance with the physical conditions.

q in Eq. (A.3), only one is independent since the integral curve is such that

$$y = y(q). \tag{A.4}$$

Substituting (A.4) into (A.3) yields an equation that can be written as

$$\Psi_{d,K}(x, q) = 0. \tag{A.5}$$

Equation (A.5) determines the dependence of x on q on the integral curve for preset values of the parameters d and K. Let us assume that there exists a boundary value d_K of the parameter d, such that, on one side of d_K (for certainty, at $d > d_K$), the root of Eq. (A.5) is real in the whole interval (A.1), while on the other side ($d < d_K$) in a certain subinterval of q from (A.1) the root is complex. Since the coefficients of Eq. (A.5) are real, the passage of d through d_K leads to the disappearance of two real roots (and the appearance of two complex-conjugate roots). At the point $d = d_K$, both these roots merge into one (denoted by x_K) for a certain value of $q = q_K$, and the function $\Psi_{dK}(x, q)$ exhibits an extremum[4]

$$\frac{d\Psi_{d,K}(x, q_K)}{dx} = 0.$$

At this point on the (q, y) plane, the integral curve touches the boundary line

$$D(y, q, d_K) = 0, \tag{A.6}$$

separating the regions of real and complex roots of Eq. (A.3) on the plane $d = d_K$.

Representing the function $\Psi_{d,K}(x, q)$ in the vicinity of the point (x_K, q_K) by a parabola, we write

$$\Psi_{d,K}(x, q) = a + b(x - x_K)^2, \tag{A.7}$$

where the coefficients a and b depend on q and d. For $d = d_K$ and $q = q_K$,

$$a = 0, \quad b = b_0 \neq 0.$$

In the vicinity of the point (d_K, q_K), we have (to within terms of higher order of smallness)

$$a = -\alpha_K(d - d_K) - \beta_K(q - q_K)^2, \quad b = b_0, \tag{A.8}$$

where α_K and β_K and constant quantities the sign of which is the same at that of b_0. Expansion (A.8) follows from the aforementioned conditions of existence of the real roots of Eq. (A.5).

From Eqs. (A.5), (A.7), and (A.8), it follows that

$$x = x_K \pm \sqrt{-\frac{a}{b_0}}, \quad \frac{dx}{dq} = \pm\frac{da}{dq}\frac{1}{2\sqrt{-ab_0}}.$$

This, together with (A.8), yields for $d = d_K$

$$x = x_K \pm \sqrt{\frac{\beta_K}{b_0}}(q - q_K), \quad \frac{dx}{dq} = \pm\sqrt{\frac{\beta_K}{b_0}}. \tag{A.9}$$

Now we will demonstrate that the point (x_K, q_K) on the plane of variables (q, x) is a saddle-type singular point for Eq. (16) and (A.3) for $d = d_K$ and, hence, the integral curve

$$x = f_K(q) \tag{A.10}$$

passing through this point is a separatrix of the family of integral curves of these equations. (The first equation (A.9) represents the segments of two separatrices.) For this purpose, let us consider the behavior of any two integral curves in the vicinity of the point (x_K, q_K). The curves are assumed to be arbitrarily close to the integral curve (A.10) with respect to the initial data and situated on different sides of this curve, so as to obey the initial conditions

$$y_0 = d_K \pm \delta|_{q=0}, \tag{A.11}$$

where δ is an arbitrarily small positive quantity.

Since the intrinsic integral curve is touching the boundary (A.6) of the domain of real roots, one of the adjacent integral curves will double-intersect the boundary (A.6) in the vicinity of the tangential point (q_K, y_K), while the other adjacent integral curve has no common points with this boundary. Therefore, the root x of Eq. (A.5) is real on the second integral curve everywhere in the vicinity of the point (q_K, y_K).

Let us introduce for these two integral curves the function (A.5) and its parabolic representation (A.7) by analogy with what was done for the intrinsic integral curve (i.e., at $\delta = 0$). As a result, all parameters of the parabola would shift by small values proportional to δ:

$$a \longrightarrow a_1 \equiv a \pm \lambda_a\delta, \quad b \longrightarrow b_1 \equiv b \pm \lambda_b\delta,$$
$$x_K \longrightarrow x_{K1} \equiv x_K \pm \lambda_K\delta,$$

where λ_a, λ_b, λ_K are certain constant coefficients. Accordingly, instead of (A.9), we obtain (to within terms of higher order of smallness)

$$\begin{gathered} x = x_{K1} \pm \sqrt{\frac{\beta_K(q - q_K)^2 + \Delta_1}{b_0}}, \\ \frac{dx}{dq} = \pm\sqrt{\frac{\beta_K}{b_0}}\frac{q - q_K}{\sqrt{(q - q_K)^2 + \Delta_1/\beta_K}}, \\ \Delta_1 \equiv \pm\lambda_a\delta. \end{gathered} \tag{A.12}$$

From (A.12), it follows that the integral curves in the vicinity of the point $q = q_K$ on the plane (q, x) exhibit a qualitative change with the sign of Δ_1/β_K: for $\Delta_1/\beta_K > 0$, the curves pass at point $q = q_K$ through the zero isocline, while for $\Delta_1/\beta_K < 0$ these curves reach the infinity iso-

---

[4] The second derivative also turning into zero at this point would be an incredible accident.

cline at the points $q = q_K \pm \sqrt{-\Delta_1/\beta_K}$. As can be readily seen, the four subfamilies of the integral curves (corresponding to two signs of the square root in (A.12) and two Δ_1 values) form together with separatrices (A.9) a characteristic saddle pattern with the singularity at (q_K, x_K).

APPENDIX 2

Let us consider the boundary of real solutions in the problem of the detonation wave structure. Solving the quadratic Eq. (10) relative to x, we obtain

$$x = \frac{Sd - S\beta y + \sqrt{Z(y, q, d)}}{(2S-1)\alpha}, \qquad (A.13)$$

where $Z(y, q, d)$ is the determinant that can be explicitly written as

$$\begin{aligned} Z(y, q, d) &= (S^2 - 2S\alpha + \alpha)d^2 \\ &+ \beta(S^2\beta + 2S\alpha - \alpha)y^2 - 2S^2\beta dy + 2\alpha^2(1-2S)q. \end{aligned} \qquad (A.14)$$

Substituting (A.14) into (19) and solving this equation relative to d yields after simple transformations the following expression for $F(y, q)$ at $d \ge 0$:

$$\begin{aligned} F(y, q) &= \frac{1}{E} \\ &\times [S^2\beta y + \alpha\sqrt{(2S-1)(2Eq - Ly^2)}], \end{aligned} \qquad (A.15)$$

where

$$E = S^2 - 2S\alpha + \alpha, \quad L = (S-1)^2\beta.$$

The coefficients E and L are positive; for E, this is seen from the identical transformation

$$E = S^2 - 2S\alpha + \alpha = (S-\alpha)^2 + \alpha\beta.$$

Let us demonstrate that $Z(y, q, d)$ with fixed y and q values is an increasing function of d. To this end, it suffices to show that the partial derivative of Z with respect to d is positive. The result of differentiating with renormalization $\alpha + \beta = 1$ can be written as

$$\frac{1}{2}\left(\frac{\partial Z}{\partial d}\right)_{y,q} = (S-1)^2\alpha d + S^2\beta(d-y). \qquad (A.16)$$

Now the positiveness of the right-hand part of (A.16) is obvious because $d > 0$, $y \le d$, and $\alpha \ge 0$ for the problem under consideration.

APPENDIX 3

Let us prove that the detonation velocity is independent of K for small values of this parameter. Upon the identical change of variables $y \equiv x + \Delta$, Eq. (10) takes the form

$$\frac{d^2}{2} = [d - (\beta\Delta + x)]Sx - \alpha q + \frac{x^2}{2} + \beta\frac{\Delta^2}{2} + \beta x\Delta. \qquad (A.17)$$

(It should be recalled that the heat evolution q is a monotonic function of time; see Eq. (14).) For small K or, accordingly, $-\Delta/x \ll 1$, Eq. (A.17) can be transformed (to within terms of higher order of smallness in $-\Delta/x$) to the following form:

$$\frac{d^2}{2} = Sdx - (S-1)\beta x_e\Delta - \alpha q - \left(S - \frac{1}{2}\right)x^2. \qquad (A.18)$$

Let us introduce the following heat evolution function:

$$\begin{aligned} q_1 &\equiv q + \frac{1}{\alpha}(S-1)\beta x_e\Delta \\ &= q_0[Q + 2K(1 - Q - \gamma\sqrt{1-Q})]. \end{aligned} \qquad (A.19)$$

In the right-hand part of (A.19), the quantities x_e and Δ are expressed in the form of (32) and (33) and $Q = q/q_0$. The function q_1 at small K monotonically increases with time and, similar to q, has a maximum value q_0. (Since $\gamma > 1$, the monotonic behavior of q_1 in the entire range of Q from 0 to 1 is ensured by the condition $K \le 3/2$.)

Equation (A.18) can be converted using the identity (A.19) into an equation of the classical detonation theory with a monotonic heat evolution function q_1:

$$\frac{d^2}{2} = Sdx - \alpha q_1 - (S - 1/2)x^2. \qquad (A.20)$$

For any sufficiently small value of K for which the heat evolution function is monotonic and, hence, exhibits a K-independent maximum value of q_0, Eq. (A.20) gives the same result (25) for the minimum detonation velocity.

APPENDIX 4

Let us derive the asymptotic formula (35). The right-hand part of (30) expresses the derivative dy/dq along the line (29), that is, along the direction in which $dF(y, q) = 0$. Differentiating (A.15) along this direction yields

$$\begin{aligned} \frac{dy}{dq} &= -\frac{\Lambda E}{S^2\beta\sqrt{2Eq - Ly^2} - \Lambda Ly}, \\ \Lambda &\equiv \alpha\sqrt{2S-1}. \end{aligned} \qquad (A.21)$$

For $K \gg 1$, all values in the right-hand part of (A.21) depend weakly on K and are close to their asymptotic values corresponding to $K = \infty$:

$$q = q_0, \quad y = d_\infty, \qquad (A.22)$$

where d_∞ is determined by formula (34). For these values of q and y, taking into account (34) and

$$\alpha + \beta = 1, \quad S = \frac{\gamma}{\gamma - 1}$$

and accomplishing simple transformations, we obtain

$$2Eq - Ly^2 = \frac{2q_0}{(\gamma-1)^2},$$
$$\Lambda Ly = \sqrt{2q_0}\alpha\beta\frac{\gamma+1}{(\gamma-1)^2}. \qquad (A.23)$$

The left-hand part of (30) expresses the derivative dy/dq along the integral curve. This derivative satisfies Eq. (16), in which the gas velocity x (with an allowance for (A.22)) is related to d_∞ by the relationship

$$x = \gamma\frac{d_\infty}{\gamma+1}.$$

(This expression is analogous to the relationship between the gas velocity and the detonation wave velocity known in the theory of normal detonation, which can be derived from Eq. (10).) Taking into account that

$$y - x = \frac{d_\infty}{\gamma+1},$$

we obtain for the derivative along the integral curve

$$\frac{dy}{dq} = -\frac{d_\infty}{K(\gamma+1)(q_0-q)}. \qquad (A.24)$$

Equating, according to (30), the expressions for derivatives (A.21) and (A.24) and taking into account (A.23), we arrive at (35) (see Section 5). Since the derivation procedure employed the asymptotic formulas (A.23) and (A.24), the result is valid to within $O[(Q^* - 1)^2]$ or $O(1/K^2)$.

REFERENCES

1. E. Youguet, J. Math. **6**, 5 (1904).
2. L. D. Landau and E. M. Lifshitz, *Mechanics of Continuous Media* (Gostekhizdat, Moscow, 1954).
3. Ya. B. Zel'dovich and S. B. Ratner, Zh. Éksp. Teor. Fiz. **11**, 170 (1941).
4. Ya. B. Zel'dovich and A. S. Kompaneets, *Theory of Detonation* (Gostekhizdat, Moscow, 1955).
5. N. M. Kuznetsov, Zh. Éksp. Teor. Fiz. **52**, 309 (1967) [Sov. Phys. JETP **25**, 470 (1967)].
6. A. N. Dremin, S. D. Savrov, V. S. Trofimov, and K. K. Shvedov, *Detonation Waves in Condensed Media* (Nauka, Moscow, 1970).
7. L. V. Al'tshuler, V. T. Ryazanov, and M. P. Speranskaya, Zh. Prikl. Mekh. Tekh. Fiz., No. 1, 122 (1972).
8. N. M. Kuznetsov, Khim. Fiz. **16**, 85 (1997).
9. N. M. Kuznetsov, Khim. Fiz. **17**, 62 (1998).
10. O. N. Davydova, N. M. Kuznetsov, V. V. Lavrov, and K. K. Shvedov, Khim. Fiz. **18**, 53 (1999).
11. L. V. Al'tshuler, V. S. Zhuchenko, and I. S. Men'shov, in *Proceedings of the Conference on Shock Waves in Condensed Media, St. Petersburg, 1998*, p. 116.
12. A. I. Ivandaev, A. G. Kutushev, and R. I. Nigmatulin, Itogi Nauki Tekh., Ser. Mekh. Zhidk. Gaza **16**, 209 (1981).
13. A. B. Gagiev and N. M. Kuznetsov, in *Modern Problems of Hydrodynamics, Aerophysics and Applied Mechanics* (MFTI, Moscow, 1986), p. 15.
14. N. M. Kuznetsov and K. K. Shvedov, Khim. Fiz. **18**, 82 (1999).
15. A. N. Dremin, *Towards Detonation Theory* (Springer-Verlag, New York, 1999).
16. S. P. D'yakov, Zh. Éksp. Teor. Fiz. **27**, 283 (1954).
17. G. M. Arutyunyan, Dokl. Akad. Nauk SSSR **185**, 778 (1969) [Sov. Phys. Dokl. **14**, 334 (1969)].
18. R. I. Nigmatulin, *Dynamics of Multiphase Media* (Nauka, Moscow, 1987), Part II.

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 533–541.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 609–618.
Original Russian Text Copyright © 2001 by Dolganov, Kats, Malinin.

Phase Transition to Anticlinic Texture in Free-Standing Smectic *C* Films

V. K. Dolganov[a, *], **E. I. Kats**[b, c], **and S. V. Malinin**[b, d]
[a]*Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
[b]*Landau Institute of Theoretical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
[c]*Institute Laue–Langevin, BP 156, Grenoble, France*
[d]*Forschungszentrum Juelich, D-52425, Juelich, Germany*
**e-mail: dolganov@issp.ac.ru*

Received May 16, 2001

Abstract—Experimental data on the optical reflectance of free-standing smectic *C* films were analyzed within the framework of a phenomenological Landau approach. At a certain temperature T_{0N} (determined from experimental data), which exceeds the known temperature T_c of the volume phase transition from smectic *A* to smectic *C* state, a surface phase transition takes place whereby molecules in the surface layer become sloped relative to the normal of the smectic layers. The transition temperatures $T_{0N}^{s,a}$ for *N*-layer films possessing synclinic (symmetric) and anticlinic (antisymmetric) textures of the order parameter (tilt angle θ) were determined. A comparison of the theoretical and experimental data allowed all parameters of the model to be determined (including critical indices of the correlation length and the surface order parameter). Three possible models of the transition from the state with transverse polarization (perpendicular to the molecular tilt plane) to the state with longitudinal polarization (parallel to this plane) are analyzed. The transition takes place at low (1°–2°) values of the order parameter θ in the middle layer of the film. © *2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Free-standing smectic films were discovered by Friedel as long ago as in 1922 [1], but detailed physical characterization of these objects began only in the late 1970s, when a large number of theoretical and experimental investigations were performed providing data on the structure, mechanical properties, and thermodynamics of the films (see, e.g., monograph [2]). Of special interest are the so-called liquid smectic films in which the layers can be considered as representing a two-dimensional (2D) liquid phase (smectic *A*) or a 2D nematic crystal (Smectic *C*).

Smectic *A* (*SmA*) phases possess a rather simple structure, with a director **n** (the unit vector determining a local axis of the average orientation of long axes of the molecules) orthogonal to the smectic layers. As a result, the only additional (with respect to an isotropic liquid) degree of freedom is related to the “solid-state” (i.e., sufficiently rigid) order in the smectic layers [3].

In smectic C (*SmC*) phases, the director **n** is inclined to make a certain polar angle θ with the normal **ν** of the layers, while the azimuthal angle φ represents a “soft” (in the Goldstone sense) part of the order parameter.

Although the above classification refers to the volume smectic phases, this approach can be applied to description of the freer standing films as well. A significant difference of films from volume phases consists in that finiteness of the former systems makes the order parameter of the films essentially inhomogeneous. In the case of *SmC* films, to which our study is devoted, this implies that the angles θ and φ vary across the film thickness. Until very recently, the only known texture in *SmC* films was the so-called synclinic (or symmetric) configuration (*s* configuration) characterized by a zero difference of the azimuthal angles in the adjacent smectic layers (δφ = 0). However, oblique smectic phases were reported in the past few years (see [4–7]) in which an anticlinic texture (*a* configuration) may exist under certain conditions. The latter texture may either correspond to a microscopically anticlinic structure with δφ = π for each pair of adjacent layers or represent a macroscopic structure in which an overall difference of the azimuthal angles for the whole film is

$$\Delta\varphi \equiv \sum_{i=1}^{N-1} \delta\varphi_{i+1,\,i} = \pi,$$

where *N i*s the total number of smectic layers in the film.

Note also a significant distinction of the free-standing films composed of oblique smectic layers from the smectic films bounded (and oriented) by solid surfaces. In the latter case, the energy of the director and polarization binding to the boundary may significantly vary with position over the interface, whereas the free-standing films may readily acquire orientation homogeneous in the film plane.

The purpose of this study was to theoretically interpret and process the experimental data available on the optical reflectance of free-standing films of a ferroelectric *SmC* phase. In Section 2, we formulate a model of the phase transition from *s* to *a* film configuration and present a phenomenological (in the Landau theory sense) description of this phase transition. In Section 3, a comparison of the theoretical profiles of the order parameter to the experimental data allows the values of all parameters involved in the theory to be determined (including indices and amplitudes characterizing the critical behavior of the surface and volume order parameters). In Section 4, these critical parameters are used to determine the temperature of the phase transition between states of the film with transverse and longitudinal polarizations. The results of these calculations agree satisfactorily with the experimental data and provide reasonable estimates for the material parameters of a liquid crystal employed in the model. Finally, Section 5 summarizes the main results of this study.

2. THEORETICAL MODEL

In order to describe the order parameter in an oblique liquid smectic phase (see, e.g., [3, 8]), it is necessary to determine the polar angle θ (measuring tilt of the director $\mathbf{n}$ relative to the normal $\boldsymbol{\nu}$ of the smectic layers) and the azimuthal angle φ (phase). These two real parameters form a two-component complex order parameter $\psi = \theta\exp(i\varphi)$ equivalent to the so-called $\mathbf{c}$ director representing a projection of the director $\mathbf{n}$ onto the plane of the smectic layer. The phase φ is the Goldstone part (degeneracy parameter) of the order parameter for all oblique smectic phases.

In the case of free-standing films (bounded by an isotropic medium—air), only the polar angle θ can be preset at the interface for the symmetry considerations. Therefore, it is this angle θ (modulus of the complex order parameter) that may exhibit inhomogeneous distribution across the film thickness. As for the azimuthal angle, both the experimental data [4–7] and the absence (for the symmetry considerations) of the surface binding energy for φ allow us to consider two possibilities: (i) *s* configuration, in which $\delta\varphi = 0$, and (ii) *a* configuration, in which either each pair of adjacent layers has $\delta\varphi = \pi$ or the total phase difference across the film is $\Delta\varphi = \pi$. For certainty, following conclusions formulated in [5], we will restrict the consideration to the second type of the *a* configuration.

The existence of both *s* and *a* configurations in the same material implies that at least two minima of the thermodynamic potential exist and, hence, a phase transition between these states is possible. It is important to note that the *s* and *a* configurations differ not only in the geometric structure of the order parameter but in their physical properties as well. It is known [3, 8] that spatially inhomogeneous director distribution unavoidably results in the so-called flexoelectric polarization $\mathbf{P}_f$

$$\mathbf{P}_f = e_1(\mathbf{n}\mathrm{div}\mathbf{n}) + e_3[\mathbf{n}\times\mathrm{rot}\mathbf{n}], \qquad (1)$$

where the coefficients e_1 and e_3 in most of the real smectic liquid crystals are on the order of 10^{-10}–10^{-11} C/m. Due to the uniaxial symmetry breakage in *SmC* films, the expression for $\mathbf{P}_f$ acquires a much more complicated form than Eq. (1), but this simple relationship between $\mathbf{P}_f$ and $\mathbf{n}$ is sufficient for our analysis.

Naturally, distributions of the flexoelectric dipole moments in the *s* and *a* configurations are different (see formula (1)). For this reason, a phase transition between *s* and *a* configurations can be induced by an external electric field. This scenario was studied by Link *et al.* [6] in free-standing achiral smectic films. Below, we will generalize this phase transition mechanism so as to include the chiral smectic phases (possessing a ferroelectric polarization along the axis perpendicular to the molecular tilt plane). In Section 3, theoretical expressions derived within the framework of the general model are compared to the experimental data.

First of all, it is necessary to demonstrate theoretically the possibility of existence of the *s* and *a* configurations in free-standing films composed of oblique smectic layers. According to the above remarks, a description of the order in the films of oblique smectic layers can be performed in terms of the polar angle θ alone. In this case, the free energy of the film (in the sense of the Landau theory) can be presented as the expansion

$$F = \int_{-L/2}^{L/2} dz\left[\frac{1}{2}A\theta^2 + \frac{1}{4}B\theta^4 + \frac{1}{2}C\left(\frac{d\theta}{dz}\right)^2\right] + F_s, \qquad (2)$$

where *A*, *B*, and *C* are some phenomenological coefficients. Note that the coefficient *A* must be zero at the point T_c of the volume phase transition from *SmA* to *SmC*. The gradient term takes into account the unavoidable inhomogeneity of the order parameter across the film thickness,[1] and F_s (surface energy) describes a physical modification of the system at the boundaries. By analogy with for-

---

[1] Within the framework of the Landau theory, we may neglect inhomogeneity of the order parameter in the plane of the layer, but the inhomogeneity across the film that is inherent in the finite system has to be taken into account in this theory.

mula (2), the surface energy F_s can also be presented as the expansion

$$F_s = \frac{1}{2}A'\left[\theta^2\left(-\frac{L}{2}\right) + \theta^2\left(\frac{L}{2}\right)\right] + \frac{1}{4}B'\left[\theta^4\left(-\frac{L}{2}\right) + \theta^4\left(\frac{L}{2}\right)\right], \quad (3)$$

where A' and B' are coefficients related to the surface properties of the film. The quantity B' can be considered as constant, while A' must go to zero at the point T_0 of the surface phase transition: $A' = \alpha'(T - T_0)$, where α' is independent of the temperature T.

The theory outlined above (in particular, formula (3)) is a modification of the de Gennes theory [9], which assumes that the surface induces ordering due to an interaction of the external field type (linear with respect to the order parameter). The linear interaction implies that molecules on the surface are always sloped. The surface energy (3) corresponds to a basically different situation: the tilt of molecules appears at a certain temperature exceeding T_c (for $T_0 > T_c$), and a decrease in the temperature leads to ordering in the inner layers. The experimental data agree with this scenario of a phase transition in free-standing films. For this reason, Eq. (3) is written with the neglect of interactions of the external field type.

Usually [9], the influence of the surface is described in terms of the so-called extrapolation length. Using the notation adopted in Eqs. (2) and (3), this quantity can be determined as $\lambda = C/A'$. Depending on the sign of λ, the phase transitions are conventionally divided into (i) ordinary (when $\lambda > 0$ and, hence, the surface suppresses the volume phase transition), (ii) extraordinary ($\lambda < 0$ and the surface stimulates the phase transition), and (iii) special (intermediate between ordinary and extraordinary, corresponding to $\lambda = \infty$, when, in the main approximation, the surface does not influence the phase transition). Note that even in the latter case the phase transition temperature in the film may differ from T_c because of the presence of the gradient term (reflecting purely geometric effects) in Eq. (2).

According to the experimental data, all the free-standing oblique smectic films studied to the present [4–7] exhibit the *SmA–SmC* phase transition at temperatures 20–30°C above the point (T_c) of the volume phase transition. This fact indicates that we deal with an extraordinary phase transition, for which $\lambda < 0$. On the other hand, according to Eq. (3), the surface phase transition for $\lambda < 0$ (i.e., for $A' < 0$) takes place at a higher temperature than the volume transition ($T_0 > T_c$). Obviously, the volume order established at $T = T_c$ will result in that the surface properties will also exhibit a certain critical behavior. In the interval $T_c < T < T_0$ (which is of most interest in this study), the volume correlation length is finite. This implies that the order parameter is described by a certain profile with a maximum on the surface. Minimization of the energy functional (2) leads to the following Euler–Lagrange equation:

$$\frac{d^2m}{dx^2} = m + m^3. \quad (4)$$

Here x and m are the dimensionless variables defined as

$$x = \frac{z}{\xi_b}, \quad m = \frac{\theta}{\theta_b}, \quad (5)$$

where the quantities ξ_b and θ_b are expressed, within the framework of the mean field theory with the free energy (2), as

$$\xi_b = \sqrt{\frac{C}{\alpha(T - T_c)}}, \quad \theta_b = \sqrt{\frac{\alpha(T - T_c)}{B}}.$$

Equation (4) must be supplemented by the boundary conditions

$$\pm\frac{C}{\xi_b}\left(\frac{dm}{dx}\right)\bigg|_{x = \pm l} + \alpha'(T - T_0)m(\pm l) + \frac{\alpha B'(T - T_c)}{B}m^3(\pm l) = 0, \quad (6)$$

where $2l \equiv L/\xi_b$ is the dimensionless film thickness.

Let us consider the case (describing our experimental data) of a free-standing film in which the surface phase transition is completed at a given temperature, while the volume phase still occurs in the *SmA* state. In describing the *SmC* phase, it is usually important to take into account the term with θ^6 in the volume free energy expansion. However, in any case (provided the temperature deviates significantly from T_c toward increase), we may ignore the nonlinear terms in Eq. (4) and find a symmetric solution to this equation in the form

$$m = \text{const} \cdot \cosh x. \quad (7)$$

Substituting expression (7) into the boundary conditions (6), we obtain an equation for the temperature T_{0N}^s below which the symmetric solution (7) corresponds to an ordered state of a smectic film composed of N layers:

$$T_{0N}^s = T_0 - \frac{C}{\alpha'\xi_b(T_{0N}^s)}\tanh\left(\frac{L}{2\xi_b(T_{0N}^s)}\right). \quad (8)$$

By the same token, we can find an antisymmetric solution

$$m = \text{const} \cdot \sinh x. \quad (9)$$

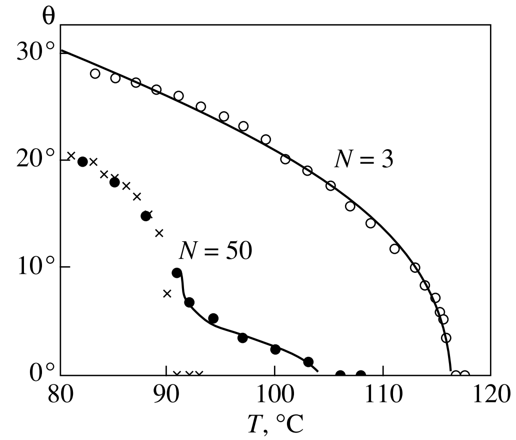


Fig. 1. Temperature dependence of the average polar angle θ of molecules in the free-standing NOBAMBC films containing $N = 3$ (open circles) and 50 (black circles) smectic layers. Crosses present the temperature variation of the slope of molecules in a bulk sample [5]. Solid curves show the results of fitting by formula (12).

This configuration also minimizes the free energy of the film and satisfies the boundary conditions below the temperature T^a_{0N} determined from the equation

$$T^a_{0N} = T_0 - \frac{C}{\alpha'\xi_b(T^a_{0N})}\coth\left(\frac{L}{2\xi_b(T^a_{0N})}\right). \tag{10}$$

The solutions to both transcendental equations were obtained numerically, indicating that the transition temperature is always higher in the *s* configuration: $T^s_{0N} > T^a_{0N}$. In the experiments with thin free-standing films, the transition from the *SmA* phase with decreasing temperature always leads to the *s* configuration. As can be readily seen from expansion (2), the *a* configuration is always metastable (in the absence of dipole–dipole interactions and external fields).

Strictly speaking, the continuous Landau–de Gennes theory expressed by Eqs. (2) and (3) must not describe the systems composed of a small number of clearly pronounced layers. This circumstance might be especially significant for thin *SmC* films in the temperature interval where the bulk correlation length is on the order of the film thickness. However, an analysis of the experimental data shows that the continuum approximation provides a sufficiently good qualitative description of the observed behavior. This is partly due to the fact that smectic liquid crystals grow, as a rule, by the so-called weak crystallization mechanism [10] in which case the smectic density modulation related to the discreteness is small as compared to the average density.

Another expression that will be required in the subsequent analysis is the order parameter profile in the *s* configuration. Using the notation adopted in Eqs. (2) and (3), this profile can be written as

$$\theta(x) = \theta_N \frac{\cosh x}{\cosh(L/2\xi_b)}, \tag{11}$$

where θ_N is the value of the order parameter on the film surface. In the next section, the theoretical model described above is applied to processing of the experimental data.

3. EXPERIMENTAL DATA TREATMENT

The angle of the director tilt in a free-standing *SmC** liquid crystal film can be determined by measuring the optical reflectance for two polarizations of the incident light beam, parallel and perpendicular to the molecular tilt plane. Since the indices of refraction in this plane (n_e) and in the perpendicular direction (n_0) are different, the intensity of reflection measured for the two polarizations allows us to determine the optical anisotropy of the film, which depends on the inclination of molecules in the smectic layer. The procedure of determination of the tilt angle of molecules in the films of arbitrary thickness is described in detail elsewhere [5]. It must be noted that, since the tilt angle depends on the distance to the film surface, the experimental data provide for an average value of the angle, with the type of averaging depending both on the film thickness and on the experimental method employed.

The simplest situation is observed for superthin films in which the correlation length of the surface ordering exceeds half of the film thickness ($\xi_b > L/2$). In this case, the tilt angle is virtually constant across the film and the experimental data can be treated in terms of a simple expression

$$\theta_N = \theta^{(0)}\left(\frac{T_{0N} - T}{T_{0N}}\right)^{\beta}, \tag{12}$$

where T_{0N} is the temperature of the transition to the *SmA* phase in an *N*-layer film. Here and below, the temperatures in power expressions are measured relative to the absolute scale.

Figure 1 shows the results of our measurements [5] of the average molecular tilt angle in the films containing $N = 3$ and 50 smectic layers. As demonstrated below, the film with $N = 3$ is characterized by $\xi > L/2$ in the entire temperature range. Therefore, the temperature dependence can be described using expression (12). The solid curve in Fig. 1 shows the results of fitting with $\beta = 0.46$, $\theta^{(0)} = 89.4°$, and $T_{0N} = 116.2°\text{C}$.

The situation with thick films is more complicated and requires special analysis. Here, an expression for the spectral dependence of the optical reflection intensity can be written as follows [11]:

$$I(\lambda) = \frac{(n^2-1)^2\sin^2(2\pi nNd/\lambda)}{4n^2 + (n^2-1)^2\sin^2(2\pi nNd/\lambda)}, \tag{13}$$

where λ is the light wavelength in vacuum. For thick *SmC** films ($N > 40$), the average tilt angle can be determined from the spectral position of the reflection minimum λ_{mC}. If the measurements are performed for the polarization perpendicular to the molecular tilt plane, the refractive index (n_0, the ordinary index of refraction) is independent of the slope. In this case, the temperature shift of the reflection minimum is related to a change in the film thickness,

$$L_C = \sum_{i=1}^{N} d_{Ci},$$

where d_{Ci} is the thickness of the *i*th smectic layer. For molecules modeled by rigid ellipsoids, the layer thickness d_{Ci} in the *SmC** phase can be represented as

$$d_{Ci} = d_A \cos\theta_i, \qquad (14)$$

where $d_A = L_A/N$ is the layer thickness in the *SmA* phase with equal interplanar spacings. A condition for the minimum phase difference at the minimum of the reflection intensity can be written as

$$\frac{2\pi n_0 d_A}{\lambda_{mC}} \sum_{i=1}^{N} \cos\theta_i = k\pi. \qquad (15)$$

According to this relationship, the experimentally determined ratio of wavelengths l_{mC}/l_{mA} corresponding to the reflection minima for the oblique and nonoblique structures is

$$\frac{\lambda_{mC}}{\lambda_{mA}} = \langle\cos\theta\rangle = \frac{1}{N}\sum_{i=1}^{N} \cos\theta_i, \qquad (16)$$

from which it follows that the tilt angle determined in a thick film represents the root-mean-square value $\langle\theta\rangle \equiv \sqrt{\overline{\theta^2}}$, or

$$\langle\theta\rangle = \left(\frac{1}{N}\sum_{i=1}^{N} \theta_i^2\right)^{1/2}. \qquad (17)$$

The expression for $\langle\theta\rangle$ can be readily obtained from the profile (11) of the order parameter in the *SmC** film:

$$\langle\theta\rangle = \frac{\theta_N(T)}{2^{1/2}\cosh(N/2\xi)}\left[1 + \frac{\xi}{N}\sinh\left(\frac{N}{\xi}\right)\right]^{1/2}, \qquad (18)$$

where $\xi = \xi^{(0)}[(T_c/(T - T_c)]^{\nu}$ is the volume correlation length for the tilt angle θ of molecules in the smectic layers (ξ is expressed in units of the layer thickness). Upon fitting formula (18) to the experimental data for a film with $N = 50$, we employed the parameters $\theta^{(0)} = 89.4°$, $\beta = 0.46$ (determined for a thin film) and the temperature of the transition to the *SmA* phase $T_{0N} = 104°\mathrm{C}$. The fitting results (Fig. 1) correspond to $\xi^{(0)} = 0.75$ and $\nu = 0.37$. Using these values of $\theta^{(0)}$, β, $\xi^{(0)}$, and ν, it is possible to calculate all structural and polarization characteristics of the film and to model the temperature dependence of the transition from transverse to longitudinal ferroelectric polarization in free-standing smectic films of various thicknesses.

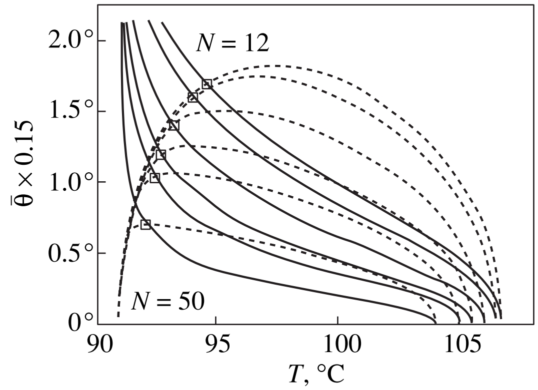


Fig. 2. Temperature dependence of the average polar angle $\bar{\theta}$ (solid curves) of molecules and the longitudinal polarization P_{0f} (dashed curves) calculated for free-standing films possessing various thicknesses (numbers of smectic layers) N = 12, 14, 19, 25, 31, and 50. Points of intersection (indicated by open squares) of the solid and dashed curves corresponding to the same N give the temperatures at which the longitudinal polarization is equal to the transverse polarization (for $e_3 = 2.5 \times 10^{-12}$ C/m and $\bar{P}_0 = \bar{\theta} \times 2.1 \times 10^{-5}$ C/m^2).

The subsequent analysis will also require an expression for the arithmetic mean of the tilt angle in the *s* configuration. This formula is also readily obtained from expression (11):

$$\bar{\theta} = \theta_N \frac{2\xi}{N}\tanh\frac{N}{2\xi}. \qquad (19)$$

4. THE TRANSITION BETWEEN STATES WITH TRANSVERSE AND LONGITUDINAL POLARIZATIONS

In the *SmC*, *SmC** (chiral ferroelectric), and *SmA** (chiral antiferroelectric) phases, the energy minima corresponding to an oblique orientation of molecules can take place at $\delta\varphi = 0$ and $\delta\varphi = \pi$. Which particular type of the structure (synclinic versus anticlinic) is formed depends on the relative magnitudes of these minima.

Now we will consider several models of the transition of the *SmC** film from the transverse to the longitudinal polarization state, which is possible for films with an antisymmetric profile of the order parameter. In contrast to [4, 5], where this structure was called the *C* configuration and remained undetermined, the first two models will consider transitions between the *s* configuration and the "inverted" *s* configuration with the tilt angle described in one half of the film (e.g., that with

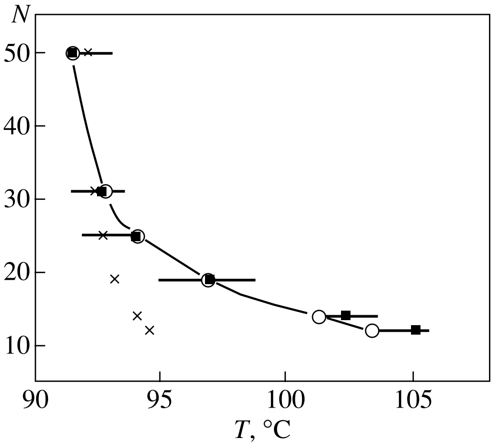


Fig. 3. The experimental temperatures of the phase transitions between the states of transverse and longitudinal polarizations in films of various thicknesses. The uncertainty (indicated by horizontal bars) is due to the temperature hysteresis revealed by experiment [5]. Symbols ×, ○, and ■ present the results of calculations using various models of the phase transition.

$x > 0$) by the “normal” formula (11), according to which $\theta(x) \propto \cosh x$, and in the other half ($x < 0$) by the relationship $\theta(-x) = -\theta(x)$. This anticlinic ($\tilde{a}$) configuration, corresponding to a break in the order parameter at $x = 0$, is not described as a whole by the Landau theory in the form of Eqs. (2) and (3). The point is that the chiral ferroelectric liquid crystals exposed to an external electric field or oriented by solid surfaces feature a competition between chirality [the contribution of which to the system energy is minimum for an inhomogeneous (spiral) orientation of molecules] and the polarization–field interaction that tends to establish a homogeneous order of the molecules; the system under such conditions is referred to as frustrated. The competing trends are especially significant in system of lower dimension (films) where one of the possible mechanisms for eliminating the frustration is the formation of defects such as domain walls analogous to the Abrikosov vortices in superconductors.

In the first model, we assume that the phase transition takes place when the magnitude of the longitudinal flexoelectric polarization becomes equal to the classical ferroelectric polarization of the film. According to formula (1), the average (over the film) longitudinal flexoelectric polarization $\bar{\mathbf{P}}_f$ for small angles in the $\tilde{a}$ configuration depends on the difference $\theta_N - \theta(0)$ between the tilt angle of molecules on the surface and in the middle layer of the film and on the flexoelectric coefficient e_3:

$$|\bar{\mathbf{P}}_f| = \frac{\pi}{180}\frac{2e_3(\theta_N - \theta(0))}{Nd}, \qquad (20)$$

where

$$\theta(0) = \frac{\theta_N}{\cosh(N/2\xi)} \qquad (21)$$

and $d \approx 3 \times 10^{-9}$ m is the smectic layer thickness.

Figure 2 (dashed curves) shows the temperature dependence of the value $P_{0f} = (\theta_N - \theta(0))/N$ proportional to the average flexoelectric polarization calculated for various film thicknesses $N = 50$, 31, 25, 19, 14, and 12. As is seen, the longitudinal flexoelectric polarization goes to zero at the temperatures of the volume and surface phase transitions. Solid curves in Fig. 2 show the temperature dependence of the average tilt angle $\bar{\theta}$ calculated by formula (19), which is assumed to be proportional to the transverse ferroelectric polarization. Reorientation of the film with $N = 50$ in the electric field and the transition from transverse to longitudinal polarization take place at $T \approx 92$°C, which corresponds to the average tilt angle $\bar{\theta} \approx 4.7°$. The magnitude of polarization at this angle for a related DOBAMBC liquid crystal is $P_0 < 10^{-5}$ C/m^2 [12, 13]. For equal transverse and longitudinal polarizations, the flexoelectric coefficient is $e_3 < 2.5 \times 10^{-12}$ C/m, which is significantly smaller as compared to typical values in liquid crystals (10^{-10}–10^{-11} C/m). These estimates indicate that the flexoelectric polarization during this transition markedly exceeds the spontaneous ferroelectric polarization P_0.

The final evidence that no transition takes place at $\bar{P}_0 = |\bar{\mathbf{P}}_f|$ is provided by a comparison of the experimental [5] and calculated temperatures of transitions in the films of various thicknesses. This comparison is illustrated in Fig. 3, where temperature intervals featuring the transition from transverse to longitudinal polarization are indicated by horizontal bars, since the experiment [5] revealed certain temperature hysteresis. The proposed model fails to describe the experimental data even provided that an unreasonably small flexoelectric coefficient e_3 is employed. For illustration, let us determine the upper estimate of the transition temperature assuming that $e_3 = 2.5 \times 10^{-12}$ C/m and $\bar{P}_0 = \bar{\theta} \times 2.1 \times 10^{-5}$ C/m^2. This example corresponds to the scale used in Fig. 2, where the points of intersection of the solid and dashed curves calculated for the same N give the temperatures at which the longitudinal polarization $|\bar{\mathbf{P}}_f|$ in the $\tilde{a}$ configuration becomes equal to the transverse polarization $\bar{P}_0$ in the s configuration. These points of intersection are plotted by crosses in Fig. 3. Although the transition temperature increases, as expected, with decreasing film thickness, the temperature dependence for any e_3 is much less pronounced as compared to the observed behavior. Taking into account a nonlinearity in the $P_0(\theta)$ value at small angles [14] even increases the discrepancy between this theoretical model and experiment. It should be recalled that the longitudinal flexoelectric polarization was calculated for the $\tilde{a}$ configuration. In the case of the anticlinic *a* configuration (9), the longitudinal polarization will be still greater because $\theta(0) = 0$.

The second proposed model assumes that the electric field produces reorientation of the molecules, driving them from an energy minimum corresponding to the synclinic configuration (*s*) to the state with anticlinic configuration ($\tilde{a}$). Since the depth of these energy minima increases with the slope of molecules, the reorientation may take place at the middle of the film where θ is small. In accordance with the Landau theory, it is natural to assume that the energy difference between $\tilde{a}$ and *s* configurations is proportional to the square polar angle $\theta^2(0)$ in the middle of the film: $\Delta F = \gamma\theta^2(0)S$, where *S* is the film area. As was mentioned above, $|\bar{\mathbf{P}}_f| \gg \bar{P}_0$ for the temperature interval of interest and the values of flexoelectric coefficients typical of liquid crystals; therefore, the $\tilde{a}$ configuration possesses a lower energy in an electric field with the strength *E*, provided that

$$\frac{\pi}{180}\frac{2e_3E(\theta_N - \theta(0))}{Nd}NdS > \gamma\theta^2(0)S,$$

from which it follows that the moment of transition corresponds to the relationship

$$\gamma = \frac{\pi e_3 E}{90}\frac{\theta_N - \theta(0)}{\theta^2(0)}.$$

Figure 4 presents the plots of $y_N \equiv (\theta_N - \theta(0))/\theta^2(0)$ versus temperature constructed for various *N* values using the parameters $\theta^{(0)}$, β, $\xi^{(0)}$, and ν determined as described above. Considering the coefficient γ in this model as the fitting parameter, we may reach agreement between the experimental data and the calculated transition temperatures (plotted by squares in Fig. 3) determined as the points of intersection of the curves for various *N* and the line $y \approx 3.5$ in Fig. 4. For $E = 10^3$ V/m and $e_3 = 10^{-11}$ C/m, the coefficient γ is 1.2×10^{-9} J/m^2 and the energy difference between the synclinic and anticlinic orientations in this films approximately amounts to 8×10^{-5} K per pair of adjacent molecules (for the number density of molecules in the smectic layer of about 6×10^{18} m^{-2}). For the second model, it is important that the state with longitudinal polarization corresponds exactly to the $\tilde{a}$ configuration.

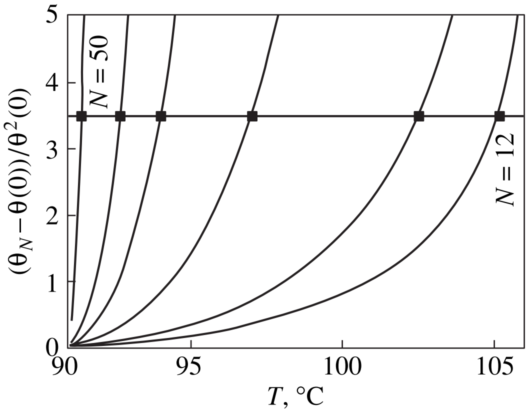


Fig. 4. Calculated temperature dependence of the quantity $y_N \equiv (\theta_N - \theta(0))/\theta^2(0)$ for free-standing films with various numbers of smectic layers *N* = 12, 14, 19, 25, 31, and 50. Black squares indicate the points of intersection of the curves with the straight line *y* = 3.5.

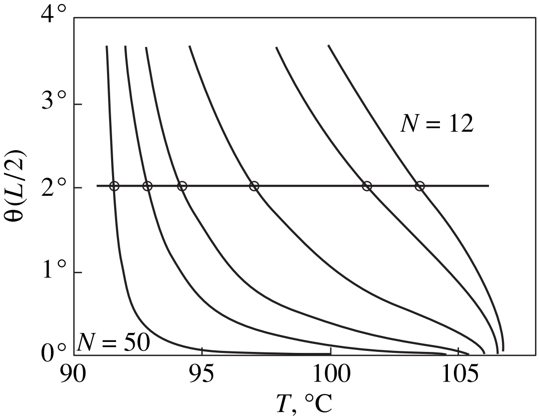


Fig. 5. Calculated temperature dependence of the tilt angle of molecules in the middle plane of the free-standing films with various numbers of smectic layers *N* = 12, 14, 19, 25, 31, and 50. Open circles indicate the points of intersection of the curves with the straight line θ(0) = 2.04°.

In the third model to be considered, it is assumed that the phase transition takes place at a certain small tilt angle in the middle layer of the film. The transition can be induced, for example, by competition of the interaction between adjacent layers and the next to adjacent layers [15, 16]. These interactions stabilize different mutual orientations of molecules in the smectic layers and, hence, may lead to the reorientation of molecules in the middle layers of the film. Such a transition may also be caused by an electrostatic interaction of charges induced by the polarization gradient between the upper and lower parts of the free-standing film [17]. The signs of electric charges induced in the two halves of the film are the same for the *s* configuration and are opposite in the antisymmetric case. In the latter case, a decrease in the electrostatic energy of the film may lead to the reorientation of molecules in the smectic layers.

Figure 5 presents the temperature dependence of the tilt angle in the middle layer versus temperature constructed for various *N* values using formula (21). The transition temperatures, determined as the points of intersection of the curves for various *N* and the line θ = 2.04°, are plotted by open squares in Fig. 3 showing that a properly selected value of the $\theta(0)$ angle (corresponding to the onset of the phase transition) provides for a good agreement between this theoretical model and experiment.

The latter two models, capable of adequately describing the temperature dependence of the phase transition, are characterized by different states of the film in a zero field. In the third model, according to which the transition takes place at a certain small value of $\theta(0)$, the film must possess a nonzero longitudinal polarization at a sufficiently high temperature even in a zero field. In the second model, where the transition is induced by the electric field, the film in a zero field at a high temperature must possess a transverse polarization. This situation was observed in SmC^* liquid crystals possessing a small ferroelectric polarization [17]. This class of liquid crystals includes NOBAMBC.

Thus, the whole combination of the experimental data allows us to select the model in which an applied electric field drives a smectic structure from one energy minimum, corresponding to a synclinic orientation of molecules in the middle smectic layer, to another minimum corresponding to the anticlinic orientation. This transition takes place virtually at the same small values of the tilt angle in the middle layer (1°–2°) as those in the third model. The model in which the transition takes place at a certain angle $\theta(0)$ may be applicable to substances with large values of the transverse ferroelectric polarization. However, the absence of a sufficiently large volume of experimental data (temperature dependence of the tilt angle, repolarization temperatures) for this class of substances hinders quantitative comparison of this model to experiment.

5. CONCLUSION

As demonstrated above, the available experimental data on the temperature dependence of the slope of molecules in free-standing can be adequately described within the framework of a phenomenological model of thin SmC^* films. Parameters involved in the theory (critical values of the correlation length, order parameter, etc.) were determined by comparing the theory with experiment. Using these empirical parameters, the macroscopic characteristic of the films were calculated, including the temperature dependence of the longitudinal ferroelectric polarization. Three possible models of the phase transition from the state with transverse polarization (symmetric profile of the order parameter) to the state with longitudinal polarization (antisymmetric profile) were analyzed. The possibility of calculating the macroscopic characteristics of the films allowed a model to be selected that most adequately describes both the temperatures of the phase transitions in an applied electric field and the state of the film in the absence of external fields.

A special consideration is necessary for determining the surface polarization and establishing the conditions for this polarization. For example, the existence of a polar binding at the boundaries must necessarily result in the slope of molecules with opposite signs on the two boundaries (a configuration). This configuration unavoidably leads to the energy of elastic deformation of the director being proportional to $1/L$. This configuration is energetically unfavorable for thin films, where a transition to the s configuration must take place with the gain in the deformation energy exceeding the loss in the energy of polar binding.

In this study, the consideration was restricted to a “minimal” model describing the transition from the state with transverse polarization to that with longitudinal polarization. However, the possibility of a microscopic anticlinic configuration (with $\delta\varphi = \pi$ for the adjacent layers) pointed out in the Introduction requires taking into account both the ferroelectric polarization P_0 and the antiferroelectric polarization P_a. The possibility of additional interactions and the manifestations of discreteness in thin films may give rise to a richer pattern of phenomena with finer effects as compared to those described and explained in this paper. For example, the observed dependence of the surface phase temperature on the film thickness in thick samples cannot be rationalized within the framework of a simple theory describing only the a configuration. However, on assuming the existence of an $\tilde{a}$ configuration, behavior of the transition temperature—a quantity extremely sensitive to the variation of parameters—can be explained only by explicitly considering the corresponding energy contributions.

Unfortunately, the experimental method employed (in fact measuring only the tilt angle averaged over the whole film) is incapable of distinguishing between many microscopic configurations. For this reason, we declined from considering such problems in more detail.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 01-02-16507), Federal Program “Statistical Physics,” and INTAS Foundation (grant no. 97-30234). One of the authors (S.V.M.) gratefully acknowledges the support from DFG (grant KO 1391/4-1).

REFERENCES

1. G. Friedel, Ann. Phys. (Paris) **18**, 273 (1922).
2. P. S. Pershan, *Structures of Liquid Crystal Phases* (World Scientific, Singapore, 1988).
3. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Clarendon, Oxford, 1993).
4. P. O. Andreeva, V. K. Dolganov, and K. P. Meletov, Pis’ma Zh. Éksp. Teor. Fiz. **66**, 442 (1997) [JETP Lett. **66**, 442 (1997)].
5. P. O. Andreeva, V. K. Dolganov, R. Fouret, *et al.*, Phys. Rev. E **59**, 4143 (1999).

6. D. R. Link, G. Natale, J. E. Maclennan, *et al.*, Phys. Rev. Lett. **83**, 3665 (1999).
7. M. Kimura, D. Kang, and C. Rosenblatt, Phys. Rev. E **60**, 1867 (1999).
8. S. Chandrasekhar, *Liquid Crystals* (Cambridge Univ. Press, New York, 1977; Mir, Moscow, 1980).
9. K. Binder, *Phase Transitions and Critical Phenomena*, Ed. by C. Domb and J. L. Leibowitz (Academic, London, 1986), Vol. 8.
10. E. I. Kats, V. V. Lebedev, and A. R. Muratov, Phys. Rep. **228**, 1 (1993).
11. M. Born and E. Wolf, *Principles of Optics* (Pergamon, Oxford, 1964; Nauka, Moscow 1973).
12. B. I. Ostrovskiĭ, A. Z. Rabinovich, A. S. Sonin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **25**, 80 (1977) [JETP Lett. **25**, 70 (1977)].
13. S. A. Pikin, *Phase Transitions in Liquid Crystals* (Nauka, Moscow, 1980).
14. S. Dumrongrattana, C. C. Huang, G. Nounesis, *et al.*, Phys. Rev. A **34**, 5010 (1986).
15. B. Rovsek, M. Cepic, and B. Zeks, Phys. Rev. E **54**, R3113 (1996).
16. B. Rovsek, M. Cepic, and B. Zeks, Phys. Rev. E **62**, 3758 (2000).
17. P. M. Johnson, D. A. Olson, S. Pankratz, *et al.*, Phys. Rev. E **62**, 8106 (2000).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 542–557.
Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 619–636.
Original Russian Text Copyright © 2001 by Zobov, Lundin, Rodionova.

The Shape of NMR Absorption and Cross-Relaxation Spectra in a Heteronuclear Spin System

V. E. Zobov, A. A. Lundin*, and O. E. Rodionova

Semenov Institute of Chemical Physics, Russian Academy of Sciences, Moscow, 117334 Russia
**e-mail: andylun@orc.ru*

Received December 9, 2000

Abstract—A dynamic theory of heteronuclear spin systems in solids at high temperatures is developed. A system of nonlinear integral equations for the spin time correlation functions is derived in the self-consistent fluctuating local field approximation taking into account corrections for the correlated local field fluctuations in real crystal lattices. The theory is applied to interpretation of the experimental data available for a LiF crystal representing a system with the nuclei of two types. The signals of free precession for Li and F nuclei, as well as the spectra of harmonic cross-relaxation, cross-polarization of the ^{6}Li isotope, and depolarization of the ^{8}Li isotope, were calculated for the magnetic field oriented along the principal crystallographic axes of LiF. The results of calculations show good agreement with experiment. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Experiments involving NMR in solids provide, depending on the selected method of signal registration, the spectra of various time correlation functions (TCFs). In the simplest case, this is the NMR absorption spectrum representing the Fourier image of a TCF called the signal of free precession, which can be directly recorded in standard pulsed NMR spectrometers. Using more sophisticated experimental schemes, it is possible to measure different variants of the cross-relaxation spectra.

The use of various cross-relaxation techniques sharply increased the sensitivity of NMR measurements, allowed the spectra of nuclear isotopes with low natural occurrence to be studied, and led to the development of two-dimensional (2D) and three-dimensional (3D) NMR Fourier transform spectroscopy that are very important for some applied (predominantly structural) investigations (see, e.g, [1]).

The cross-relaxation spectra, as well as all other NMR spectra of solids, reflect the process of attaining equilibrium in a system of many interacting bodies under various conditions. Therefore, the problems of describing the shapes of the TCF spectra measured by various NMR techniques are essentially the partial cases of a central problem of nonequilibrium statistical mechanics. In heteronuclear systems (i.e., in substances containing nuclei of various types with significantly different Larmor frequencies), an important role in the process of attaining equilibrium between various subsystems belongs to the shape of far branches (wings) of the spectra, as characterized by the rate of decay with increasing frequency detuning. The corresponding bibliography and an analysis of the relevant experimental data can be found in [2].

It must be noted that a correct calculation of the shape of the wings poses significant additional requirements to the theory as compared to the approach to calculation of the central part of the spectrum (cf. [3–5] and [6]).

In 1969, McArthur, Hahn, and Walstedt [7] experimentally observed for the first time an exponential decay of the cross-relaxation rate with increasing detuning under double resonance conditions at large detunings (on the wing of the spectrum). Subsequently, the exponential shape of wings of the NMR spectra was repeatedly observed in experiment for numerous solids studied under various NMR techniques (see, e.g., [2, 8]).

The first qualitative explanations of the observed experimental facts were proposed in [6, 9] for homonuclear spin systems and in [2] for the heteronuclear case. The theory developed in [2, 6] was based on the diagram expansion of TCFs, with the main approximation adopted there assuming the presence of an infinitely large number Z of the equivalent nearest neighbors surrounding a given nucleus in the sublattice, which was achieved in the limit of the space of infinite dimension ($d \longrightarrow \infty$). All characteristics of the system were expressed through a single scaling parameter, the second moment of the spectrum.

For a heteronuclear spin system considered in [2] in the same approximation as that used for the homonuclear case in [6], we obtained analytical estimates of the frequency asymptotes of the TCF spectra observed in various NMR experiments involving Li nuclei, which were performed predominantly with the external magnetic fields oriented along the [111] direction of a LiF crystal—a traditional model crystal with a heteronuclear spin system. The agreement between theory and experiment was achieved in [2] only on a qualitative level.

In this study, much more correct results were obtained owing to the transition to real 3D lattices (d = 3). Now it is possible, in particular, to follow the preasymptotic behavior of the wings. This is important in most cases for description of the available experimental data, for example, of the β-NMR spectra reported in [10, 11].

2. MAIN EQUATIONS FOR TIME CORRELATION FUNCTIONS

As is well known [3], the main interaction responsible for broadening of the NMR spectra in nonmetallic diamagnetic solids, thus determining both dynamic and thermodynamic properties of the nuclear spin subsystem, is the secular part of the magnetic dipole–dipole interaction between nuclei:

$$H = \sum_{\alpha} H^{(\alpha)}_{\text{hom}} + \sum_{\alpha > \beta} H^{(\alpha,\beta)}_{\text{het}}, \tag{1}$$

$$H^{(\alpha)}_{\text{hot}} = \sum_{i>j} b^{(\alpha)}_{ij} \left\{ S^{(\alpha)}_{zi} S^{(\alpha)}_{zj} - \frac{1}{4}[S^{(\alpha)}_{+i} S^{(\alpha)}_{-j} + S^{(\alpha)}_{-i} S^{(\alpha)}_{+j}] \right\}. \tag{2}$$

Hamiltonian (2) describes the interaction inside a (homogeneous) spin subsystem containing the nuclei of type α, while

$$H^{(\alpha,\beta)}_{\text{het}} = \sum_{k,m} b^{(\alpha,\beta)}_{km} S^{(\alpha)}_{zk} S^{(\beta)}_{zm} \tag{3}$$

is the secular part of the (heterogeneous) magnetic dipole–dipole interaction between nuclei of types α and β. As usual, the dipole coefficients $b^{(\alpha)}_{ij}$ and $b^{(\beta)}_{km}$ are determined by the mutual arrangement of interacting spins and by the direction of external magnetic field [3]:

$$\begin{aligned} b^{(\alpha)}_{ij} &= \frac{\gamma^2_\alpha h^2}{r^3_{ij}}(1 - 3\cos^2\theta_{ij}), \\ b^{(\alpha,\beta)}_{km} &= \frac{\gamma_\alpha \gamma_\beta h^2}{r^3_{km}}(1 - 3\cos^2\theta_{km}). \end{aligned} \tag{4}$$

Here, we use the standard notation [3]: $\mathbf{r}_{ij}$ is the internuclear radius vector connecting ith and jth spins; θ_{ij} is the angle of this vector relative to the external magnetic field; and γ_α and γ_β are the gyromagnetic ratios of the nuclei of types α and β, respectively.

The Fourier image of the NMR absorption spectrum for the nuclei of type α, representing the free precession signal arising upon application of a π/2 pulse to the equilibrium spin system, is proportional to TCF:

$$\Gamma^{(\alpha)}(t) = \frac{\text{Sp}(S^{(\alpha)}_x(t) S^{(\alpha)}_x)}{\text{Sp}(S^{(\alpha)2}_x)}, \tag{5}$$

where $S^{(\alpha)}_x$ is the total x component of the spins of type α and $S^{(\alpha)}_x(t)$ function is a solution to the Heisenberg equation

$$\frac{dS^{(\alpha)}_x(t)}{dt} = \frac{i}{\hbar}[H, S^{(\alpha)}_x] \tag{6}$$

with Hamiltonian (1). Below, we will also use the supplementary autocorrelation functions for the longitudinal and transverse components of individual spins:

$$\Gamma^{(\alpha)}_0(t) = \frac{\text{Sp}(S^{(\alpha)}_{xi}(t) S^{(\alpha)}_{xi})}{\text{Sp}(S^{(\alpha)2}_{xi})}, \tag{7}$$

$$\Gamma^{(\alpha)}_z(t) = \frac{\text{Sp}(S^{(\alpha)}_{zi}(t) S^{(\alpha)}_{zi})}{\text{Sp}(S^{(\alpha)2}_{zi})}. \tag{8}$$

The transverse local field components are determined by the interaction with spins of the same type,

$$h^{q(\alpha)}_i = -\frac{1}{2}\sum_j b^{(\alpha)}_{ij} S^{(\alpha)}_{qj}, \quad q = x, y, \tag{9}$$

whereas the longitudinal local field components contain a contribution due to the interaction (3) between spins of different types,

$$h^{z(\alpha)}_i = \sum_j b^{(\alpha)}_{ij} S^{(\alpha)}_{zj} + \sum_{\beta \neq \alpha} \sum_k b^{(\alpha,\beta)}_{ik} S^{(\beta)}_{zk}. \tag{10}$$

Previously [2, 12–14], we demonstrated that the correlation functions of the local field components (9) and (10) in a system with large number Z ($Z \longrightarrow \infty$) of the equivalent nearest neighbors surrounding a given spin in the lattice are proportional to TCFs (7) and (8):

$$g^{(\alpha)}_x(t) = \frac{\text{Sp}\{h^{x(\alpha)}_i(t) h^{x(\alpha)}_i\}}{\text{Sp}E} = \frac{1}{4}\Delta^2_{\alpha\alpha}\Gamma^{(\alpha)}_0(t), \tag{11}$$

$$\begin{aligned} g^{(\alpha)}_z(t) &= \frac{\text{Sp}\{h^{z(\alpha)}_i(t) h^{z(\alpha)}_i\}}{\text{Sp}E} \\ &= \Delta^2_{\alpha\alpha}\Gamma^{(\alpha)}_z(t) + \sum_\beta \Delta^2_{\alpha\beta}\Gamma^{(\beta)}(t), \end{aligned} \tag{12}$$

where E is the unit matrix and

$$\begin{aligned} \Delta^2_{\alpha\alpha} &= \frac{1}{3}S^{(\alpha)}(S^{(\alpha)} + 1)\sum_j b^{(\alpha)2}_{ij}, \\ \Delta^2_{\alpha\beta} &= \frac{1}{3}S^{(\beta)}(S^{(\beta)} + 1)\sum_n b^{(\alpha,\beta)2}_{mn} \end{aligned} \tag{13}$$

are the average squares of the two contributions to the local field component at the spin of type α. Analogous expressions for the field components at the spins of type β are obtained by the obvious change of indexes. A reason for rejecting the cross-correlation contributions proportional to terms of the type $\mathrm{Sp}\{S_{zi}^{(\alpha)}(t)S_{zj}^{(\alpha)}\}$ in TCFs (11) and (12) is the structure of the coefficients of expansion in powers of time for these functions.

A coefficient at the zero power of time in the cross TCF is zero, while that at the autocorrelation TCF is unity. The ratio of the quadratic terms is determined by the ratio of lattice sums of the type [2]

$$\varepsilon = \sum_{i,j,k} b_{ik}^{(\alpha,\beta)} b_{jk}^{(\alpha,\beta)} b_{ij}^{(\alpha)2} \Big/ \sum_{i,j,k} b_{ik}^{(\alpha,\beta)} b_{ij}^{2(\alpha)} = S_3'/S_1S_1'. \quad (14)$$

For cubic lattices of the NaF and LiF types in an external magnetic field oriented along the crystallographic directions [100], [110], and [111], this ratio is very small (see Table 2 below). The coefficients at high powers of time contain more complicated lattice sums. However, the structure of the expansion coefficients for the cross TCFs is such that the corresponding lattice sums necessarily contain a loop of bonds, while the contributions to the autocorrelation functions involve the terms possessing a treelike (loopless) structure. Estimates and numerical calculations showed [2] that, for large Z ($d \longrightarrow \infty$), an increase in the size or number of loops in the lattice sum only increases the difference of values of the lattice sums of the two types.

In the case of a heteronuclear spin–spin interaction described by Hamiltonian (1), a TCF of the product of more than two local field operators $h_i^{q(\alpha)}(t)$ represents a sum of the products of all possible pairwise averages (11) and (12) [2], analogous to those in the homonuclear case [12, 13].

This property, together with the independence of various contributions to the local field, is essentially a definition of the Gaussian stochastic process [15]. In the case when the time variation of contributions to the local field is caused by the spin–spin interaction (1), it is the rejection of lattice loops that leads to the independence of fluctuations. The condition of absolute independence of the time fluctuations of the local field components is, strictly speaking, valid only for the model lattices in which the infinite limit with respect to the number of nearest neighbors Z is achieved in the infinite limit of the space dimension (i.e., in hypercubic lattices of large dimensionality), rather than in the limit of, for example, infinite radius of the interaction.

Nevertheless, we will assume (based on the aforementioned numerical estimates and the results of lattice sum calculations in various cases) that the local magnetic field acting upon a spin in a cubic lattice is a three-dimensional (three-component) Gaussian stochastic field.

Equations for the TCF of a spin precessing in an arbitrary Gaussian stochastic field were derived in [13]. Taking into account that, in a heteronuclear spin system, such equations should be written for the spins of each type, we arrive at a system of integral equations

$$\frac{d}{dt}\Gamma_q^{(\alpha)}(t) = -\int_0^t K_q^{(\alpha)}(t-t')\Gamma_q^{(\alpha)}(t')dt', \quad (15)$$

where index α enumerates the spin subsystems and index $q = 0, z, \ldots$ refers to the spin projections (0 is the x component of the spin; the x component of the total spin enters into these equations without index). The memory function $K_q^{(\alpha)}(t)$ is determined by an infinite series [2]. Without losing generality, below we will assume the system to contain the spins of only two types.

Equations (15) were derived for the interaction with an arbitrary magnetic anisotropy. Using axially symmetric Hamiltonian (1), with the purpose of taking into account predominantly the longitudinal local field component, significantly improves convergence of the series for the memory function in a homonuclear spin system [5, 6, 16–18].

In connection with this, by analogy with [2, 6, 17], it is expedient to transform Eqs. (15) to the following form:

$$\begin{gathered}\Gamma_q^{(\alpha)}(t) = G^{(\alpha)}(t) - \int_0^t G^{(\alpha)}(t-t') \\ \times \int_0^{t'} \{K_q^{(\alpha)}(t'') - K_{q0}^{(\alpha)}(t'')\}\Gamma_q^{(\alpha)}(t'-t'')dt'dt'',\end{gathered} \quad (16)$$

where

$$G^{(\alpha)}(t) = \exp\left\{-\sum_\beta \Delta_{\alpha\beta}^2 \int_0^t\int_0^{t'} \Gamma_z^{(\beta)}(t'')dt'dt''\right\} \quad (17)$$

is the TCF of a spin rotating in a local field possessing only the longitudinal component (analogous to that employed in the Anderson model [19]). Equations for $\Gamma_z^{(\alpha)}(t)$ are used without transformations.

With a view to describing the experimental results for LiF crystals, we explicitly write the main equations for a nuclear spin subsystem of this crystal in the lowest approximation for the memory function:

$$\begin{gathered}G^{(1)}(t) = \exp\left\{-\Delta_{11}^2\int_0^t\int_0^{t'} dt'dt''\Gamma_z^{(1)}(t'')\right. \\ \left. -\Delta_{12}^2\int_0^t\int_0^{t'} dt'dt''\Gamma_z^{(2)}(t'')\right\},\end{gathered}$$

$$
\begin{aligned}
G^{(2)}(t) &= \exp\left\{-\Delta_{21}^2\int_0^t\int_0^{t'} dt'dt''\Gamma_z^{(1)}(t'') - \Delta_{22}^2\int_0^t\int_0^{t'} dt'dt''\Gamma_z^{(2)}(t'')\right\},\\
\Gamma_z^{(1)}(t) &= 1 - \frac{1}{2}\Delta_{11}^2\int_0^t\int_0^{t'} dt'dt''\Gamma_0^{(1)2}(t'-t'')\Gamma_z^{(1)}(t''),\\
\Gamma_z^{(2)}(t) &= 1 - \frac{1}{2}\Delta_{22}^2\int_0^t\int_0^{t'} dt'dt''\Gamma_0^{(2)2}(t'-t'')\Gamma_z^{(2)}(t''),\\
\Gamma_0^{(1)}(t) &= G^{(1)}(t) - \frac{1}{4}\Delta_{11}^2\int_0^t dt' G^{(1)}(t-t')\\
&\times\int_0^{t'} dt''\Gamma_0^{(1)}(t'-t'')\Gamma_z^{(1)}(t'-t'')\Gamma_0^{(1)}(t''),\\
\Gamma_0^{(2)}(t) &= G^{(2)}(t) - \frac{1}{4}\Delta_{22}^2\int_0^t dt' G^{(2)}(t-t')\\
&\times\int_0^{t'} dt''\Gamma_0^{(2)}(t'-t'')\Gamma_z^{(2)}(t'-t'')\Gamma_0^{(2)}(t'').
\end{aligned}
\tag{18}
$$

Here and in what follows, indexes 1 and 2 refer to lithium and fluorine nuclei, respectively.

Table 1 presents the quantities $\Delta_{\alpha\beta}^2$ in Eqs. (11)–(13) and (18) calculated for three main orientations of the constant magnetic field relative to the principal crystallographic axes using the lattice sums taken from [20]. The moments were calculated with neglect of the ^{6}Li isotope. Consequences of the presence of a small amount of ^{6}Li nuclei (natural occurrence, 7.2%) in a LiF crystal will be discussed below.

3. THE SHAPE OF THE NMR ABSORPTION SPECTRA IN LITHIUM FLUORIDE

The orientation dependence of the NMR absorption spectra of the β-active ^{8}Li nuclei in a LiF crystal was thoroughly studied in [10, 11]. An extremely high sensitivity of the method employed allowed the shape of wings of the spectra to be determined with high precision: this point was given special attention in the papers cited.

Table 1. Average squares of the homo- and heteronuclear contributions to the longitudinal local fields in LiF

Quantity	$\mathbf{H}_0 \parallel [111]$	$\mathbf{H}_0 \parallel [110]$	$\mathbf{H}_0 \parallel [100]$
Δ_{22}^2, (rad/s)^2	2838×10^6	2581×10^6	1809×10^6
$\Delta_{21}^2/\Delta_{22}^2$	0.3065	1.9559	9.7209
$\Delta_{12}^2/\Delta_{22}^2$	0.0613	0.3912	1.9444
$\Delta_{11}^2/\Delta_{22}^2$	0.1455	0.1455	0.1455

Since the sensitivity of the β-NMR spectroscopy allows the absorption signals to be measured for a single nucleus (under real experimental conditions, the number of ^{8}Li nuclei simultaneously present in the sample does not exceed 10^8), we will assume that a local field acting upon the ^{8}Li nuclei has only the longitudinal component. Therefore, the ^{8}Li NMR absorption signal is adequately determined by the Anderson TCF of the type $G^{(1)}(t)$ from system (18), since the probability of finding nearest neighbors of the same type is very small (due to the small concentration of ^{6}Li and ^{8}Li in the sample).

At the same time, we should take into account a difference between the time scales of decay of the TCF $G^{(1)}(t)$ for ^{7}Li and ^{8}Li. The ratio of the $\Delta_{\alpha\beta}^2$ values for these nuclei is

$$
\mu = \gamma_{^8\mathrm{Li}}^2/\gamma_{^7\mathrm{Li}}^2 \approx 0.14. \tag{19}
$$

The same value is obtained for the ratio of exponents of TCFs in system (18). From this, we derive a simple relationship between the correlation functions proper:

$$
G_{^8\mathrm{Li}}^{(1)} = [G_{^7\mathrm{Li}}^{(1)}(t)]^{\mu}. \tag{20}
$$

For calculating the β-NMR absorption spectra, we numerically solved the system of interrelated integral equations (18). The solution was obtained by the iteration method, with the initial approximation for $G^{(\alpha)}(t)$ functions taken in the following form:

$$
\begin{aligned}
G_0^{(1)}(t) &= \exp\{-\Delta_{11}^2t^2/2 - \Delta_{12}^2t^2/2\},\\
G_0^{(2)}(t) &= \exp\{-\Delta_{21}^2t^2/2 - \Delta_{22}^2t^2/2\}.
\end{aligned}
\tag{21}
$$

Upon substituting functions (21) into system (18), we may start the iterative procedure. For large times, a difference between the decay rates of TCFs of the transverse and longitudinal spin components may lead to the instability development. When the signs of instability appeared, the iterative procedure was terminated and new initial TCFs were taken equal to the $G^{(\alpha)}(t)$ functions numerically calculated by that time instant, which significantly increased the time interval for stable calculations. In this way, the TCFs were calculated over the time interval from 0 to 800 μs for the three main orientations of the constant magnetic field relative to the

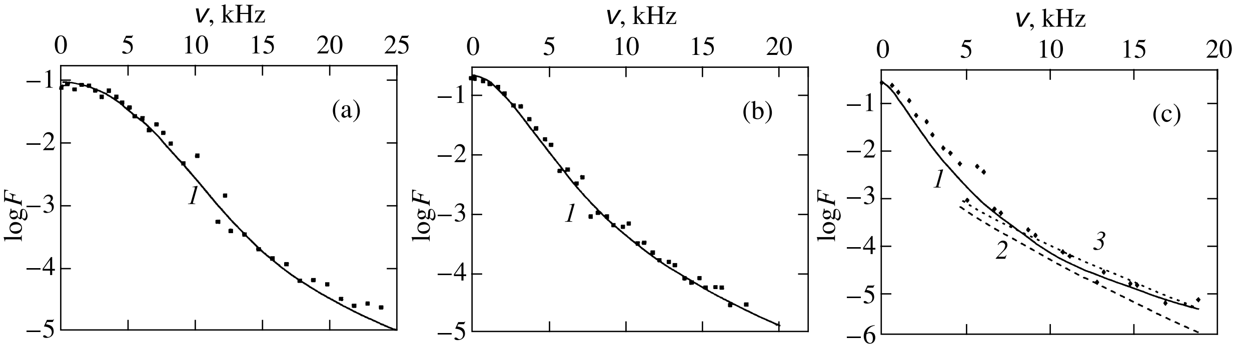


Fig. 1. The β-NMR spectra of ^{8}Li nuclei in LiF for the external magnetic field oriented along (a) [100], (b) [110], and (c) [111] crystallographic directions. Points present the experimental data taken from [10, 11]; solid curves*1* show the Fourier images of the right-hand part of formula (20) for the TCFs calculated using system (18); dashed curves *2* and *3* are calculated using an asymptotic formula from [2] with $\tau_0 = 55$ and 45 μs, respectively.

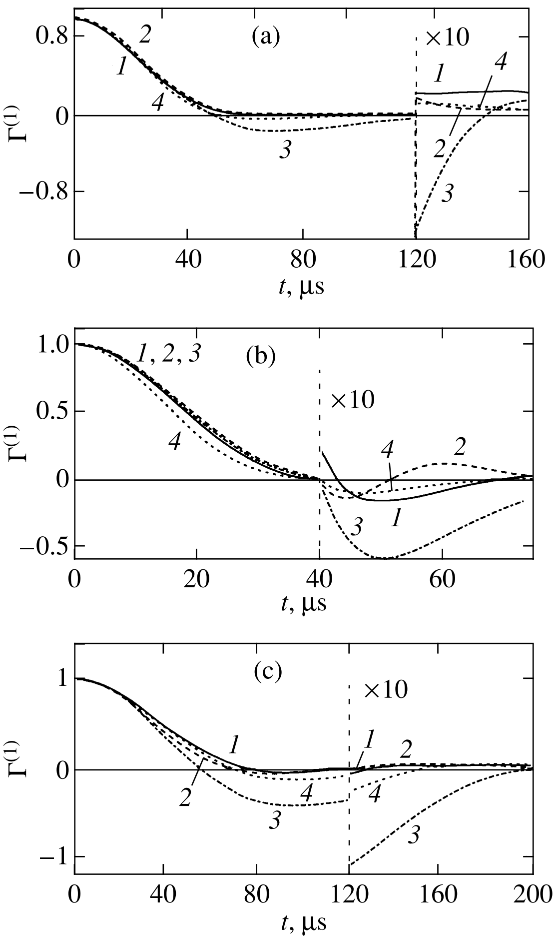


Fig. 2. Free precession signals of ^{7}Li nuclei in LiF for $\mathbf{H}_0 \parallel$ [100] (a), [110] (b), and [111] (c). Curves *1* present the experimental data from [22]; curves (*2*–*4*) show the results of calculations using (*2*) formula (28), (*3*) formula (25) with $A^2 = 1.25\Delta_{11}^2$, and (*4*) formula (25) with $A^2 = 0.5\Delta_{11}^2$. Since some of the curves virtually coincide, amplitudes behind the vertical dashed lines are increased by a factor of 10.

principal crystallographic axes. Then the NMR absorption spectra were numerically calculated by conducting the Fourier transformation. The results of calculations are presented in comparison to the experimental data in Fig. 1. As can be seen, the calculated shape of wings in the TCF $G^{(1)}(t)$ spectra fits well to the experimentally measured values [10, 11].

Unfortunately, no detailed experimental investigations of the ordinary NMR absorption spectrum shape (including wings) were performed on a level comparable to that achieved for the homonuclear system of spins [8, 21] in a simple cubic lattice of the CaF$_2$ type. However, the signals of free precession from ^{7}Li and fluorine nuclei measured over a time interval containing 2–3 zeros for the main orientations of the constant magnetic field were reported in [22]. The observed signals were described by the correlation functions $\Gamma(t)$. For calculating these functions, system (18) must be supplemented by a pair of the corresponding equations. The necessary equations can be obtained from the two last equations in system (18) by substituting 5/4 for the coefficient at the integral. Figures 2 and 3 show the free precession signals calculated using these equations and system (18) in comparison to the experimental data taken from [22]. As is seen, there is only a qualitative agreement between the experiment and our calculation. This result is by no means surprising because, in contrast to the wings the character of which is determined predominantly by the multispin combination processes involving only spins in a far environment [6, 17], the central part of the NMR spectrum (responsible for oscillations in the free precession signal) is determined by interactions with the nearest neighbors, the number of which Z is finite rather than infinite, as was assumed in deriving system (18). For example, the spectrum of the longitudinal component of the local field created by such neighbors is binomial (rather than Gaussian) with considerable discreteness (ignoring the transverse flip-flop interaction [23, 24]).

For this reason, the “free-particle Green’s functions” introduced in [16, 17] have to be replaced for the nearest neighbors, in the simplest case (see the considerations below), by TCFs corresponding to a binomial distribution

$$G_0^{(\alpha)}(t) = R'_{\alpha\alpha}(t)R'_{\alpha\beta}(t), \tag{22}$$

$$R'_{\alpha 2}(t) = \prod_{f\neq 0}\cos\frac{b_{0f}^{(\alpha,2)}t}{2}, \tag{23}$$

$$R'_{\alpha 1}(t) = \prod_{l\neq 0}\left(\frac{1}{2}\cos\frac{b_{0l}^{(\alpha,l)}t}{2}+\frac{1}{2}\cos\frac{3b_{0l}^{(\alpha,l)}t}{2}\right), \tag{24}$$

where subscript 0 refers to a certain fixed spin type α and indexes *f* and *l* run over values corresponding to the fluorine and lithium positions, respectively.

By analogy with what was done for the homonuclear spin system [6, 17], Eqs. (18) derived for a heteronuclear system can be linearized. For the free precession signals, we obtain an equation of the pairwise interaction model [5, 6, 17, 23, 25]:

$$\Gamma^{(\alpha)}(t) = G_0^{(\alpha)}(t) - A^2\int_0^t G_0^{(\alpha)}(t-t')\int_0^{t'} G_0^{(\alpha)}(t'-t'')\Gamma^{(\alpha)}(t'')dt'dt''. \tag{25}$$

It should be emphasized that the nonlinearity of system (18) was of principal significance for correct description of the wings of the TCF spectra [2, 6, 17].

Generally speaking, the binomial spectral components of TCFs (23) and (24) must be broadened as a result of the transverse interactions of spins creating a local field at the selected spin. Methods of taking into account the effect of the transverse interaction on the longitudinal local field component are considered below (see the section devoted to the three-spin cross-relaxation, Eqs. (43)–(46)).

The simple Eq. (25) provides for a quite good agreement with experiment when a binomial distribution is used as the initial spectrum in the iterative procedure. However, the agreement can be improved by taking into account corrections to the irreducible operator $G^{(\alpha)}(t)$ due to the four-vertex irreducible diagrams [6, 17]. Since the structure of these corrections is rather complicated [6] and depends significantly on the selection of the initial approximation, we will employ a modified equation for the free precession signal in LiF with a much simpler form of corrections:

$$\Gamma^{(\alpha)}(t) = G_0^{(\alpha)}(t) - k^{\alpha}\int_0^t R_{\alpha\beta}(t')\dot{R}_{\alpha\alpha}(t')\Gamma^{(\alpha)}(t-t')dt', \tag{26}$$

$$\dot{R}_{\alpha\alpha}(t) = d/dt(R_{\alpha\alpha}(t)), \quad R_{\alpha\beta}(t) = R'_{\alpha\beta}(t),$$

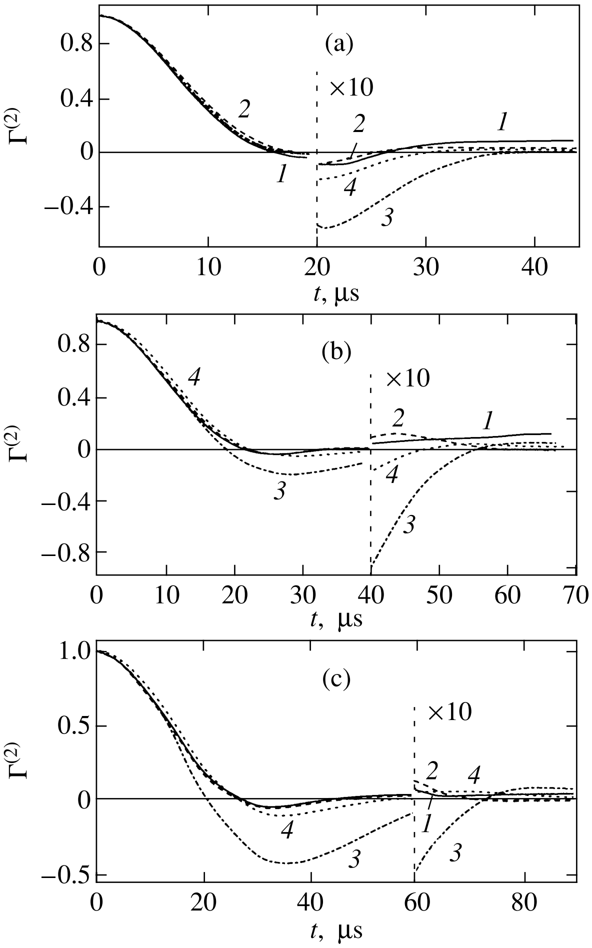


Fig. 3. Free precession signals of ^{19}F nuclei in LiF. Notation is the same as in Fig. 2, except for $A^2 = 1.25\Delta_{22}^2$ (*3*) and $0.5\Delta_{22}^2$ (*4*).

where $R_{\alpha\alpha}(t)$ is obtained from $R'_{\alpha\alpha}(t)$ entering into (22) by substituting $\lambda b_{0k}^{(\alpha\beta)}$ for $b_{0k}^{(\alpha\beta)}$. Correctness of the second moment of the spectrum is ensured by a strict relationship between the parameters $k^{(\alpha)}$ and λ in Eq. (26):

$$k^{(\alpha)} = \frac{9}{4\lambda^2-1}. \tag{27}$$

Equation (26), which was originally used for description of the free precession signal in [18, 26, 27], satisfactorily described this signal in CaF$_2$ with λ = 1.225 and 1.19.

For determining the corrections to Eq. (26), we represent (by analogy with what was done in the case of a homonuclear system [8]) $G^{(\alpha)}(t)$ as a solution to Eq. (15) with the nucleus $Q^{(\alpha)}(t)$ determined by the series $K^{(\alpha)}(t)$ with rejected terms due to the vertices corresponding to the interaction between transverse spin projections and the longitudinal projections multiplied by the factor λ. Combining the Laplace transforms of Eq. (15) and (26), we obtain

$$\begin{aligned}\Gamma^{(\alpha)}(t) &= R_{\alpha\beta}(t)R_{\alpha\alpha}(t)\\ &+ k^{(\alpha)}\int_0^t R_{\alpha\beta}(t')\dot{R}_{\alpha\alpha}(t')\Gamma^{(\alpha)}(t-t')dt'\\ &- \int_0^t \Phi^{(\alpha)}(t-t')\Gamma^{(\alpha)}(t')dt',\end{aligned} \tag{28}$$

where

$$\begin{aligned}\Phi^{(\alpha)}(t) &= \int_0^t \{K^{(\alpha)}(t') - Q^{(\alpha)}(t')\}G^{(\alpha)}(t-t')dt'\\ &+ k^{(\alpha)}R_{\alpha\beta}(t)\dot{R}_{\alpha\alpha}(t).\end{aligned}$$

Equation (28) is formally exact, provided that complete series are retained for $K^{(\alpha)}(t)$ and $Q^{(\alpha)}(t)$, and allows the required corrections to be determined. Acting as in the case of the homonuclear system [8], we obtain for the contribution of four-vertex diagrams to $\Phi^{(\alpha)}(t)$

$$\begin{aligned}\Phi^{(\alpha)}(t) &= \frac{k_4\Delta_{\alpha\alpha}^2}{\lambda^2}\int_0^t (t-t')R_{\alpha\beta}(t-t')R_{\alpha\beta}(t')\\ &\times\{\dot{R}_{\alpha\alpha}(t-t')R_{\alpha\alpha}(t') + R_{\alpha\alpha}(t-t')\dot{R}_{\alpha\alpha}(t')\}dt'.\end{aligned} \tag{29}$$

Appendix B gives expressions that can be used for comparing the exact contributions to the fourth moment of the spectrum to the approximate contributions determined using Eq. (28).

The parameter k_4, ensuring correctness of the homonuclear contribution to the fourth moment of the spectrum, is $k_4 = 9S_3/8S_1^2$, where S_3 and S_1 are the lattice sums calculated in [20] (see also Appendix B). Explicit expressions for the nucleus of Eq. (28) employed in the calculations for three main orientations of the external magnetic field are presented with the corresponding constants in Appendix A.

As can be seen in Fig. 2, the solution of Eq. (28) provides for a rather good agreement with experimental data for the free precession signal [22].

4. THE RATE OF THREE-SPIN CROSS-RELAXATION IN LITHIUM FLUORIDE

Here, we will consider the process of three-spin cross-relaxation in LiF between the Zeeman subsystems of ^{7}Li and ^{19}F nuclei depending on the applied constant magnetic field H_0. This process was experimentally studied by Pershan [28]. According to the thermodynamic theory [29, 30], employing methods of perturbation theory (the applicability of which improves with increasing detuning), the cross-relaxation rate is described by the expression

$$1/T_{12} = Af(\omega), \tag{30}$$

where

$$f(\omega) = \frac{1}{\pi}\int_0^\infty F(t)\cos\omega t dt, \tag{31}$$

$$F(t) = \mathrm{Sp}\{V^+(t)V^-\}/\mathrm{Sp}(V^+V^-), \tag{32}$$

$$V^+ = \sum_{\substack{i,j,k\\ i\neq j}} C_{ij/k}S_{+i}^{(1)}S_{+j}^{(1)}S_{-k}^{(2)} + \sum_{i,k} C_{ii/k}S_{+i}^{(1)2}S_{-k}^{(2)}. \tag{33}$$

Here V^+ is the operator responsible for the cross-relaxation processes [28, 30] in the second order of perturbation theory and V^- is the Hermitian conjugate to V^+. The coefficients $C_{ij/k}$ represent the products of various coefficients of the nonsecular part of the dipole–dipole interaction inducing the process of simultaneous reversal of two ^{7}Li nuclei and one fluorine nucleus. In the notation of [28], these coefficients are as follows:

$$\begin{aligned}C_{ij/k} &= \alpha_{ij/k} + \beta_{ij/k}, \quad C_{ii/k} = \beta_{ii/k},\\ \alpha_{ij/k} &= 2\alpha_1E_{ij}D'_{ik} - 2\alpha_3C_{ij}B'_{ik},\\ \beta_{ij/k} &= \frac{1}{2}\alpha_2(B'_{ik}C'_{jk} + B'_{jk}C'_{ik}),\\ \alpha_1 &= \frac{1}{2\pi}\left(\frac{1}{2\omega_1} + \frac{1}{\omega_2}\right) = \frac{1}{2\hbar H_0}\frac{\gamma_2 + 2\gamma_1}{2\gamma_1\gamma_2},\\ \alpha_2 &= \alpha_3 = \frac{1}{2\hbar H_0}\frac{\gamma_2}{\gamma_1(\gamma_2-1)},\end{aligned} \tag{34}$$

where $\omega_1 = \gamma_1H_0$ and $\omega_2 = \gamma_2H_0$. The quantity A in formula (30) is the square modulus of the matrix perturbation element (33). The frequency ω in Eq. (31) is determined by the relationship

$$\omega = (\gamma_2 - 2\gamma_1)H_0, \tag{35}$$

and $f(\omega)$ is the form function of the cross-relaxation spectrum, which is normalized to unity with respect to the area under curve. The time dependence of operator $V^+(t)$ in formula (32) is determined by the secular part of the dipole–dipole interaction (1).

Upon rejecting the lattice sums with loops and other small terms, we arrive at an expression [2]

$$F_3(t) = \Gamma_0^{(2)}(t)\Gamma_0^{(1)2}(t), \tag{36}$$

which is valid in the same approximation as is system (19). The TCF spectrum calculated by Eq. (36) with the aid of system (18) is presented in Fig. 4; as is seen, this

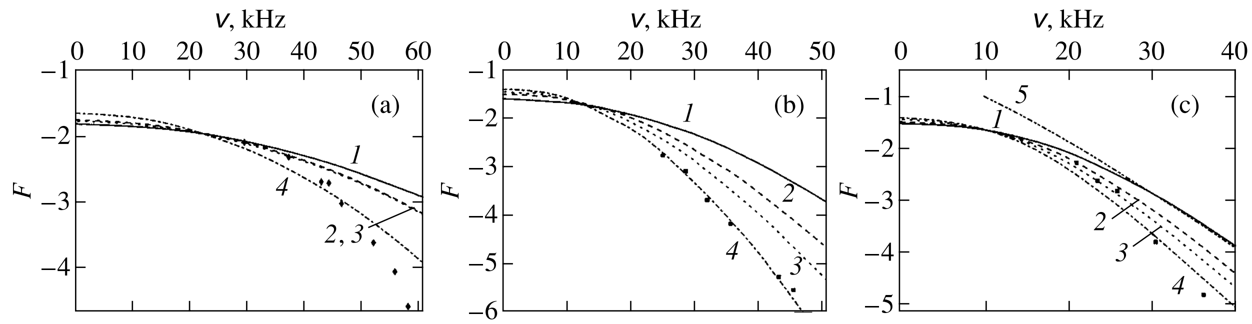


Fig. 4. Wings of the three-spin cross-relaxation spectra of LiF calculated for the external magnetic field oriented along (a) [100], (b) [110], and (c) [111] crystallographic directions. Points present the experimental data from [28]; curves *1* show the Fourier spectra approximated by formula (36) with TCF calculated using system (18); curves *2–4* are the same spectra as *1*, but calculated with various sets of scaling factors (δ) and parameters (ν and ν'): (a) δ = 0.8 (*2*), 0.6 (*3*), 0.1 (*4*) and ν = ν'= 1; (b) δ = 0.64 (*2*), 0.5 (*3*), 0.125 (*4*) and ν = ν'= 1; (c) δ = 0.8, ν = 0.79, ν'_1 = 0.71 (*2*); δ = 0.7, ν = ν'= 1 (*3*); δ = 0.6, ν = ν' = 1 (*4*); curve *5* in (c) corresponds to the asymptotic formula from [2] with τ_0 = 45 μs.

approximation provides for only a qualitative agreement with experiment [28].

The main reason for the lack of a quantitative fit of this theory to experiment is that the three-center correlation function (32) exhibits an extremely high sensitivity (related to spatial correlations of the local fields) with respect to details of the crystal lattice topology (this sensitivity is much more pronounced as compared to that of the one-center function $G^{(1)}(t)$ in system (18)). For a limited number of nearest neighbors, the description of NMR spectra for a particular LiF lattice in terms of Eq. (19) requires certain corrections.

First, it is necessary to modify the second moment of the resulting spectrum as indicated by Pershan [28]. Let us study how much it is necessary to calculate corrections in the simplest case of a constant longitudinal field, that is, when the analysis is restricted to only the H_{zz} interaction in Hamiltonian (1):

$$\begin{aligned} F_3^{(0)}(i, j, k|t) &= \mathrm{Sp}\{\exp(iH_{zz}t)S_{-k}^{(2)}S_{+i}^{(1)}S_{+j}^{(1)} \\ &\quad\times\exp(-iH_{zz}t)S_{+k}^{(2)}S_{-i}^{(1)}S_{-j}^{(1)}\} \\ &\quad\times[\mathrm{Sp}(S_{+k}^{(2)}S_{-k}^{(2)})\mathrm{Sp}(S_{-i}^{(1)}S_{+i}^{(1)})]^{-1} \\ &= G^0(i, j, k, F|t)G^0(i, j, k, L|t)G^0(i, j, k, |t), \end{aligned} \tag{37}$$

where

$$G^0(i, j, k, F|t) = \prod_{f\neq k}\cos\left\{\frac{1}{2}(b_{fl}^{(1,2)} + b_{fj}^{(1,2)} - b_{fk}^{(2)})t\right\},$$

$$\begin{aligned} G^0(i, j, k, L|t) = \prod_{l\neq i, j}\left\{\frac{1}{2}\cos\left[\frac{1}{2}(b_{li}^{(1)} + b_{lj}^{(1)} - b_{lk}^{(1,2)})t\right]\right. \\ \left. + \frac{1}{2}\cos\left[\frac{3}{2}(b_{li}^{(1)} + b_{lj}^{(1)} - b_{lk}^{(1,2)})t\right]\right\}, \end{aligned}$$

$$\begin{aligned} G^0(i, j, k, |t) &= \left\{\frac{1}{2} + \frac{1}{2}\cos[(b_{ik}^{(1,2)} - b_{ij}^{(1)})]t\right\} \\ &\quad\times\left\{\frac{1}{2} + \frac{1}{2}\cos[(b_{jk}^{(1,2)} - b_{ij}^{(1)})]t\right\}. \end{aligned}$$

In Eq. (37), indexes *f* and *k* run over values corresponding to the positions of fluorine nuclei, while *l*, *i*, and *j* refer to the lithium positions. The dipole–dipole interaction coefficients $\{b_{mn}(k)\}$ are obviously expressed through the above-introduced coefficients of Hamiltonian (1). These coefficients are proportional to the values introduced by Pershan [28]:

$$b_{fi}^{(1,2)} = A'_{fi}, \quad b_{li}^{(1)} = 2A_{li}, \quad b_{jk}^{(2)} = 2A''_{fk}.$$

In this approximation, the resulting cross-relaxation spectrum is obtained by summing all three-center functions $F_3^0(i, j, k|t)$ over the lattice with the weights $C_{ij/k}(C^*_{ij/k} + C^*_{ji/k})$, where an asterisk denotes the complex conjugate. Of course, the presence of a transverse interaction and local field fluctuations will result in replacing cosines in (37) by much more complicated constructions varying depending on the particular triads. However, the aforementioned correlation effect is qualitatively well illustrated by relationship (37).

The approximation of uncorrelated local fields implies that we ignore in (37) both the interaction between spins *i*, *j*, *k* and the interaction of these spins with nearest neighbors. Since each of the spins *i*, *j*, *k* in this approximation is surrounded by a different set of neighbors, the TCFs $G^0(i, j, k, F|t)$ and $G^0(i, j, k, L|t)$ split into the products of separate contributions due to spins *i*, *j*, *k*:

$$F_3^0(i, j, k|t) = R_{11}^2(t)R_{12}^2(t)R_{22}(t)R_{21}(t), \tag{38}$$

where $R_{\alpha\beta}(t)$ are the functions determined above (see also Appendix B).

Table 2. Parameters of the local field fluctuation decrease and the corresponding lattice sum ratios

Quantity	$\mathbf{H}_0 \parallel [100]$		$\mathbf{H}_0 \parallel [110]$		$\mathbf{H}_0 \parallel [111]$	
	fcc	fcc'	fcc	fcc'	fcc	fcc'
S_2/S_1^2	0.068	–	0.225	–	0.066	–
$S_2'/S_1'^2$	–	0.239	–	0.182	–	0.126
S_3/S_1^2	0.226	–	0.099	–	0.144	–
$S_3'/S_1'S_1$	–	0.071	–	0.087	–	0.288
ν	0.71	–	0.68	–	0.79	–
ν'	–	0.93	–	0.91	–	0.71

Taking into account the spatial correlation effects for the ratio of second TCF moments (37) and (38), we obtain (see Appendix B)

$$\delta = M_2/M_2^{(n)} \approx 0.64–0.86$$

depending on the external field orientation ($M_2^{(n)}$ is the second moment determined with the corrections described above). As can be seen in Fig. 3, introduction of the scaling factor significantly improves the agreement of theory and experiment.

Besides changing the second moment of the TCF spectrum, the aforementioned correlation effect leads to a decrease in the magnitude of fluctuations in the local field acting upon the selected spin in a real lattice [8] as compared to that determined in the approximation of an infinite number of neighbors. For elucidating this effect, let us consider the TCF of a longitudinal local field component in a lattice of finite dimension:

$$\langle h_i^{z(\alpha)}(t)h_i^{z(\alpha)}(0)\rangle = \frac{1}{3}\sum_{\beta,\, j\neq i}{}' S^{(\beta)}(S^{(\beta)}+1)(b_{ij}^{(\alpha\beta)})^2\Gamma_{zj/i}^{(\beta)}(t)$$
$$+\frac{1}{3}\sum_{\substack{j\neq i\\ k,\,\beta}}{}' S^{(\beta)}(S^{(\beta)}+1)b_{ij}^{(\alpha\beta)}b_{ik}^{(\alpha\beta)}\Gamma_{zjk/i}^{(\beta)}(t). \tag{39}$$

Table 3. Parameters of the second moment decrease due to the correlation effects and the corresponding lattice sum ratios

Quantity	$\mathbf{H}_0 \parallel [111]$	$\mathbf{H}_0 \parallel [110]$	$\mathbf{H}_0 \parallel [100]$
C	1.14	1.35	2.35
C'	20.71	2.74	0.0742
E	5.276	5.70	7.708
$S_3''/S_1'^2$	0.0378	0.148	0.0499
$S_3'''/S_1'^2$	0.741	0.179	–0.0410
δ	0.73	0.64	0.86

The first term in (39) contains the autocorrelation function of the z projection of spin j for the β-type nuclei. The second term contains the cross TCF of two spins j and k of the same type β, since the flip-flop transitions between spins of various types were ignored. The slash in the subscript denotes exclusion of the interaction with the selected spin i. Expression (12) for the main approximation is obtained from Eq. (39) upon excluding restrictions concerning the interaction and rejecting the second term.

For evaluating the effect, let us expand (39) in powers of time

$$\langle h_i^{z(\alpha)}(t)h_i^{z(\alpha)}(0)\rangle = \Delta_{\alpha\alpha}^2 - \frac{t^2\Delta_{\alpha\alpha}^4}{4}\left\{1-\frac{S_2}{S_1^2}-\frac{S_3}{S_1^2}\right\}$$
$$+\Delta_{\alpha\beta}^2 - \frac{t^2\Delta_{\alpha\beta}^2\Delta_{\beta\beta}^2}{4}\left\{1-\frac{S_3'}{S_1S_1'}\right\}O(t^4), \tag{40}$$
$$\Delta_{\alpha\alpha}^2 = \frac{4(\gamma_\alpha\hbar)^4}{3a^6}S^{(\alpha)}(S^{(\alpha)}+1)S_1,$$
$$\Delta_{\alpha\beta}^2 = \frac{4(\gamma_\alpha\gamma_\beta\hbar^2)^2}{3a^6}S^{(\beta)}(S^{(\beta)}+1)S_1',$$

where a is the lattice constant (see Appendix B). The terms S_3, S_3' characterize the correlation of contributions to the local field. Estimates of the correlation effect with respect to the lattice sums for the second moment are presented in Tables 2 and 3. Notations for the lattice sums $\{S_i\}$ are taken from [20]. The effect of the fluctuation decrease is most readily taken into account by retaining the form of the field correlator (12) and taking TCF $\Gamma_z^\beta(t)$ in a power of ν_β ($\nu_\beta < 1$):

$$g_z^{(\alpha)}(t) = \Delta_{\alpha\alpha}^2[\Gamma_z^{(\alpha)}(t)]^{\nu_\alpha} + \sum_\beta \Delta_{\alpha\beta}^2[\Gamma_z^{(\beta)}(t)]^{\nu_\beta}. \tag{41}$$

The exponents in Eq. (41) can be determined by equating the coefficients at t^2 in the TCF expansions (41) and (40):

$$\nu = \nu_\alpha = 1 - \frac{S_2}{S_1^2} - \frac{S_3}{S_1^2}, \quad \nu' = \nu_\beta = 1 - \frac{S'_3}{S'_1 S_1}.$$

The results of calculations of the parameter of fluctuation decrease are presented in Table 2.

Introduction of the scaling factor δ and/or the exponent ν describing the local field fluctuation decrease hardly complicates calculation of the spectra as compared to the procedure of solving the initial system (18). As can be seen from the spectra calculated for the field oriented along [110] and [111] in Fig. 3, this modification of the basic theory leads to a good coincidence with the experimental results. However, description of the wings for the field oriented in the [100] direction requires further correction taking into account the crystal structure in more detail.

As demonstrated in Appendix B, the main contribution to the cross-relaxation rate is due to the spin triads in Eq. (37) in which two lithium nuclei are connected to fluorine. For the [100] field orientation, the main contribution in the lattice sum (B.3) is due to the triads of nearest neighbors with the interaction B'_{ik} directed along the [100] axis and the interaction C'_{kj} along the cube diagonal. For the external field oriented along the [100] direction, the local field from the nuclei of other type significantly exceeds the homonuclear contribution. In this heteronuclear interaction, 98% of the contribution to the second moment is provided by the interaction with six nearest neighbors. For this reason, the main correlation effect in (37) consists in excluding the heteronuclear interaction with spin triad from (38). Taking these considerations into account, we replace expression (38) by the following:

$$F_3(t) = R_{11}^2(t) R'_{12}(t) R_\Delta(t) R_{22}(t) R'_{21}(t). \qquad (42)$$

The so-modified functions $\{R(t)\}$ are determined below (see formulas (43), (45), and (47)). The homonuclear contributions to the longitudinal local field component are still determined using TCFs in the form of Eq. (17). For the contributions of lithium nuclei to the local field at fluorine nuclei, let us replace the component $G^{(2)}(t)$ due to cross interaction by the product of contributions due to six neighbors (by analogy with formula (24)) and exclude the "forbidden" contributions from this product. As a result, we obtain

$$R'_{21}(t) = (1/2)^5 [F_{11}(t) + F_{13}(t)]^4 (F_{21} + F_{23}). \qquad (43)$$

For calculating TCFs $\{F_{ql}(t)\}$ according to (39) [where $q = 2$ for the spins whose internuclear vectors to the selected spin are parallel to the external field (there are two such spins) and $q = 1$ for the spins with perpendicular internuclear vectors (there are four such spins); the spin projection is $l/2 = 1/2$ or $3/2$], we use equations of the type (15):

$$\dot{F}_{ql} = -q^2 l^2 m^2 \int_0^t \Gamma_z^{(1)}(t - t') F_{ql}(t') dt', \qquad (44)$$
$$F_{ql}(0) = 1,$$

where $m^2 = \lambda \Delta_{21}^2/60$ is the dipole–dipole interaction coefficient renormalized to the factor λ and expressed in terms of the second moment.

For determining the contributions of fluorine nuclei to the local field at lithium nuclei, we have to exclude the local field contribution from a nucleus close to one of the lithium nuclei (that occurring on the opposite end of the cube diagonal, which has its own independent environment). Here, we obtain the expression

$$R'_{12}(t) = L_2^3(t) L_1^8(t), \qquad (45)$$

where

$$\dot{L}_q(t) = -q^2 b^2 \int_0^t \Gamma_{zq}^{(2)}(t - t') L_q(t') dt', \qquad (46)$$
$$b = \Delta_{12}^2/12.$$

As was noted in Section 3, Eqs. (43–(46) allow us to adequately take into account the local field fluctuations producing broadening of the components of binomial distribution (24), In particular, we obtain an expression describing the exponential wing of the NMR absorption spectrum. In the approximation of constant local fields, the TCFs $\{F_{ql}(t), L_q(t)\}$ represent usual cosines (Fourier images of the binomial distribution components) and $R_\Delta(t)$ is merely the function $G^0(i, j, k, | t)$ (for simplicity, only a strong interaction of the neighboring lithium and fluorine nuclei is retained):

$$R_\Delta(t) \approx 0.5\left[1 + \cos\left(\sqrt{(4/3)\Delta_{12}^2 t}\right)\right]. \qquad (47)$$

The autocorrelation functions $\Gamma_0^{(\alpha)}(t)$ in Eqs. (19) and (36) are close in form to the Anderson TCFs (17), differing only in width of the spectrum (by 1/4 fraction of a small homonuclear contribution to the second moment). To simplify the calculation, we may replace the TCF $\Gamma_0^{(\alpha)}(t)$ in (36) by the Anderson functions (43), (45), and (47) with corrected second moment.

Thus, we have arrived at Eq. (42). Note that, if the spatial correlation effects in (43) were not taken into account, this equation would evidently transform into (38). The results of calculations are presented in Fig. 3c. As can be seen, the experimental spectrum for the [100] orientation decays on the wing more rapidly than does the theoretical curve. A special feature of the interaction for this orientation consists in that (see Table 1)

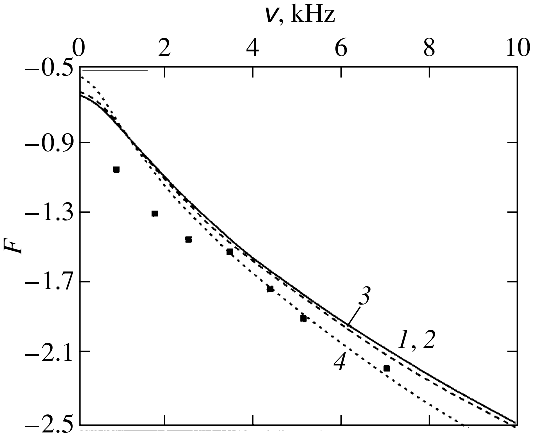


Fig. 5. A wing of the cross-polarization spectrum of LiF for the external magnetic field oriented along the [111] direction. Points present the experimental data from [34]; curve *1* show the Fourier spectra approximated by formula (49) with TCF calculated using system (19); curve *2* (virtually coinciding with *1*) and curves *3* and *4* are the same spectra as *1*, but calculated with scaling factors (δ) = 1 (*2*), 0.95 (*3*), 0.7 (*4*) and the parameters ν = 0.79, ν' = 0.71 (*2*), ν = ν' = 1 (*3*), and ν = ν' = 0.44 (*4*).

the heteronuclear interaction significantly exceeds the homonuclear one. The result is unexpected to even a greater extent, since even the constant longitudinal local field approximation (37) (essentially using a binomial distribution for description of the spectrum) yields a wider spectrum poorly coinciding with experiment. Apparently, the spectrum measured at this orientation acquires the properties of an inhomogeneously broadened spectrum. The cross-relaxation at the wing involves a few resonance triads, while the retardation is related to an insufficient rate of the spectral diffusion of the main fraction of nuclei toward these triads.

5. CROSS-POLARIZATION IN LITHIUM FLUORIDE

The single-spin cross-relaxation processes, usually referred to as the cross-polarization, seem to be studied most exhaustively [7, 32–36]. This is related to the fact that conditions for the observation of this effect are created much more readily than the conditions for, e.g., three-spin cross-relaxation.

As a rule, the cross-polarization is observed in a rotating coordinate system. Such experiments are performed with a crystal containing the isotope (with a low natural occurrence) of a nucleus possessing a nonzero magnetic moment. Engelsberg and Nordberg [34] studied the cross-polarization related to the interaction (3) between nuclei of the rare isotope ^{6}Li and the dipole–dipole reservoir of ^{7}Li and ^{19}F nuclei. Orienting the spin quantization axis of ^{6}Li along the blocking field (x axis in the rotating coordinate system), we obtain the perturbation in the following form:

$$V^{\pm} = \frac{1}{2}\sum_{\beta}\sum_{k,m} b_{km}^{(\alpha,\beta)} S_{\pm k}^{(\alpha)} S_{zm}^{(\beta)}. \tag{48}$$

For a sufficiently large amplitude of the blocking field, $\gamma_3 H_1 \gg \sqrt{\Delta_{31}^2 + \Delta_{32}^2}$ (index 3 denotes ^{6}Li nuclei), the rate of the process is determined by expressions (30) and (31) with the perturbation (48), $\omega = \gamma_3 H_1$, and $A = \pi(\Delta_{31}^2 + \Delta_{32}^2)$. The correlation function $F(t)$ according to formula (39) is expressed through the correlation function of the longitudinal local field at the ^{6}Li nucleus.

Ignoring the lattice sums with loops ($d \longrightarrow \infty$), we obtain

$$F(t) = \frac{\Delta_{12}^2 \Gamma_z^{(2)}(t) + \Delta_{11}^2 \Gamma_z^{(1)}(t)}{\Delta_{21}^2 + \Delta_{11}^2}. \tag{49}$$

In a real LiF lattice with the [111] field orientation, the loops and, hence, the cross-correlation functions produce a significant contribution (see Table 2). The magnitude of this contribution can be estimated using expansion (40).

A correction related to the cross TCF is readily introduced for a far wing of the function (49) determined by the nearest singular point on the imaginary time axis [2]. This singularity is in common for all TCFs related by Eqs. (18). The main parts of the TCF in the vicinity of this singularity are determined by the coefficients [2]

$$\Gamma_{z/k}^{(\alpha)} \approx \frac{C_{z\alpha}}{(\tau_0 + it)^{\zeta_\alpha}},$$
$$\Gamma_{zij/k}^{(\alpha)} \approx -\frac{(b_{ij}^{(\alpha)})^2 C_{z\alpha}}{\sum_j (b_{ij}^{(\alpha)})^2 (\tau_0 + it)^{\zeta_\alpha}}. \tag{50}$$

The quantities $C_{z\alpha}$, τ_0, and ζ_α are determined in [2, Table 2]. The form of coefficients in the right-hand part of formulas (50) follows from the equations for this correlation function and, as can be readily verified, satisfies the condition of conservation of the total z projection of the total spin

$$\Gamma_z^{(\alpha)}(t) + \sum_j \Gamma_{zij/k}^{(\alpha)}(t) = 1.$$

Substituting (50) into (39) and (31), we obtain

$$
\begin{aligned}
f(\omega) \approx \frac{\exp(-\tau_0|\omega|)}{\Delta_{12}^2 + \Delta_{11}^2}\Bigg\{ & \frac{\Delta_{12}^2 C_{z2}|\omega|^{\zeta_2 - 1}}{\Gamma(\zeta_2)}\left(1 - \frac{S_3'}{S_1 S_1'}\right) \\
& + \frac{\Delta_{11}^2 C_{z1}|\omega|^{\zeta_2 - 1}}{\Gamma(\zeta_1)}\left(1 - \frac{S_2}{S_1^2} - \frac{S_3}{S_1^2}\right)\Bigg\}.
\end{aligned}
\tag{51}
$$

A decrease in the coefficients leads to lowering of the wing as compared to the ideal case of the lattice of infinite dimension. However, the observed results are not adequately described by the asymptotic formula (51), because the experiment still did not reach the frequency range corresponding to the far wing of the spectrum. For this reason, we calculate the spectrum using (by analogy to the previous section, see Eq. (41)) expression (49) for $F(t)$ in the form of a sum of the autocorrelation functions $\Gamma_z^{(\alpha)}(t)$ in a power of $\nu_\alpha < 1$. Determining ν_α by the first term of the expansion (40) in powers of time, we obtain the values listed in Table 2. As can be seen in Fig. 5, this approach provides for a good agreement between theory and experiment for the wing of the spectrum.

However, when the amplitude of the blocking field decreases to a level comparable to the local fields at the ^{6}Li nuclei, the experimental spectrum shows a significant decrease in the cross-polarization rate as compared to the result calculated using perturbation theory. This is probably related to the fact that, in a weak field, only a part of the (rather than the whole) perturbation induces the cross-polarization transitions. Indeed, for a time on the order of T_2, the spin of ^{6}Li rotates about the effective field vector that is a sum of the field H_1 directed along the x axis of the rotating coordinate system and a local field directed along the z axis. The corresponding transient magnetization oscillations are well known [37]. After the establishment of a quasi-equilibrium, the spin of ^{6}Li is directed along the effective field making an angle below 90° to the z axis. Thus, rotations of the ^{6}Li spin are induced by a part of the local field perpendicular to the effective field, rather than by the whole local field. However, it should be noted that the local field consists of the components created by the ^{19}F and ^{7}Li nuclei. As is seen from the data in Table 1, the field of the ^{19}F nuclei is smaller than that of the ^{7}Li nuclei ($\Delta_{12}^2 < \Delta_{11}^2$) but varies at a much greater rate in time ($\Delta_{11}^2 \ll \Delta_{22}^2$) The rapidly fluctuating field generated by fluorine nuclei is not involved in the effective field formation and does not contribute to decrease in the magnitude of interaction (as it was in the experiments with CaF$_2$ [7]). At the same time, the field created by the ^{7}Li nuclei may well account for the effects under consideration.

6. CONCLUSION

In concluding, it should be noted that the approach developed above offers a rigorous microscopic theory for model lattices of infinite dimension. For real systems, the theory is essentially a self-consistent Gaussian fluctuating local field approximation. The proposed modification of the general theory allowed us to pass to real three-dimensional lattices, and the introduced corrections provided for an adequate description for almost all of the classical experiments with LiF crystals representing a heteronuclear spin system. Apparently, only description of the three-spin cross-relaxation spectra for the [100] field orientation requires some additional refinement.

A principal distinctive feature of the proposed approach is that all results necessary for the description (at least qualitative but, after small correction, quantitative as well) of the TCF spectra (both considered above and any other) immediately follow from the theory without additional assumptions. Therefore, the theory is closed. The proposed theory provides a common basis for interpretation of the TCF spectra observed by various methods, predicting an exponential decay on the wing with a common exponent for all TCF of a given crystal.

ACKNOWLEDGMENTS

The authors are grateful to V.V. Eremin and V.O. Zavel’skiĭ for their help in processing of the experimental data from [28], to F.S. Dzheparov and S.V. Stepanov for kindly providing the experimental β-NMR spectra, to F.S. Dzheparov for valuable advice made upon preliminary consideration of a part of this paper, and to M.A. Popov for his help in the lattice sum calculations for Table 3.

This study was supported by the Russian Foundation for Basic Research, project nos. 98-03-33176 and 99-02-18214.

APPENDIX A

Here we present explicit expressions for the memory function entering into Eq. (28) for various orientations of the external constant magnetic field relative to the crystallographic axes used in our calculations.

Fluorine nuclei

$\mathbf{H}_0 \parallel [100]$:

$$G^{(2)}(t) = R_{21}(t)R_{22}(t), \quad R_{22}(t)\exp\{-\lambda^2\Delta_{22}^2 t^2/2\},$$

$$R_{21} = \left(\frac{1}{2}\cos 6mt + \frac{1}{2}\cos 2mt\right)^2$$

$$\times\left(\frac{1}{2}\cos 3mt + \frac{1}{2}\cos mt\right)^4 \exp\left\{-\frac{1}{2}(1-d')\Delta_{21}^2 t^2\right\},$$

$$k_4 = \frac{9S_3}{8S_1^2} = 0.23\times\frac{9}{8}, \quad m^2 = \frac{d'\Delta_{21}^2}{60},$$

where $d' = 0.98$ is the fraction in the second moment due to the six nearest neighbors selected. Here and below, the dipole–dipole coefficients are expressed in fractions of the second moment given in Table 1.

$\mathbf{H}_0 \parallel [110]$:

$$G^{(2)}(t) = R_{21}(t)R_{22}(t),$$

$$R_{22}(t) = \cos^2 Mt\cos^2 2Mt\exp\left\{-\frac{1}{2}(1-d_c)\lambda^c\Delta_{22}^2 t^2\right\},$$

$$M^2 = \lambda^2 d_c\Delta_{22}^2/10, \quad d_c = 0.81,$$

$$R_{21} = \left(\frac{1}{2}\cos 3mt + \frac{1}{2}\cos mt\right)^4$$

$$\times\left(\frac{1}{2}\cos 6mt + \frac{1}{2}\cos 2mt\right)^2 \exp\left\{-\frac{1}{2}(1-d')\Delta_{21}^2 t^2\right\},$$

$$m^2 = \frac{d'\Delta_{21}^2}{60}, \quad k_4 = \frac{9}{8}\times 0.1, \quad d = 0.85.$$

$\mathbf{H}_0 \parallel [111]$:

$$G^{(2)}(t) = R_{21}(t)R_{22}(t),$$

$$R_{21} = \left(\frac{1}{2}\cos 3mt + \frac{1}{2}\cos mt\right)^6$$

$$\times\left(\frac{1}{2}\cos 9mt + \frac{1}{2}\cos 3mt\right)^2 \exp\left\{-\frac{1}{2}(1-d')\Delta_{21}^2 t^2\right\},$$

$$R_{22} = \exp\left(-\frac{\lambda^2\Delta_{22}^2 t^2}{2}\right), \quad m^2 = \frac{d'\Delta_{21}^2}{120},$$

$$k_4 = \frac{9}{8}\times 0.14, \quad d' = 0.64.$$

Lithium nuclei

$\mathbf{H}_0 \parallel [100]$:

$$G^{(1)}(t) = R_{12}(t)R_{11}(t),$$

$$R_{12}(t) = \cos^2 2bt\cos^4 bt\exp\left\{-\frac{1}{2}(1-d')\Delta_{12}^2 t^2\right\},$$

$$R_{11} = \exp\left(-\frac{\lambda^2\Delta_{11}^2 t^2}{2}\right), \quad b^2 = \frac{d'\Delta_{12}^2}{12},$$

$$k_4 = \frac{9}{8}\times 0.23, \quad d' = 0.98.$$

$\mathbf{H}_0 \parallel [110]$:

$$G^{(1)}(t) = R_{12}(t)R_{11}(t),$$

$$R_{12}(t) = \cos^2 2bt\cos^4 bt\exp\left\{-\frac{1}{2}(1-d')\Delta_{12}^2 t^2\right\},$$

$$b^2 = \frac{\Delta_{12}^2 d'}{12}, \quad d' = 0.85,$$

$$R_{11} = \left(\frac{1}{2}\cos 3Mt + \frac{1}{2}\cos Mt\right)^2$$

$$\times\left(\frac{1}{2}\cos 6Mt + \frac{1}{2}\cos 2Mt\right)^2 \exp\left\{-\frac{1}{2}(1-d_c)\Delta_{11}^2\lambda^2 t^2\right\},$$

$$M^2 = \frac{\lambda^2 d_c\Delta_{11}^2}{50}, \quad k_4 = \frac{9}{8}\times 0.1, \quad d_c = 0.81.$$

$\mathbf{H}_0 \parallel [111]$:

$$G^{(1)}(t) = R_{12}(t)R_{11}(t),$$

$$R_{12}(t) = \cos^2 3bt\cos^6 bt\exp\left\{-\frac{1}{2}(1-d_c)\Delta_{12}^2 t^2\right\},$$

$$R_{11}(t) = \exp\left(-\frac{\Delta_{11}^2\lambda^2 t^2}{2}\right),$$

$$b^2 = \frac{d'\Delta_{12}^2}{24}, \quad k_4 = \frac{9}{8}\times 0.14, \quad d' = 0.65.$$

APPENDIX B

Stokes and Ailion [20] calculated several lattice sums that are important for understanding the spin–spin relaxation processes taking place under the action of homo- and heteronuclear interactions described by the dipole–dipole Hamiltonian (1):

$$S_1 = \frac{1}{4}\sum_j (1-3\cos^2\theta_{ij})^2\left(\frac{a}{r_{ij}}\right)^6,$$

$$S_1' = \frac{1}{4}\sum_{j'} (1-3\cos^2\theta_{ij'})^2\left(\frac{a}{r_{ij'}}\right)^6,$$

$$S_2 = \frac{1}{16}\sum_j (1-3\cos^2\theta_{ij})^4\left(\frac{a}{r_{ij}}\right)^{12},$$

$$S'_2 = \frac{1}{16}\sum_{j'} (1-3\cos^2\theta_{ij'})^4\left(\frac{a}{r_{ij'}}\right)^{12}, \qquad \text{(B.1)}$$

$$S_3 = \frac{1}{16}\sum_{j,k} (1-3\cos^2\theta_{ij})^2(1-3\cos^2\theta_{ik}) \times (1-3\cos^2\theta_{jk})\left(\frac{a}{r_{ij}}\right)^6\left(\frac{a}{r_{ik}}\right)^3\left(\frac{a}{r_{jk}}\right)^3,$$

$$S'_3 = \frac{1}{16}\sum_{j,k'} (1-3\cos^2\theta_{ij})^2(1-3\cos^2\theta_{ik'}) \times (1-3\cos^2\theta_{jk'})\left(\frac{a}{r_{ij}}\right)^6\left(\frac{a}{r_{ik'}}\right)^3\left(\frac{a}{r_{jk'}}\right)^3,$$

where *a* is the lattice constant. For the LiF crystal under consideration, the summation index in the expressions for S_i (B.1) corresponds to the sites of the fcc lattice occupied by the spins of one type and that in the expressions for S'_i, to the sites of the fcc' lattice occupied by the spins of another type. Here and below, we retain the notation as in [20]. Since the formulas for the cross-relaxation (e.g., (33)) include the terms of the nonsecular interaction, we obtained the lattice sums with a different angular dependence:

$$C = \frac{9}{4S_1}\sum_j \sin^2\theta_{ij}\cos^2\theta_{ij}\left(\frac{a}{r_{ij}}\right)^6,$$

$$C' = \frac{9}{4S_1}\sum_{j'} \sin^2\theta_{ij'}\cos^2\theta_{ij'}\left(\frac{a}{r'_{ij}}\right)^6,$$

$$E = \frac{9}{4S_1}\sum_j \sin^4\theta_{ij}\left(\frac{a}{r_{ij}}\right)^6, \qquad \text{(B.2)}$$

$$S'''_3 = \frac{1}{16C'}\sum_{j',k'} 9\sin^2\theta_{ij'}\cos^2\theta_{ij'}(1-3\cos^2\theta_{ik'}) \times (1-3\cos^2\theta_{j'k'})\left(\frac{a}{r_{ij'}}\right)^6\left(\frac{a}{r_{ik'}}\right)^3\left(\frac{a}{r_{j'k'}}\right)^3,$$

and the sum

$$S''_3 = \frac{1}{16}\sum_{j',k'} (1-3\cos^2\theta_{ij'})^2(1-3\cos^2\theta_{ik'}) \times (1-3\cos^2\theta_{j'k'})\left(\frac{a}{r_{ij'}}\right)^6\left(\frac{a}{r_{ik'}}\right)^3\left(\frac{a}{r_{j'k'}}\right)^3,$$

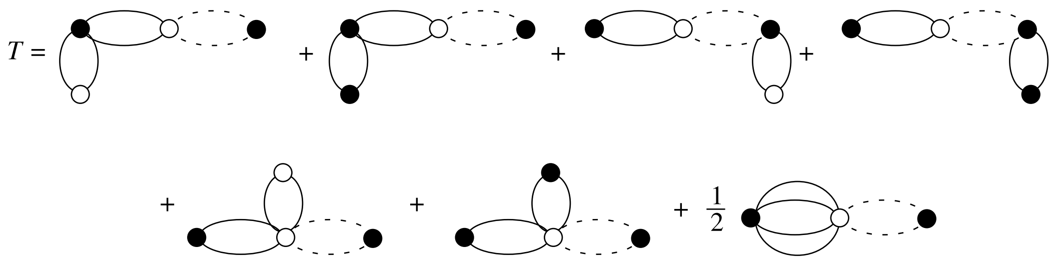
corresponding to the previous angular dependence (not presented in [20]). These lattice sums, calculated for ten coordination spheres (containing 21^3 neighbors), are presented in Table 3.

The three-spin interaction in (33) is represented by the sums of two terms with the coefficients $\alpha_{ik/j}$ and $\beta_{ik/j}$ differing qualitatively by the positions of fluorine nuclei: Li–F–Li and F–Li–Li. Using the above lattice sums, we can estimate the ratio of the aforementioned contributions to the square modulus of the matrix element of the cross-relaxation interaction in (30) for these lithium–fluorine configurations. This ratio amounts to 2.7 for $\mathbf{H}_0 \parallel [100]$, 6.3 for $\mathbf{H}_0 \parallel [110]$, and 1.4 for $\mathbf{H}_0 \parallel [111]$. Therefore, the maximum contribution to the cross-relaxation rate is due to the Li–F–Li configuration. This contribution in (30) is proportional to the lattice sum

$$W = \sum_{i\neq j}\left|B'_{ik}C'_{jk}\right|^2 = \frac{1}{4}\hbar^4\Delta_{12}^4 C' - \sum_i \left|B'_{ik}C'_{ik}\right|^2. \qquad \text{(B.3)}$$

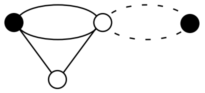
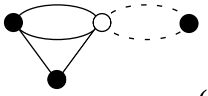
In (B.3), the lattice sum with one summation is much smaller compared to the first term and will be ignored below.

Let us estimate a decrease in the second moment of the cross-relaxation spectrum in the constant local field approximation for the (B.3) contribution. Taking the coefficient at $t^2/2$ in the cosine expansion (37), multiplying this coefficient by $\left|B'_{ik}C'_{jk}\right|^2$, and summing over *i* and *j*, we obtain a large number of lattice sums corresponding to different schemes of bonding. For illustrative purposes, we pass to the graph representation. The sites occupied by lithium and fluorine atoms are denoted by black and open circles, respectively. The bonds $b_{ik}^{(\alpha)}$, $b_{ik}^{(\alpha,\beta)}$, and $B'_{ik} = -4b_{ik}^{(\alpha,\beta)}$ are indicated by solid lines, and C'_{jk}, by dashed lines. In this notation, the contributions of terms without loops are as follows:

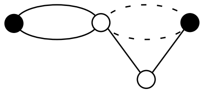
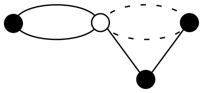


(B.4)

For the terms with a loop of three sites, we obtain

$P_1 = -2$  -2 

(B.5)

$P_2 = -2$  -2 

The terms with a loop of four sites will be ignored. Expressing the terms in (B.4) through the lattice sums presented above, we obtain

$$T = \frac{1}{4}\Delta_{12}^2\hbar^4 C'(\Delta_{22}^2 + 2\Delta_{11}^2 + 2\Delta_{12}^2 + \Delta_{21}^2) - \Delta_{22}^2\frac{4S_2'}{S_1'^2}, \quad \text{(B.6)}$$

where the last term represents a sum of the last diagram (B.4) with compensating corrections arising during the transformation of the lattice sums to the first and penultimate diagrams. The other compensating corrections are small and can be ignored.

By the same token, we obtain

$$P_1 = \frac{1}{2}\Delta_{12}^6 C'\left(\frac{\gamma_2}{\gamma_1} + \frac{5\gamma_1}{\gamma_2}\right)\frac{S_3''}{S_1'^2},$$

$$P_2 = \frac{1}{2}\Delta_{12}^6 C'\left(\frac{\gamma_2}{\gamma_1} + \frac{5\gamma_1}{\gamma_2}\right)\frac{S_3'''}{S_1'^2}. \quad \text{(B.7)}$$

As pointed out above, the correlations vanish in the limit of $\{Z, d\} \longrightarrow \infty$. In this limit, the ratios $S_2'/S_1'^2$, $S_3''/S_1'^2$, $S_3'''/S_1'^2$, and the like go to zero. Therefore, in the absence of correlations, we obtain

$$T_1 = \frac{1}{4}\Delta_{12}^2\hbar^4 C'(\Delta_{22}^2 + 2\Delta_{11}^2 + 2\Delta_{12}^2 + \Delta_{21}^2). \quad \text{(B.8)}$$

The coefficient δ describing a decrease in the second moment due to the correlation effects is equal to the ratio of the sum of contributions (B.6) and (B.7) to (B.8). Making these calculations, we obtain (see Table 3):

$$\delta = 1 - \frac{\left[\frac{4S_2'^2}{S_1'^2} + 2\left(\frac{\gamma_2}{\gamma_1} + \frac{5\gamma_1}{\gamma_2}\right)\left(\frac{S_3''}{S_1'^2} + \frac{S_3'''}{S_1'^2}\right)\right]\Delta_{12}^2}{\Delta_{22}^2 + 2\Delta_{11}^2 + 2\Delta_{12}^2 + \Delta_{21}^2}. \quad \text{(B.9)}$$

In concluding, we also write for the convenience of readers the expressions for the first two moments of the NMR absorption spectrum [38] of LiF in the notation adopted in this paper:

$$M_2^{(2)} = \frac{9}{4}\Delta_{22}^2 + \Delta_{21}^2,$$

$$M_4^{(2)} = M_4^{22} + M_4^{(21,\,21)} + M_4^{(22,\,21)} + M_4^{(21,\,11)},$$

$$\frac{16M_4^{(22)}}{(9\Delta_{22}^2)^2} = \frac{7}{3} - \frac{4S_2}{2S_1^2} + \frac{2S_3}{3S_1^2}, \quad \frac{M_4^{(21,\,21)}}{\Delta_{21}^4} = 3 - \frac{34S_2'}{(5S_1')^2},$$

$$\frac{4M_4^{(22,\,21)}}{9\Delta_{21}^2\Delta_{22}^2} = 6 - \frac{10}{9}\left(1 - \frac{S_3'}{S_1S_1'}\right),$$

$$\frac{4M_4^{(21,\,11)}}{9\Delta_{11}^2\Delta_{21}^2} = \frac{2}{9}\left(1 - \frac{S_3'}{S_1S_1'}\right);$$

and an expression for the fourth moment following from the solution of Eq. (28):

$$\frac{16M_4^{(22)}}{(9\Delta_{22}^2)^2} = 1 + \frac{8\lambda^2}{9}\left(1 + \frac{\nu_2}{4\lambda^2}\right) - \frac{8\lambda^2 S_2}{9S_1^2} + \frac{16}{27}k_4,$$

$$\frac{M_4^{(21,\,21)}}{\Delta_{21}^4} = 3 - \frac{34S_2'}{(5S_1')^2}, \quad \frac{4M_4^{(21,\,11)}}{9\Delta_{11}^2\Delta_{21}^2} = \frac{2}{9}\nu_1,$$

$$\frac{4M_4^{(22,\,21)}}{9\Delta_{21}^2\Delta_{22}^2} = \frac{6 + 4k^{(2)}}{1 + k^{(2)}} = 6 - 2\left(1 - \frac{4\lambda^2}{9}\right).$$

This is the general result taking in account fluctuations of the longitudinal local field component. A transition to the static (binomial) distribution of the longitudinal local fields is performed as ν_a tends to zero.

REFERENCES

1. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of NMR in One and Two Dimensions* (Clarendon Press, Oxford, 1987; Mir, Moscow, 1990).
2. V. E. Zobov and A. A. Lundin, Zh. Éksp. Teor. Fiz. **106**, 1097 (1994) [JETP **79**, 595 (1994)].
3. A. Abragam, *The Principles of Nuclear Magnetism* (Clarendon Press, Oxford, 1961; Inostrannaya Literatura, Moscow, 1963).
4. B. N. Provotorov, T. P. Kulagina, and G. E. Karnaukh, Zh. Éksp. Teor. Fiz. **113**, 967 (1998) [JETP **86**, 527 (1998)].
5. A. A. Lundin and A. V. Makarenko, Zh. Éksp. Teor. Fiz. **87**, 999 (1984) [Sov. Phys. JETP **60**, 570 (1984)].
6. A. A. Lundin, A. V. Makarenko, and V. E. Zobov, J. Phys.: Condens. Matter **2**, 10131 (1990).
7. D. A. McArthur, E. L. Hahn, and R. E. Walstedt, Phys. Rev. **188**, 609 (1969).
8. V. E. Zobov, M. A. Popov, Yu. N. Ivanov, and A. I. Lifshits, Zh. Éksp. Teor. Fiz. **115**, 285 (1999) [JETP **88**, 157 (1999)].
9. V. E. Zobov and A. A. Lundin, Pis'ma Zh. Éksp. Teor. Fiz. **43**, 418 (1986) [JETP Lett. **43**, 536 (1986)].
10. M. I. Bulgakov, A. D. Gul'ko, F. S. Dzheparov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **58**, 614 (1993) [JETP Lett. **58**, 592 (1993)].
11. Yu. G. Abov, A. D. Gul'ko, F. S. Dzheparov, *et al.*, Fiz. Élem. Chastits At. Yadra **26**, 1654 (1995) [Phys. Part. Nucl. **26**, 692 (1995)].

12. V. E. Zobov, Teor. Mat. Fiz. **77**, 426 (1988).
13. V. E. Zobov, Teor. Mat. Fiz. **84**, 111 (1990).
14. V. E. Zobov and M. A. Popov, Zh. Éksp. Teor. Fiz. **103**, 2129 (1993) [JETP **76**, 1062 (1993)].
15. N. G. Kampen, Physica (Amsterdam) **143**, 215 (1974).
16. A. A. Lundin and B. N. Provotorov, Zh. Éksp. Teor. Fiz. **70**, 2201 (1976) [Sov. Phys. JETP **43**, 1149 (1976)].
17. V. E. Zobov, A. A. Lundin, and A. V. Makarenko, Preprint No. 436F, IF SO AN SSSR im. L. V. Kirenskogo (Kirenskii Institute of Physics, Siberian Division, Academy of Sciences of USSR, Krasnoyarsk, 1987).
18. G. E. Karnaukh, A. A. Lundin, B. N. Provotorov, and K. T. Summanen, Zh. Éksp. Teor. Fiz. **91**, 2229 (1986) [Sov. Phys. JETP **64**, 1324 (1986)].
19. P. W. Anderson and P. R. Weiss, Rev. Mod. Phys. **25**, 269 (1953).
20. H. T. Stokes and D. C. Ailion, Phys. Rev. B **15**, 1271 (1977).
21. M. Engelsberg and I. J. Lowe, Phys. Rev. B **10**, 822 (1974).
22. B. T. Gravely and J. T. Memory, Phys. Rev. B **3**, 3426 (1971).
23. A. A. Lundin, Zh. Éksp. Teor. Fiz. **102**, 352 (1992) [Sov. Phys. JETP **75**, 187 (1992)].
24. A. A. Lundin and A. V. Makarenko, in *Nuclear Magnetic Resonance and Crystal Structure* (Inst. Fiz. Sib. Otd. Akad. Nauk SSSR im. L. V. Kirenskogo, Krasnoyarsk, 1984), p. 40.
25. A. A. Lundin and A. V. Makarenko, Fiz. Tverd. Tela (Leningrad) **29**, 1229 (1987) [Sov. Phys. Solid State **29**, 702 (1987)].
26. K. W. Becker, T. Plefka, and G. Sauermann, J. Phys. C **9**, 4041 (1976).
27. G. Sauermann and M. Wiegand, Physica B (Amsterdam) **103**, 309 (1981).
28. P. S. Pershan, Phys. Rev. **117**, 109 (1960).
29. M. Goldman, *Spin Temperature and Nuclear Magnetic Resonance in Solids* (Clarendon Press, Oxford, 1970; Mir, Moscow, 1972).
30. L. L. Buishvili, G. V. Kobakhidze, and M. G. Menabde, Zh. Éksp. Teor. Fiz. **84**, 138 (1983) [Sov. Phys. JETP **57**, 80 (1983)].
31. D. E. Demco, J. Tegenfeldt, and J. S. Waugh, Phys. Rev. B **11**, 4133 (1975).
32. D. V. Lang and P. R. Moran, Phys. Rev. B **1**, 53 (1970).
33. U. Haeberlen, *High Resolution NMR in Solids* (Academic, New York, 1976; Mir, Moscow, 1980); M. Mehring, *High Resolution NMR Spectroscopy in Solids* (Springer-Verlag, New York, 1976; Mir, Moscow, 1980).
34. M. Engelsberg and R. E. Norberg, Phys. Rev. B **5**, 3395 (1972).
35. J. T. Markert and R. M. Gotts, Phys. Rev. B **36**, 6993 (1987).
36. V. A. Atsarkin, Fiz. Tverd. Tela (Leningrad) **27**, 656 (1985) [Sov. Phys. Solid State **27**, 406 (1985)].
37. A. Losche, *Kerninduktion* (Deutsche Verlag der Wissenschaften, Berlin, 1957; Inostrannaya Literatura, Moscow, 1963).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 558–567.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 637–648.
Original Russian Text Copyright © 2001 by Demidov, Karzanov, Markov, Sdobnyakov.

Ion Irradiation Induced Spontaneous Hypersonic Long-Range Stimulation of Silicon Nitride Synthesis in Silicon

E. S. Demidov, V. V. Karzanov, K. A. Markov, and V. V. Sdobnyakov
Nizhni Novgorod State University, Nizhni Novgorod, 603600 Russia
e-mail: ett@phys.unn.runnet.ru

Received January 22, 2001

Abstract—We studied the nature of the effect of medium-energy ion implantation on the defect system of a crystal target over distances exceeding by three to four orders of magnitude the average projected range of ions in the target material. Recently, we discovered an especially strong manifestation of this long-range effect in crystal targets: argon ion bombardment stimulated the formation of a Si_3N_4 phase in nitrogen-saturated layers of a silicon wafer, the effect being observed at a distance of up to 600 μm away from the ion stopping zone. An analysis of changes in the electrical and optical properties of the nitrogen-saturated layer depending on the argon ion dose, in comparison to the morphology development on the ion-irradiated silicon surface, suggests that sufficiently effective pulsed sources of hypersonic (in the initial propagation stage) shock waves appear in the Ar^+ ion stopping zone. These shock waves arise as a result of the jumplike formation and evolution of a network of dislocation loops and argon blisters, accompanied by explosions of the blisters. These processes probably proceed in a self-synchronized or spontaneous manner. Argon in the blisters occurs at $T = 773$ K in a solid state under a pressure of 4.5×10^9 Pa, the blister energy reaching up to 5×10^8 eV. Estimates show that the synchronized explosions of blisters in the region of a nitrogen-saturated layer at the rear side of a 600-μm-thick silicon wafer may produce a peak pressure at the wave front exceeding 10^8 Pa, which is sufficient to cause the experimentally observed changes. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

About three decades ago, investigations involving silicon crystals irradiated by argon ions at an average energy of 40 keV and an ion flux density of 5×10^{13} cm^{-2} s^{-1} revealed the motion of dislocations over distances of several hundred microns, which is three to four orders of magnitude larger than the average projected range R_p of ions in the target material [1, 2]. Subsequently, it was established that the long-range effect is also manifested by a redistribution of interstitial impurity (chromium and iron) atoms possessing diffusion mobility in silicon at room temperature [3, 4]. Then, it was found that the long-range effect can also lead to a redistribution of impurities more tightly bound to a target crystal lattice, in particular, impurity atoms of the iron group in A^3B^5 compounds (GaAs, InP, GaP) [5, 6]. It was established that the irradiation with 40-keV argon ions can stimulate the process of oxygen redistribution in silicon [7] and even the formation of a stoichiometric Si_3N_4 phase in Si_xN_y layers in silicon crystals saturated with nitrogen during preliminary irradiation with nitrogen ions [8, 9].

In this paper, we will consider the nature of the long-range effect of medium-energy ion irradiation taking into account the laws of the Si_3N_4 phase formation and the results of a recent atomic force microscopy (AFM) investigation [10] of a silicon crystal surface irradiated with argon ions. We believe that most probable is an acoustic mechanism of the long-range effect. A significant increase in the electric resistance and the characteristic (for Si_3N_4 phase) optical absorption in the nitrogen-saturated silicon layer is observed (together with the aforementioned manifestations of the long-range effect) in silicon crystals irradiated with inert gas ions to a total dose of $\Phi \sim 10^{16}$–10^{17} cm^{-2}. For $\Phi \geq 10^{16}$ cm^{-2}, a network of dislocation semiloops arises under the argon-implanted layer, while the coalescence of incorporated argon in this layer leads to the formation of large gas bubbles called blisters. It is suggested that, as the ion dose Φ increases, the jumplike processes of the formation and evolution of the dislocation network and gas blisters and the explosion of blisters give rise to the acoustic pressure pulses. In all stages of this process, the pulsed pressure exhibits amplification due to self-synchronized or spontaneous character of the jumplike events.

A comparative analysis of the visible dimensions of blisters (or craters formed upon blister explosions) in the AFM images, the balance of implanted argon, and the melting and boiling temperatures of inert gases as functions of the pressure indicates that a static pressure of argon in a blister immediately before explosion may reach up to 4.5×10^9 Pa. This peak pressure level is close to that typical of the strongest packets of hypersonic shock waves (in the initial propagation stage) with a duration of 3×10^{-9} s detected after blister explosions. Estimates showed that a peak pressure after a

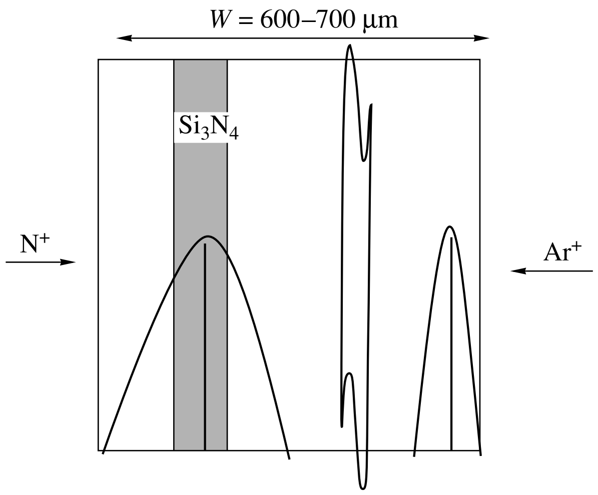


Fig. 1. Schematic diagrams illustrating the double sequential irradiation of silicon samples with nitrogen and argon. Nitrogen saturation conditions: $E = 150$ keV; $R_p^N = 370$ nm; $\Delta R_p = 90$ nm; $\Phi = (1–5) \times 10^{17}$ cm^{-2}; T_{imp}, room temperature. Argon irradiation conditions: $E = 40$ keV; $R_p^N = 41$ nm; $\Delta R_p = 18.1$ nm; $\Phi = (10^{15}–3 \times 10^{17})$ cm^{-2}; $T_{imp} = 500$°C.

spontaneous blister explosion in the nitrogen-saturated layer at the rear side of a 600-μm-thick silicon wafer reaches 10^7–10^8 Pa. Thus, the known [11] threshold pressure (10^7 Pa) of nitrogen necessary for initiating the reaction of Si_3N_4 phase synthesis in a silicon powder under static conditions at 300°C is reached.

It must be noted that successful experiments on the synthesis of a dielectric Si_3N_4 phase, as well as elucidation of the nature of the long-range effect of ion irradiation, are important both for applications (development of the silicon-on-insulator technology for microelectronics) and for the fundamental knowledge of the ion sputtering processes, radiation damage and fracture of materials, and extremal properties of materials (high-energy materials science).

2. EXPERIMENT DESCRIPTION AND PRELIMINARY REMARKS

The experiments involving double sequential irradiation with nitrogen and argon (Fig. 1) were performed on 600- to 700-μm-thick single crystal silicon wafers, representing the materials both with high resistance (100 Ω cm *n*-Si or 2000 Ω cm *p*-Si) used for the optical absorption measurements and with very low resistance (0.005 Ω cm *n*-Si) for the electrical measurements. In the first stage, a silicon wafer was irradiated from one side with 40- or 150-keV nitrogen ions at a dose of up to 5×10^{17} cm^{-2} at a temperature of 400°C for creating a subsurface or buried nitrogen-saturated layer. Then, the wafer was irradiated from the opposite side with 40-keV argon ions at a low flux density of 5×10^{13} cm^{-2} s^{-1} to a dose varied from 10^{15} to 3×10^{17} cm^{-2}. This exposure was performed at a target temperature of 500°C, which was found in preliminary experiments to be sufficient for the Si_3N_4 phase formation under the action of argon ion bombardment.

The IR absorption spectra were measured on a UR-20 spectrophotometer and Bruker IFS-113V Fourier spectrometer in the wavenumber range from 650 to 1300 cm^{-1} containing all characteristic absorption bands of the Si_3N_4 phase (the main peak observed in the region of 850 cm^{-1}). The IR absorption and electrical (transverse resistance of the nitrogen-saturated silicon layer) measurements were described in detail elsewhere [8, 9].

Figure 2 shows the effect of increasing argon ion dose on the intensity of the IR absorption band characteristic of Si_3N_4 and the average electric resistivity of the nitrogen-saturated silicon layer. As is seen, the two curves correlate well with each other: significant changes start at a dose Φ between 10^{16} and 3×10^{16} cm^{-2} and cease between 3×10^{16} and 10^{17} cm^{-2}. It should be noted that our experiments showed that similar changes in the properties of the nitrogen-saturated silicon layer subjected to thermal annealing in the absence of argon ion irradiation were observed after treatment at a tem-

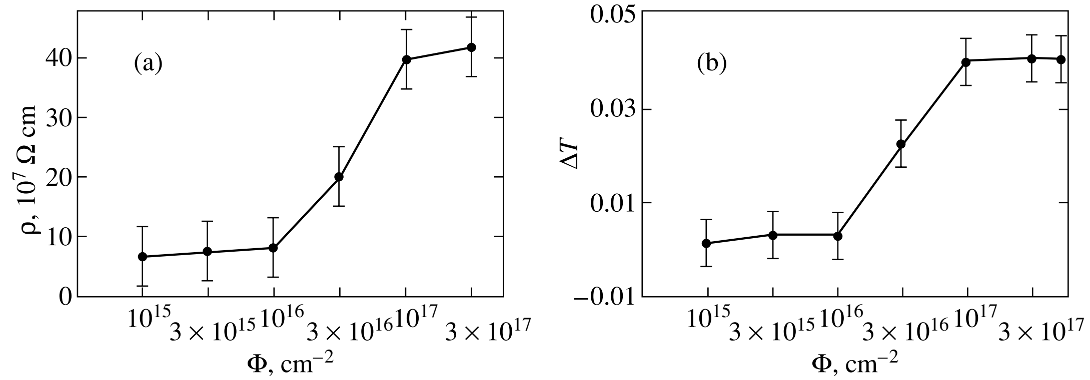


Fig. 2. Plots of the (a) average resistivity ρ and (b) intensity Δ*T* of the IR absorption band characteristic of Si_3N_4 versus increasing argon ion dose Φ in the nitrogen-saturated silicon layer.

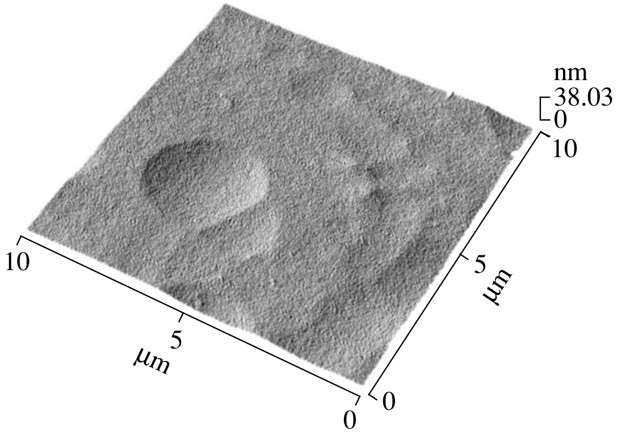


Fig. 3. An AFM image of the surface of a nitrogen-saturated silicon sample irradiated with 40-keV argon ions at T = 500°C to $\Phi = 10^{15}$ cm^{-2}.

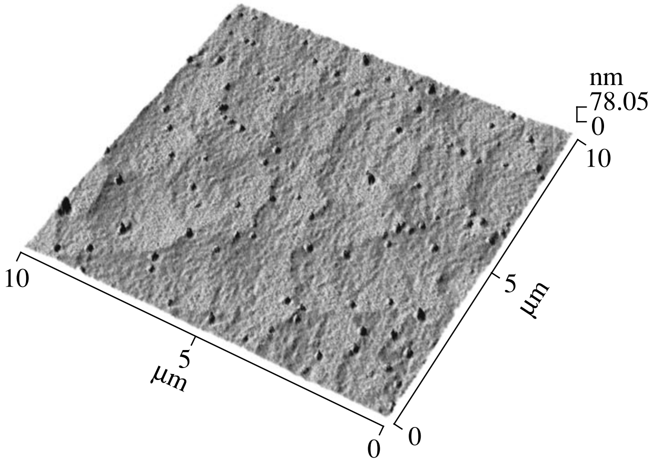


Fig. 4. An AFM image of the surface of a nitrogen-saturated silicon sample irradiated with 40-keV argon ions at T = 500°C to $\Phi = 10^{16}$ cm^{-2}.

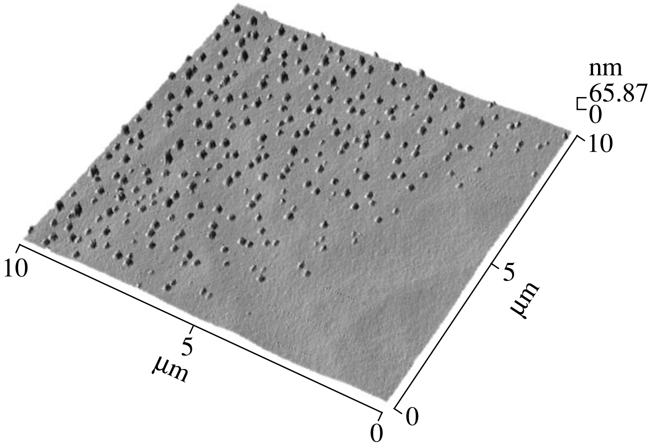


Fig. 5. An AFM image of the surface of a nitrogen-saturated silicon sample irradiated with 40-keV argon ions at T = 500°C to $\Phi = 3 \times 10^{16}$ cm^{-2} (the image is taken from an area adjacent to the region shadowed by the sample holder during ion irradiation).

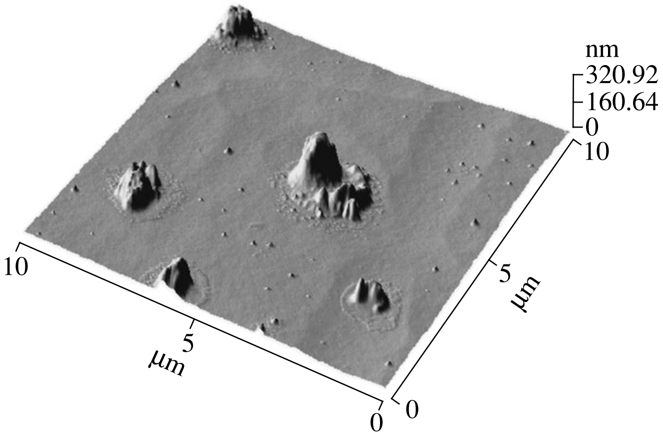


Fig. 6. An AFM image of the surface of a nitrogen-saturated silicon sample irradiated with 40-keV argon ions at T = 500°C to $\Phi = 3 \times 10^{16}$ cm^{-2} and then stored for three months in air at room temperature.

perature of 1000°C. Thus, ion irradiation of the rear side of a 600- to 700-μm-thick silicon sample allows the temperature required for a dielectric Si_3N_4 phase formation to be reduced down to 500°C.

The surface of silicon samples irradiated with argon ions under the same conditions (E = 40 keV, T = 500°C) as the samples presented in Fig. 2 was studied by AFM on an Accurex Topometrix scanning-probe microscope. The measurements were performed using a silicon nitride probe in a contact mode in air at room temperature.

It was found that ion irradiation to a dose as small as $\Phi = 10^{15}$ cm^{-2} leads to the disappearance of small (~2 nm deep) scratches present on the surface of unirradiated samples. The AFM image in Fig. 3 demonstrates the well-known phenomenon of surface swelling [12] related to the amorphization of silicon and the formation of clusters or small bubbles of dissolved argon. The latter are manifested by fine surface roughnesses with dimensions on the order of 10 nm.

When the ion dose is increased to $\Phi = 10^{16}$ cm^{-2} (Fig. 4), large argon bubbles (blisters) appear, which are similar to those observed in silicon irradiated with neon ions [13]. As the irradiation dose grows further to $\Phi = 3 \times 10^{16}$ cm^{-2} (Fig. 5), the size and surface density of blisters increase, while the swelling vanishes. The surface between blisters exhibits smoothening (also observed in [13]) up to a mirror finish.

Figure 6 demonstrates a remarkable pattern observed on the surface of a silicon single crystal also (as that presented in Fig. 4) irradiated to $\Phi = 3 \times 10^{16}$ cm^{-2} and then stored for three months in air at room temperature. This image indicates that the process of coalescence

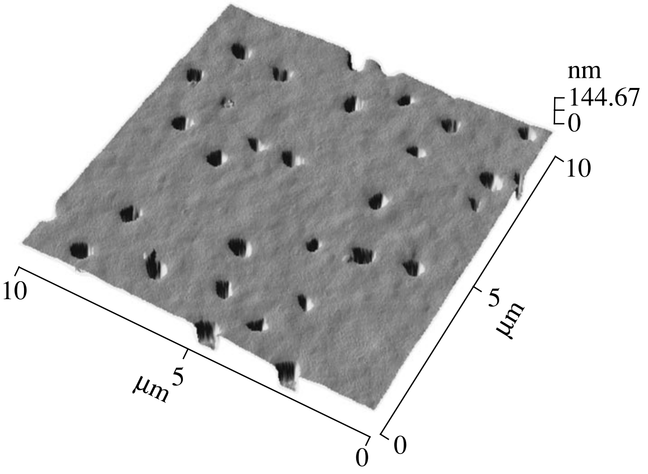


Fig. 7. An AFM image of the surface of a nitrogen-saturated silicon sample irradiated with 40-keV argon ions at T = 500°C to $\Phi = 10^{17}$ cm^{-2}.

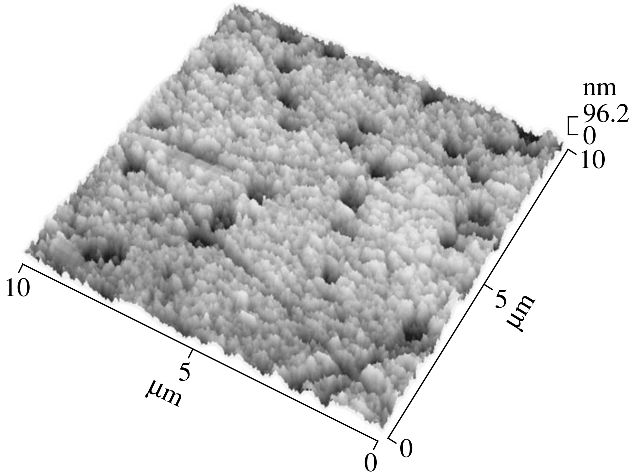


Fig. 8. An AFM image of the surface of a nitrogen-saturated silicon sample irradiated with 40-keV argon ions at T = 500°C to $\Phi = 3 \times 10^{17}$ cm^{-2}.

proceeds even at room temperature: the surface density of blisters drops 100 times, while the average dimensions increase to reach 250 nm in height and 1000 nm in diameter; at the same time, the total volume of blisters (i.e., the volume of encapsulated argon) remains the same. These phenomena show evidence of, first, a high mobility related to the process of blister evolution and a high mobility of argon and silicon atoms and, second, a superplasticity of the layer of silicon amorphized as a result of the argon ion irradiation. Indeed, for a blister cap thickness of 40 nm and a diameter of about a fraction of micron, the deformation is on the order of 10%, which is far beyond the possible level of purely elastic deformation in solids. Apparently, this superplasticity may also account for the aforementioned smoothening of the sample surface between blisters. Another remarkable phenomenon is the healing of the amorphized layer surface at the sites where blisters have disappeared. The angular shape of blisters observed in Fig. 4 apparently indicates that argon in the blisters occurs in a crystalline solid state, in agreement with the experimental evidence at $T \approx 300$ K [14] and the opinion of Ezhevskiĭ *et al.* [15].

A not less striking pattern is observed in Fig. 7 showing an AFM image of the silicon surface irradiated with argon ions to a dose of $\Phi = 10^{17}$ cm^{-2}. This image displays holes (craters) remaining after opening of the blisters. As can be seen, virtually all blisters in this sample are open. The surface density of blisters in this stage (prior to explosion) is smaller, while the dimensions of blisters are greater, as compared to those in Fig 5. The craters exhibit sharp angles and edges, nearly vertical walls, and almost flat bottom; the sample surface between craters is smooth. Thus, the craters appear as equally "young," which indicates that the explosion took place at $\Phi \sim 10^{17}$ cm^{-2} and, most probably, virtually simultaneously. The craters possess equal depths of about 100 nm, which is approximately 2.5 times the average projected range of 40-keV argon ions (R_p = 41 nm). As is well known, irradiated silicon wafers contain, under the argon implanted layer, a layer enriched with vacancies and a still deeper layer enriched with intrinsic interstitial silicon atoms [16]. The crater depth of 100 nm suggests that implanted argon in the samples irradiated to large doses migrates toward the vacancy-rich layer and accumulates there, at a depth of about $2.5R_p$. The aforementioned high mobility of blisters is indicative of a significant permeability of this, apparently porous, silicon layer. According [13], a porous layer is also formed in the samples irradiated with neon ions.

Finally, Fig. 8, showing data for a sample irradiated to $\Phi = 3 \times 10^{17}$ cm^{-2}, indicates that an increase in the ion dose above 10^{17} cm^{-2} does not lead to the appearance of new large blisters. The AFM image shows evidence of erosion in both smooth areas and craters, which is based on the formation and explosion of small argon blisters. This behavior is quite reasonable: the existing craters allow the incorporated argon to escape readily from the porous layer.

Table 1 summarizes the results of processing of the experimental AFM data (Figs. 3–8), presenting the parameters of blisters and craters. Also indicated is the thickness W_r of a silicon layer sputtered during the irradiation, which was calculated according to [17]. Data in the fifth column (lines 4 and 5) confirm almost exact conservation of the total volume of blisters in the course of their evolution (for crystals irradiated to the same dose $\Phi = 3 \times 10^{16}$ cm^{-2}). The sixth line indicates the proposed parameters of blisters immediately before explosion (for $\Phi = 6 \times 10^{16}$ cm^{-2}), determined from a comparison of the parameters of craters (two bottom lines) assuming that virtually all implanted argon (for

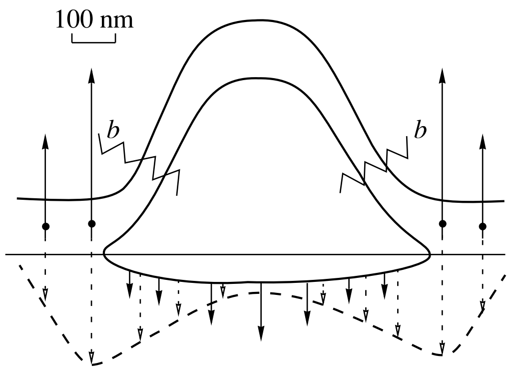


Fig. 9. Schematic diagram illustrating a proposed shape of the argon blister prior to explosion. Black and open arrows show the distribution of mechanical stresses immediately before and after explosion. Broken lines *b* indicate expected sites of the blister cap fracture.

$\Phi = 10^{16}$ cm^{-2}) is spent for the blister formation and that the argon density in the blisters is constant. The latter assumption is valid if argon in the blisters occurs in the least compressible solid state both at room temperature (as in [14]) and during the irradiation at 500°C (our case). It should be noted that this temperature (500°C = 773 K) is much higher than the critical temperature of Ar (T_c = 150.65 K) [18]. This implies that only two states of argon may exist at 500°C (liquid and gaseous states are indistinguishable).

Figure 9 shows a schematic diagram of the proposed shape and size of blisters prior to explosion. The blister cap thickness depicted in this figure is about 40 nm, as estimated with an allowance for the crater depth (100 nm at $\Phi = 10^{17}$ cm^{-2}, which corresponds to the lower boundary of a porous silicon layer), the porous silicon layer thickness ($\Delta R_p \approx 20$ nm), and shift of the surface as a result of the atomic sputtering (~20 nm), and a decrease in the cap thickness by a factor of about 1.5 as a result of the deformation to a nearly hemispherical shape with a curvature radius of about 300 nm. The latter curvature radius at a blister cap thickness of 40 nm corresponds to a 10% material deformation level (indicated above to be far beyond the known elasticity limits of solid materials).

Ezhevskiĭ *et al.* [15] estimated the inert gas pressure in a blister in silicon at $T \approx 300$ K at approximately 2.5×10^9 Pa by considering the blister cap bent as an elastic membrane. Under our experimental conditions ($T \sim 773$ K), taking into account the aforementioned superplasticity of amorphized silicon, this estimate appears to be inadequate. We determined the upper possible limit for the static pressure in blisters from a numerical analysis of the relationship between melting point and pressure for inert gases. It was established that all the known inert elements (except for helium) obey a simple relationship (Fig. 10):

$$T_m = T_{m0} + Ap^k \qquad (1)$$

with the same exponent $k = 0.9$ and the parameters T_{m0} and A indicated in Table 2. The higher the pressure, the better the fit of experimental points (reference data taken from [18]) to the relationship (1). These results allow us to extrapolate Eq. (1) up to 773 K and estimate the upper limit of the static pressure in a blister as $p_{0b} = 4.5 \times 10^9$ Pa; this estimate is not much higher than that obtained in [15] for 300 K. For this pressure, with an allowance for the compressibility of solid argon [18], we can estimate the density $\rho = 2.5$ g/cm^3 that is greater approximately by half than the value ($\rho_{\text{Ar}} = 1.65$ g/cm^3) at normal atmospheric pressure.

The estimated static pressure 4.5×10^9 Pa of argon in a blister with the configuration depicted in Fig. 9 cor-

Table 1. Average parameters of blisters and craters in Si crystals irradiated with Ar ions

Φ, cm^{-2}	D_b, D_c, nm	h_b, d_c, nm	N_b, N_c, cm^{-2}	N_bV_b, N_cV_c, nm cm^2	W_r, nm	ρ/ρ_{Ar}
			Blisters			
10^{16}	150	60	1.1×10^8	0.99	3.5	3
3×10^{16}	200	40	4.5×10^8	5.9	10.5	1.4
*3×10^{16}	1000	250	5×10^6	6.5	10.5	1.2
**6×10^{16}	600	250	2.5×10^7	12	21	1.4
			Craters			
10^{17}	450	100	2.5×10^7	5	35	
3×10^{17}	600	60	2×10^7	4.3	105	

Note: D_b and D_c are the diameters, N_b and N_c the surface densities, h_b and d_c the height and depth, V_b and V_c the volumes, and N_bV_b and N_cV_c the total volumes of blisters and craters, respectively (V_b was calculated as the volume of hemisphere $\pi h_b D_b^2/6$); W_r is the thickness of the layer removed (sputtered) by ion irradiation; $\rho_{\text{Ar}} = 2.5$ g/cm^3 is the solid argon density under the static pressure of $p_s = 4.5 \times 10^9$ Pa; ρ is the argon density calculated assuming that all implanted argon is spent for the blister formation; asterisk refers to the parameters determined after a three-month room-temperature storage; double asterisk indicates predicted values.

responds to a tensile stress of about 2.5 × 10^{10} Pa in the blister cap. This value is ten times the ultimate strength of silicon single crystals and approaches the value for thin carbon or sapphire dislocation-free crystals (whiskers) [18]. This increased value of the argon-bombardment-amorphized silicon seems to be quite realistic, since the dislocation mechanism of fracture is inoperative in the amorphous material and the surface microcracks or, in our case, nanocracks (limiting the strength of glasses) are healed under the action of ion irradiation.

The unusual combination of superhigh plasticity and strength is probably related to the presence of a considerable amount of silicon monomers (similar to carbon monomers), possessing a high tensile strength and pliability with respect to bending deformation, in the argon-bombardment-amorphized silicon. The existence of a quasi-one-dimensional allotropic form of silicon was recently experimentally established [19, 20]. This allotropic form can appear in amorphous silicon at 500°C [19] and was observed in silicon samples irradiated with neon ions [20].

The fact that argon escapes from blisters as a result of explosion, rather than leaks gradually, is confirmed by the well-known experimental phenomenon of peeling observed in irradiated materials exhibiting blistering [12]. Proceeding from the proposed blister shape (Fig. 9), this mechanism can be explained by three reasons. First, despite a high plasticity of the amorphized silicon, some shear stresses are certainly developed in this material, reaching a maximum at the base of the blister cap. Second, the rate of silicon sputtering is higher for an oblique incidence of the ion beam, which also indicates that maximum sputtering takes place at the cap base. Third, for a blister cap thickness close to the average projected ion range R_p, the implanted argon would “cut” the cap at the base (the site is indicated in Fig. 9), where the cap thickness projection onto the direction of normal to the crystal surface exceeds R_p.

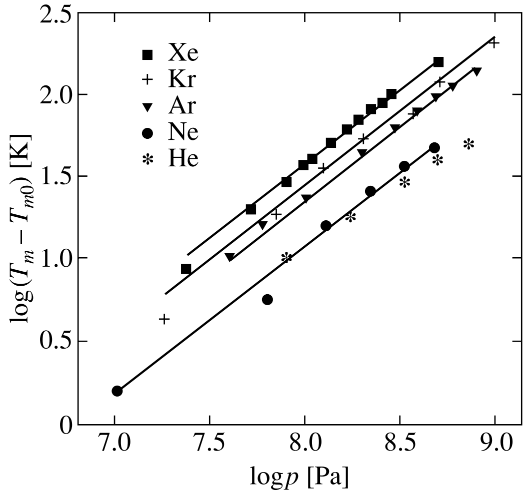


Fig. 10. Plots of the melting temperature versus pressure for various inert gases. Points show the experimental data from reference book [18]. Solid lines are calculated by the equation $T_m = T_{m0} + Ap^k$ with $k = 0.9$.

3. DISCUSSION OF THE MECHANISM OF THE LONG-RANGE EFFECT OF ION IRRADIATION

Incorporation of a single medium-energy ion into a crystal leads to the development of extremely high peak temperatures (~10^8 K) and pressures (10^{12} Pa). The process of ion stopping gives rise to numerous factors acting upon the crystal, including the production of electromagnetic radiation with a frequency spectrum up to the X-ray range, electron–hole pairs, intrinsic defects (vacancies and interstitial atoms), and acoustic pulses. However, almost all these factors decay within a 15- to 20-μm-thick layer. Both highly and poorly conducting crystals exhibit a universal maximum transparence for the acoustic waves. However, a principal difficulty in developing an acoustic model for the long-range effect of ion irradiation is related to a very short time of action (the ion stopping time is 10^{-13}–10^{-14} s [21]) and very small size of a region of the acoustic wave emission (on the order of a few nanometers). Estimates made in [22] for the argon ions with a primary energy of E = 40 keV and an ion flux density of 10^{14} cm^{-2} s^{-1}, assuming the displacement cascade of silicon atoms to be characterized by a relaxation time of 10^{-11} s and an excited region radius of 5 nm, showed that the pressure of acoustic pulses over distances of several hundred microns can be only on the order of 10^5 Pa. This pressure is insufficient even to overcome a threshold of the dislocation mobility equal to 10^6–10^7 Pa [21].

The appearance of a silicon crystal surface with traces of exploded blisters imaged in Fig. 7 gives some grounds for explaining the long-range effect. First, the ion irradiation of silicon leads to the formation of rather large blisters with a high argon pressure, possessing a considerable accumulated energy. Calculations show that the potential energy of the compressed argon contained in a blister prior to explosion amounts to 5 × 10^8 eV. The duration of explosion, estimated at a few tenths of nanosecond, corresponds to the lengths of acoustic waves close to the blister size. Thus, all conditions are

Table 2. Parameters of the melting temperature–pressure relationship for inert gases

	T_{m0}, K	A, $\text{Pa}^{-0.9}$ K
He	0	
Ne	24.4	7.5 × 10^{-7}
Ar	83.8	1.4 × 10^{-6}
Kr	116	1.78 × 10^{-6}
Xe	161.3	2.34 × 10^{-6}

provided for a highly effective acoustic emission. Moreover, such acoustic waves will readily penetrate, experiencing almost no absorption, through a 600-μm-thick crystal depicted in Fig. 1. The second ground is the possibility that a self-synchronized or spontaneous process of the blister explosion wave propagation over the crystal surface may take place, which would result in the formation of a wave front featuring a peak pressure multiply enhanced as compared to a single partial source.

The blisters begin to explode, as indicated above, in the interval of ion doses from 3×10^{16} to 10^{17} cm^{-2}, while changes in the properties of the nitrogen-saturated layer already take place (see Fig. 2) in the region of doses between 10^{16} and 3×10^{16} cm^{-2}. We believe that the formation and evolution of blisters prior to explosion, as well as the formation and evolution of a dislocation network under the ion-implanted layer, also contribute to the pulsed acoustic emission. The crystal region near the ion-irradiated surface represents a system of alternating contracted and expanded layers. The uppermost layer is contracted due to the incorporation of excess argon; the next layer is rich in vacancies and, hence, expanded; this is followed by another compressed layer rich in interstitial silicon atoms. This distribution of stresses must lead to the formation of localized near-surface dislocation loops or semiloops emerging at the surface. The intrinsic dislocation density of a crystal is typically 10^4 cm^{-2}, which is much smaller as compared to the surface density of blisters observed in Figs. 4–6. Probably, the centers of argon coalescence are represented just by dislocation semiloops appearing as a result of the stress release near the irradiated layer in the course of the argon ion dose growth. Thus, the evolution of blisters and the formation of dislocations are interrelated processes that are likely to possess a jumplike character. As the ion dose increases, a close to critical state is developed before each jumplike change. During the blister explosion, as well as in all other stages of the system development, a self-synchronized propagation of the wave of acoustic source switching is highly probable.

The self-synchronized process is a variant of the mechanism of a nondecaying acoustic wave propagation in a crystal proposed by Pavlov *et al.* [22]. In the original variant, the wave energy losses were compensated by the mechanical discharge of nonequilibrium defects present in a real crystal occurring in a nearly critical state. In our case, the potential sources occurring in a subcritical state are distributed over a crystal surface and are controlled and renewed by ion irradiation of the crystal.

The problem of describing a fanlike wave of the acoustic source switching (“domino effect”), propagating along the surface in which these sources reside, and a superposition of the generated pulsed acoustic wave packets is very interesting but complicated and requires special analysis. Here, our aim is to obtain some estimates within the framework of a simplified model. In particular, we neglect the crystal anisotropy and consider the simplest variant, when all sources are switched on simultaneously. This is possible provided that dispersion in attaining the critical state by numerous sources would be sufficiently small, so that a subset of the primary sources would almost simultaneously operate under the action of thermal fluctuations or impinging ions. Under these conditions, an almost plane wave would form at a large distance inside the crystal with a peak pressure

$$p = \gamma p_0, \tag{2}$$

where p_0 is the peak pressure on the surface of a hemispherical source with the radius r_S and γ is the geometric coefficient of the pressure attenuation. The latter coefficient can be expressed as

$$\gamma = \pi r_S^2 N_S, \tag{3}$$

where N_S is the surface density of the acoustic emission sources, which reflects a uniform distribution of the pressure from individual sources over the wave front. In the case of harmonic oscillations, the interference gives rise to a nearly plane wave with the front parallel to the surface featuring a uniform distribution of coherent point sources. In a system of pulsed sources with the upper frequency much lower compared to the Debye frequency, the dispersion is negligibly small and all phase relationships between the spectral components are retained in the pulsed wave front. In other words, the wave front repeats the time variation of the pressure at the almost point sources.

Consider the case when the process starts with one source S_1, after which the sources switch one another by the generated acoustic pulses. Here, a wave with a cone-shaped front would form, at least at a large distance from the first source S_1, which is analogous to the wave from simultaneously switching imaginary sources situated on the surface of a cone with the vertex at the source S_1. The cone angle is determined by the ratio of the surface and volume waves. A special feature of the system is that the geometric divergence of the spherical wave from each individual source and the physical attenuation due to the acoustic energy absorption are determined by the actual position of the source on the crystal surface. At a large distance from S_1, the curvature of the wave front is small; that is, the front is almost flat and is inclined at 45° to the crystal surface (provided that the velocity of switching sources is equal to the velocity of sound in the crystal volume). In this case, the peak pressure at the wave front is determined by expressions (2) and (3) because the projected density of sources per unit wave front area is the same.

Judging from the complicated pattern of mechanical stresses schematically depicted in Fig. 9, each blister explosion gives rise to a number of complex oscillation modes. In particular, a wave with a concave front featuring a caustic effect of the hypersound focusing is

Table 3. Parameters of sources of acoustic pressure pulses

Source type	N_s, 10^7 cm^{-2}	r_s, nm	p_0, Pa	τ, s	$C_{\rm pr}$	$p_{\rm fs}$, Pa	γ	p_f, Pa
Formation and evolution of dislocation network	2.5–45	250–1000	10^6–10^7	$(0.5–2) \times 10^{-10}$	0.54–0.97	1.5×10^4	1	4×10^5–10^7
Formation and evolution of clusters	2.0–45	75–300	2×10^9	$(0.5–2) \times 10^{-10}$	0.54–0.97	$(1–10) \times 10^5$	0.05–0.08	4×10^7–1.4×10^8
Explosion of blisters	2.0–2.5	300	2×10^9	3×10^{-10}	0.98	10^6	0.05–0.07	10^8–1.4×10^8

Note: N_s is the surface density of sources; r_s is the source radius; p_0 is the initial peak pressure; τ is the pressure pulse duration; $C_{\rm pr} = 10^{-\alpha W/20}$ is the attenuation coefficient (W is the crystal thickness); $p_{\rm fs} = (r_s/W)C_{\rm pr}p_0$ is the final peak pressure due to a single source; γ is the geometric decay coefficient; $p_f = \gamma C_{\rm pr}p_0$ is the final peak pressure for self-synchronized sources. For dislocations: $r_s = 0.5\sqrt{N_s}$ and $\tau = 2r_s/v_s$ (sound velocity $v_s = 5 \times 10^5$ cm/s).

possible [23, 24]. However, at large distances from each source, an almost spherical compression wave is likely to be most important. This wave arises as a result of the mechanical recoil reaction accompanying acceleration of the gas (argon escaped from the exploding blister) and the blister cap. In expression (1), we took $p_0 = 2 \times 10^9$ Pa (rounded half of the static pressure estimate $p_{0b} = 4.5 \times 10^9$ Pa obtained above) by analogy with the case of exploding sphere [24] where half of the energy is spent for the internal wave. The same p_0 value was used to characterize the formation and jumplike growth of the argon blisters, where the gas volume also exhibits a sharp increase. Selection of the initial pressure $p_0 \approx 2 \times 10^9$ Pa implies that evolution and explosion of the blisters is accompanied by the formation of acoustic wave packets representing, at least initially, “soft” shock waves [23].

In order to provide for the motion of dislocations, the pressure must be on the order of 10^6–10^7 Pa. This implies that moving dislocations generate pressure pulses of the same magnitude, that is, $p_0 = 10^6$–10^7 Pa. The radius of the region of dislocation stresses amounts to 10^3–10^4 interatomic distances [21]. Judging by the density of decorating blisters or craters (Figs. 4–7), the dislocations are closely packed. For this reason, the value of γ in (1) was taken equal to unity.

The physical attenuation caused by the acoustic energy dissipation was evaluated at the upper frequency of the spectrum,

$$f = \frac{1}{2\pi\tau}, \tag{4}$$

in which most of the energy is transferred by an acoustic pulse with the duration τ. For the explosion of blisters, an estimate of 3×10^{-10} s was calculated as the time during which the center of mass of argon shifts by half of the height of the blister under the action of an average accelerating pressure of 2×10^9 Pa. This displacement corresponds to doubling of the volume of argon escaped from the crater. For dislocations, which were assumed to move at a velocity of sound v_s, the lower boundary of τ was estimated as

$$\tau = \frac{1}{v_s N_d^{1/2}}, \tag{5}$$

where N_d is the density of dislocations; it was assumed that this value is equal to $N_d = N_b$, that is, to the density of decorating blisters. The average velocities of the transverse and longitudinal acoustic waves in silicon are 5×10^3 and 9×10^3 m/s, respectively [18]. Since the motion of dislocation loops is likely to be limited by slower transverse waves, we selected $v_s = 5 \times 10^3$ m/s. The same value was taken for blisters, by virtue of the aforementioned relationship in evolution of the dislocations and blisters.

Estimated as described above, the τ values (presented in Table 3) correspond to a hypersonic frequency range. Extrapolating of the temperature dependence of the hypersound absorption coefficient α [25, Fig. 4.10] for the longitudinal waves with a frequency of $f \approx 0.5$ GHz (according to Mason and Bateman [25, (4.77)]) to the temperature of our samples (773 K) irradiated with argon ions, we obtained $\alpha = 2.5$ dB/cm. For smaller τ values (Table 2), we used the (most “unfavorable”) Akhiezer law $\alpha \sim f^2$, although a comparison of the data $f = 0.48$ GHz [25, Fig. 4.10] and $f = 9$ GHz [25, Table 4.5] for 62 K rather suggests the Landau–Rumer law $\alpha \sim f$. When the temperature increases, the situations (judging by [18, Fig. 7.47]) becomes intermediate.

With an allowance for the hypersound attenuation (see Eqs. (2) and (3)), the resulting final peak pressure p_f at the rear side of a silicon plate with the thickness W is determined by the expression

$$p_f = C_{pr}p = C_{pr}\gamma p_0, \tag{6}$$

where C_{pr} is the physical pressure attenuation coefficient (expressed in dB). According to standard definition, this coefficient obeys the relationship $-20\log C_{pr} = \alpha W$. Table 3 gives the values of p_f for $W = 600$ μm. For comparison, Table 3 also gives the final pressure p_{fs} of a single source calculated by formula (6) with the geometric factor $\gamma = r_s/W$ according to the law $p \sim 1/r$, where r is the distance to the source. As demonstrated in [22], this law is valid at a distance $r \gg r_s$ for a source generating sufficiently short pressure pulses such that $v_s\tau \ll r$.

As can be seen from the data in Table 3 for a single source of hypersound, the p_{fs} value may just reach a threshold level for the dislocation motion in the region of the nitrogen-saturated crystal layer in the variant with a change in the state of blisters. With a self-synchronization mechanism, this threshold is reached for all three variants presented in Table 3. The increase in the average p_f value on the passage from top to bottom line in the last column of Table 3 is consistent with variation of the physical properties of the nitrogen-saturated layer (Fig. 2) and with the proposed sequence of processes (based on the AFM image analysis, see Figs. 3–7) from the formation and evolution of dislocations and blisters to the explosion of blisters leading to the generation of hypersonic pulses. The maximum peak pressure (above 10^8 Pa) in Table 3 exceeds by at least one order of magnitude the known [11] static threshold nitrogen pressure (10^7 Pa) for ignition of the Si_3N_4 phase synthesis using silicon powder in N_2 atmosphere at 300°C. Validity of the comparison of this process to a reaction in the nitrogen-saturated silicon layer may be related to the probable existence of nitrogen bubbles in this layer. The gas pressure in such bubbles can be high, much in excess of 10^7 Pa. A fraction of nitrogen enters into a chemical reaction with silicon already in the stage of nitrogen implantation. Another fraction reacts with silicon in the course of the sample heating to 500°C. The unreacted nitrogen remaining in the bubbles at a subthreshold pressure is consumed under the action of hypersonic pressure pulses generated in the argon stopping zone by sources with the parameters indicated in Table 3.

Another possible contribution to the mechanism of stimulation of the Si_3N_4 formation is related to the presence of stresses and a dislocation network in the region of the nitrogen-saturated layer. This region must also contain alternating contracted and expanded layers, similar to those formed near the argon ion stopping zone. Under the action of hypersonic pulses, the edge and/or mixed dislocations migrate (see [21]) so as to generate vacancies and interstitial atoms. This favors increased mobility of silicon and nitrogen atoms, accelerates the formation of Si_3N_4 phase, and (which is of special value from technological standpoint) leads to leveling of the properties of the nitrogen-saturated layer and to decreasing mechanical stresses in this layer and in the neighboring silicon crystal regions. It is important to note that the pressure effects are developed in a dynamic regime. The geometric length $v_s\tau$ of the acoustic pulses is close to the size of individual dislocation loops near the nitrogen-saturated layer. This layer, together with dislocations, effectively absorbs a large (due to small τ) acoustic power. For this reason, we did not consider the reflection of hypersound from the nitrogen-saturated layer of silicon. However, such a reflection might play a significant role in the long-range effect of ion irradiation observed, as mentioned in the Introduction, under different experimental conditions.

4. CONCLUSION

Thus, an analysis of the experimental data concerning variation of the properties of a nitrogen-saturated silicon layer and the results of the AFM examination of the argon blistering allowed a realistic explanation of the long-range effect observed in nitrogen-saturated single crystal silicon wafers at a distance of several hundred microns from the surface irradiated with argon ions. Realism of the proposed explanation is based (besides a high initial pressure and large accumulated energy) on (i) the optimum parameters of the sources (size, duration of action) ensuring effective generation of the pressure pulses reaching the rear side of a 600-μm-thick silicon wafer without significant absorption of the hypersound energy and (ii) the possibility of self-synchronized or spontaneous source switching leading to a strong amplification of the hypersound pulses produced by the explosion of blisters, with a peak pressure exceeding 10^8 Pa in the region of the nitrogen-saturated layer.

It must be noted that argon seems to be an optimum inert gas for the long-range effect of stimulated Si_3N_4 synthesis at 500°C. With neon ions of the same energy, more readily penetrating into silicon and reaching a greater depth, much greater doses will be required to provide for the explosion of blisters. Irradiation with the neon ions of lower energy would not ensure the formation of blisters with same size and cap thickness as those in the case of argon: according to Eq. (1) and Fig. 10, neon at 500°C is characterized by a higher limiting pressure. Difficulties in keeping neon under high pressure at this temperature are also related to increased permeability of silicon for smaller atoms. Our experiments showed that irradiation of the samples at T = 500°C with neon or silicon (producing no blistering) does not lead to a pronounced long-range effect. For heavier krypton and especially xenon, the limiting gas pressure in the blisters at 500°C will be significantly lower (see Eq. (1) and Fig. 10). In order to provide for a sufficient strength of the blister cap, higher ion energies will be required. With a lower pressure jump and a greater atomic mass, the process of gas escape after blister explosion will be slower and longer.

Besides the long-range effect, the unique possibility of generating high-power hypersonic self-synchronized pulses, the superplasticity and superstrength of silicon amorphized by argon ion irradiation, and the possibility of keeping an inert gas in a solid state up to T = 500°C, the phenomenon under consideration provides for a deeper insight into the physics of ion sputtering of solids. For ion energies of about tens kiloelectronvolts, the atomic sputtering by the well-known mechanism of the cascade momentum transfer from projectile to target atoms is supplemented by removal of the target material in the form of large polyatomic particles representing the caps of exploded blisters. A decrease in the ion energy leads to a decrease in the thickness of a mechanically stressed layer and, hence, to an increase in the density of dislocation loops and decorating blisters; the blisters become smaller in size and possess thinner caps. Therefore, the polyatomic particles sputtered from the target decrease with the ion energy. In this context, it is clear that the rate of ion sputtering of a given material depends on the strength of the amorphized target layer. For the sputtering by individual atoms, the effective process involving heavy ions with minimum energies at an oblique incidence is characterized by a smaller projection of the ion range onto the surface normal. Effective sputtering is favored by using chemically inert gases featuring most pronounced blistering.

ACKNOWLEDGMENTS

The authors are grateful to A.F. Khokhlov, M.F. Churbanov, D.I. Tetel’baum, V.Ya. Demikhovskiĭ, S.V. Gaponov, N.N. Salashchenko, V.V. Kurin, A.I. Titov, V.K. Kiselev, I.B. Khaĭbullin, G.A. Kachurin, N.T. Bagraev, P.A. Aleksandrov, S.I. Rembeza, A.M. Orlov, V.M. Vorotyntsev, G.P. Pashev, Yu.M. Gryaznov, S.A. Krivelevich, Mititaka Terasawa, V.D. Skupov, D.L. Zagorskiĭ, and V.N. Perevezentsev for their interest in this study, valuable remarks, and fruitful discussions. Special thanks to G.A. Maksimov and D.O. Filatov for the opportunity of performing experiments and help in the Center for Scanning-Probe Microscopy, and to D.I. Kuritsyn for his help in optical measurements.

This study was supported by the Ministry of Education of the Russian Federation within the framework of the Program “Fundamental Research in Nuclear Physics and the Physics of Ionizing Radiation Beams” (project no. 97-12-9.2-4) and the Program “Scientific Research in Higher Education for Commercial Technologies” (project no. 01.12.01.15).

REFERENCES

1. G. I. Uspenskaya, N. N. Abramova, D. I. Tetel’baum, *et al.*, in *Physical Principles of Ion-Beam Doping* (Gorki, 1972), p. 96.
2. P. V. Pavlov, V. I. Pashkov, V. M. Genkin, *et al.*, Fiz. Tverd. Tela (Leningrad) **15**, 2857 (1973) [Sov. Phys. Solid State **15**, 1914 (1973)].
3. P. V. Pavlov, E. S. Demidov, and G. V. Zorina, Fiz. Tekh. Poluprovodn. (Leningrad) **21**, 984 (1987) [Sov. Phys. Semicond. **21**, 603 (1987)].
4. E. S. Demidov, V. V. Karzanov, and P. V. Pavlov, Fiz. Tekh. Poluprovodn. (Leningrad) **23**, 548 (1989) [Sov. Phys. Semicond. **23**, 342 (1989)].
5. P. V. Pavlov, E. S. Demidov, and V. V. Karzanov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **26**, 1148 (1992) [Sov. Phys. Semicond. **26**, 643 (1992)].
6. E. S. Demidov, A. B. Gromoglasova, and V. V. Karzanov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 1025 (2000) [Semiconductors **34**, 983 (2000)].
7. P. V. Pavlov, E. S. Demidov, and V. V. Karzanov, Vysokochist. Veshchestva, No. 3, 31 (1993).
8. E. S. Demidov, V. V. Karsanov, and K. A. Markov, in *Book of Abstract of 11th International Conference on Ion Beam Modification of Materials, IBMM98, Amsterdam, 1998*, P 13.11, p. 100.
9. E. S. Demidov, V. V. Karzanov, and K. A. Markov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 170 (2000) [Semiconductors **34**, 163 (2000)].
10. V. V. Karzanov, K. A. Markov, S. Yu. Zubkov, *et al.*, in *Proceedings of the All-Russia Conference “Microprobing-99”, Nizhni Novgorod, 1999*, p. 185.
11. A. Makino, in *Chemistry of Synthesis by Combustion*, Ed. by M. Koizumi (Mir, Moscow, 1998), p. 247, translated from Japanese.
12. I. A. Abroyan, A. N. Andronov, and A. I. Titov, in *Physical Foundations of Electronic and Ionic Technology* (Vysshaya Shkola, Moscow, 1984), p. 320.
13. A. A. Ezhevskiĭ, A. F. Khokhlov, G. A. Kruglov, *et al.*, Vestn. Nizhegor. Univ., Ser.: Fiz. Tverd. Tela, No. 2, 51 (1998).
14. L. A. Margues *et al.*, J. Appl. Phys. **81**, 3 (1997).
15. A. A. Ezhevskiĭ, A. F. Khokhlov, G. A. Maksimov, *et al.*, in *Proceedings of the All-Russia Conference “Microprobing-99”, Nizhni Novgorod, 1999*, p. 159.
16. P. V. Pavlov, D. I. Tetelbaum, V. D. Skupov, *et al.*, Phys. Status Solidi A **94**, 395 (1986).
17. A. Pronyavichus and Yu. Dudonis, in *Modification of Solid Properties by Ion Beams* (Mokslas, Vilnius, 1980), p. 241.
18. *Handbook of Physical Quantities,* Ed. by I. S. Grigoriev and E. Z. Meilikhov (Énergoatomizdat, Moscow, 1991; CRC Press, Boca Raton, 1997).
19. A. F. Khokhlov, A. I. Mashin, and D. A. Khokhlov, Pis’ma Zh. Éksp. Teor. Fiz. **67**, 646 (1998) [JETP Lett. **67**, 675 (1998)].
20. A. I. Mashin and A. F. Khokhlov, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33**, 1001 (1999) [Semiconductors **33**, 911 (1999)].
21. P. V. Pavlov and A. F. Khokhlov, in *Physics of Solid State* (Vysshaya Shkola, Moscow, 1985), p. 384.
22. P. V. Pavlov, Yu. A. Semin, V. D. Skupov, and D. I. Tetel’baum, Fiz. Tekh. Poluprovodn. (Leningrad) **20**, 503 (1986) [Sov. Phys. Semicond. **20**, 315 (1986)].
23. *Physical Encyclopedia,* Ed. by A. M. Prokhorov (Sovetskaya Éntsiklopediya, Moscow, 1984), p. 944.
24. G. B. Whitham, in *Linear and Nonlinear Waves* (Wiley, New York, 1974; Mir, Moscow, 1977), p. 568.
25. J. W. Tucker and V. W. Rampton, *Microwave Ultrasonics in Solid State* (North-Holland, Amsterdam, 1972; Mir, Moscow, 1975).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 568–578.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 649–660.
Original Russian Text Copyright © 2001 by I.G. Kuleev, I.I. Kuleev.

SOLIDS
Structure

Normal Phonon–Phonon Scattering Processes and the Thermal Conductivity of Germanium Crystals with Isotope Disorder

I. G. Kuleev* and I. I. Kuleev
Institute of Metal Physics, Ural Division, Russian Academy of Sciences, Yekaterinburg, 620219 Russia
e-mail: kuleev@imp.uran.ru

Received April 4, 2001

Abstract—The influence of the normal phonon–phonon scattering processes on the thermal conductivity was theoretically studied for germanium crystals with various degrees of the isotope disorder. The theory takes into account redistribution of the phonon momentum in the normal scattering processes both inside each oscillation branch (Simons mechanism) and between various phonon oscillation branches (Herring mechanism). Contributions to the thermal conductivity due to the drift mobility of the longitudinal and transverse phonons are analyzed. It is shown that the momentum redistribution between longitudinal and transverse phonons according to the Herring relaxation mechanism leads to a significant suppression of the drift motions (and to the corresponding drop in contribution to the thermal conductivity) of the longitudinal phonons in isotopically pure germanium crystals. The results of the thermal conductivity calculations involving the Herring relaxation mechanism agree well with the experimental data available for germanium crystals with various degrees of the isotope disorder. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Extensive development of modern semiconductor technologies and microelectronics poses the task of searching for new materials possessing unusual physical properties. In this context, of special interest are the investigations in the field of isotope design [1] and the study of the physical properties of isotope-enriched germanium, silicon, and diamond crystals [2–9] widely used in modern microelectronics. The experimental investigations of thermal conductivity and thermal emf [2–4] performed for germanium crystals with various degrees of the isotope disorder showed that the maximum values of thermal conductivity in isotopically pure samples containing 99.99% ^{70}Ge are ten times and the absolute values of thermal emf are two times greater than the values observed in analogous crystals with the natural isotope composition.

In recent investigations [2, 10], a detailed theoretical analysis of the obtained experimental results was performed within the framework of the generalized Callaway model [11]. Using this approach, Asen-Palmer *et al.* [2] separated the contributions from longitudinal and transverse phonons to the thermal conductivity and determined the parameters characterizing various mechanisms of the phonon scattering in germanium crystals featuring the isotope disorder. The same generalized Callaway model was employed for calculating the thermal conductivity of isotope-enriched silicon and diamond [5, 6, 8–10]. As will be demonstrated below, this model corresponds to the assumption that relaxation of the phonon momentum proceeds separately inside each branch of the phonon spectrum. In [2], a difficulty (the same as in [11]) was encountered in the attempt to interpret the experimental data taken from [12] without separating the contributions from longitudinal and transverse phonons. The problem was that the values of the thermal conductivity of germanium calculated in the region of maximum using both the Callaway model [11] and the generalized variant [2] were considerably greater than the experimentally measured values.

For this reason, Holland [13] rejected the Callaway model [11] and demonstrated that, in calculating the thermal conductivity of germanium and silicon crystals, one must separate the contributions from longitudinal and transverse phonons because the latter phonons are characterized by a strong dispersion and the Debye temperatures of the two oscillation branches are significantly different. However, Holland included the relaxation frequencies of the normal phonon–phonon scattering processes (N processes) in the resistive relaxation frequencies. He managed to describe quite satisfactorily the temperature dependence of the thermal conductivity of germanium and silicon crystals with the natural isotope composition in a broad temperature range by doubling the number of fitting parameters. Despite this result, the Holland theory [13] was in a definite sense inferior to the Callaway model [11], because Holland did not take into account a special role of the normal phonon–phonon scattering processes.

It is well known [11, 14, 15] that the momentum of phonons involved in the N process is conserved and, hence, these processes do not contribute to the momentum relaxation and, hence, to the thermal resistance. However, the N processes form a nonequilibrium phonon distribution function and provide for relaxation of the phonon system to a locally equilibrium distribution at an average drift rate. By redistributing the energy and momentum between various phonon modes, the *N* processes prevent strong deviations of each phonon mode from the equilibrium distribution. These processes are accompanied by a change in the relative role of various resistive processes of the phonon momentum relaxation (scattering on defects and boundaries of the sample, phonon–phonon umklapp processes). Thus, this scattering mechanism plays a significant role in relaxation of the total momentum of the phonon system and produces a significant effect on the thermal conductivity in isotopically pure germanium samples at low temperatures, when the phonon–phonon umklapp processes are inactivated to a considerable extent.

However, in this case, an analysis of the role of *N* processes in the scattering of phonons belonging to different oscillation branches must be performed more thoroughly than was done previously [2–10]. It should be noted that the two-parametric Holland approximation [13] is valid in cases when the phonon relaxation frequencies in the *N* processes ($\nu^{\lambda}_{phN}(q)$) for each branch λ of the phonon spectrum are significantly smaller than the resistive frequencies ($\nu^{\lambda}_{phR}(q)$) of the phonon relaxation related to the phonon scattering on phonons in the umklapp processes and on the defects and boundaries of the sample. In the opposite limiting case realized (according to data in [2]) for the longitudinal phonons in the germanium samples studied, it is necessary to make an allowance for a drift in the phonon system related to the *N* process. In taking this factor into account, the nonequilibrium phonon subsystem is described with the aid of six parameters (including four relaxation frequencies, $\nu^{\lambda}_{phN}(q)$ and $\nu^{\lambda}_{phR}(q)$) and the drift velocity $\mathbf{u}_{\lambda}$. Description of the nonequilibrium phonon system in an extended basis provides for a more correct analysis of the phonon momentum relaxation and the thermal conductivity in isotopically pure germanium.

The purpose of this study was to elucidate the influence of the phonon drift related to the *N* process on the thermal conductivity of germanium with isotope disorder. The main attention was paid to the analysis of various models and approximations used to describe the relaxation and the momentum redistribution for the longitudinal and transverse phonons as a result of the *N* process both inside each oscillation branch (Simons mechanism [16]) and between various branches (Herring mechanism [17]). The effects of the normal phonon–phonon scattering on the mutual electron–phonon drag in metals and degenerate semiconductors were considered elsewhere [18, 19]. In this paper, in contrast to [18, 19], we will consider in more detail the Herring relaxation mechanism leading to redistribution of the phonon momentum between various oscillation branches. The role of this mechanism, which plays a major role in the *N* process in germanium, significantly increases in the case of decrease in the isotope disorder.

It will be demonstrated that the problem considered by Holland [13] and the problems arising in the treatment of experimental data [2, 12] are removed if we correctly take into account the *N* process involved in the relaxation according to Herring. In this case, redistribution of the phonon momentum between transverse and longitudinal phonons in the normal scattering processes leads to a significant suppression of the drift of longitudinal phonons in isotopically pure germanium (99.99% Ge), mostly at the expense of the transverse phonon scattering. This mechanism of the momentum redistribution accounts for a significant decrease in the contribution of longitudinal phonons to the thermal conductivity. As a result, the maximum values of the total thermal conductivity in isotopically pure germanium (99.99% Ge) decrease 1.5 times as compared to the values obtained in [2] within the framework of the generalized Callaway model. Thus, there is no need in introducing additional mechanisms of the phonon scattering on dislocations [2] and/or increasing the number of fitting parameters in the theory.

2. EFFECT OF THE NORMAL PHONON–PHONON SCATTERING ON THE MOMENTUM RELAXATION IN A NONEQUILIBRIUM PHONON SYSTEM

A kinetic equation for the nonequilibrium phonon distribution function $N^{\lambda}(\mathbf{q}, \mathbf{r})$ with an allowance for the *N* process is as follows [18]:

$$\begin{aligned}\mathbf{v}^{\lambda}_{\mathbf{q}}\nabla_{\mathbf{r}}N^{\lambda}_{\mathbf{q}} &= -(N^{\lambda}_{\mathbf{q}} - N^{0}_{q\lambda})\nu^{(1)\lambda}_{ph} \\ &- (N_{\mathbf{q}}{}^{\lambda} - N(\mathbf{q}, \mathbf{u}_{\lambda}))\nu^{\lambda}_{phN} + I_{phe}(N^{\lambda}_{\mathbf{q}}, f_0(\varepsilon_k)).\end{aligned} \tag{1}$$

Here $\mathbf{v}^{\lambda}_{\mathbf{q}} = s_{\lambda}\mathbf{q}/q$ is the group velocity of the acoustic phonons with the polarization λ; $\mathbf{u}_{\lambda}$ is the phonon drift velocity; $f_0(\varepsilon_k)$ are the Fermi distribution functions; $\nu^{\lambda}_{phN}(q)$ is the phonon relaxation frequency in the *N*-process; $\nu^{(1)\lambda}_{ph}(q) = \nu^{\lambda}_{phU}(q) + \nu^{\lambda}_{phi}(q) + \nu^{\lambda}_{phB}(q)$ is the frequency including all the nonelectron resistive phonon relaxation frequencies, caused by the phonon scattering on phonons in the umklapp processes and on the defects and boundaries of the sample; and I_{phe} is the phonon–electron collision integral [19]. Equation (1) takes into account that the *N* process leads the phonon system to a local equilibrium Planck distribution at an

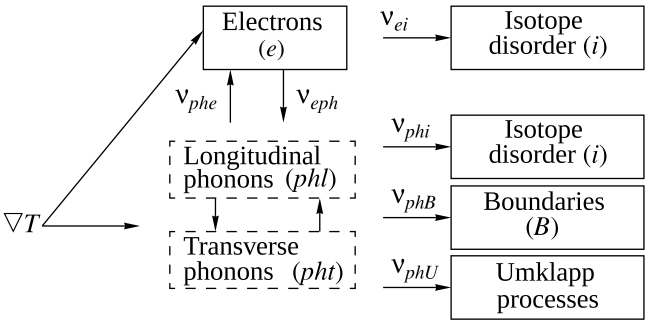


Fig. 1. A schematic diagram illustrating the phonon momentum relaxation in an electron–phonon system with an allowance for the normal phonon scattering within the framework of the Herring mechanism.

average drift velocity $\mathbf{u}_\lambda$ that can be different for the phonons possessing different polarizations λ [14, 15]:

$$N(\mathbf{q}, \mathbf{u}_\lambda) = \left[\exp\left(\frac{\hbar\omega_{p\lambda} - \hbar\mathbf{q}\cdot\mathbf{u}_\lambda}{k_B T}\right) - 1\right]^{-1} \approx N_{q\lambda}^0 + \frac{\hbar\mathbf{q}\cdot\mathbf{u}_\lambda}{k_B T} N_{q\lambda}^0 (N_{q\lambda}^0 + 1), \tag{2}$$

(here $N_{q\lambda}^0$ is the Planck function).

Figure 1 shows a schematic diagram illustrating the redistribution and relaxation of momentum acquired by the phonon system due to the temperature gradient. The mechanisms of the electron–phonon relaxation characterized by the frequencies ν_{eph} and ν_{phe} lead to the momentum redistribution inside the electron–phonon system, whereby electrons interact only with the long-wave phonons. The problem of renormalization of the phonon thermal conductivity related to the mutual entraining of electrons and phonons is not considered here. The magnitude of this effect is small because the thermal conductivity is determined from the condition of zero total current through the sample, in which case the average velocity of the ordered motion of electrons in any physically small volume of the sample is zero as well. Therefore, a momentum transferred from this ordered electron and phonon motion to the phonon subsystem is small and the influence of the nonequilibrium phonon motion on this subsystem through the conduction electron subsystem can be ignored [18, 19]. The phonon scattering on the isotope disorder (ν_{phi}) and on the boundaries (ν_{phB}) and the phonon–phonon umklapp process (ν_{phU}) lead to the relaxation of momentum of the phonon system. The N processes redistribute momentum between various phonon modes, leading to a phonon drift at an average velocity $\mathbf{u}_\lambda$. The phonon drift velocity is determined by all thermally excited phonons; the main contribution to this velocity (as well as to the thermal conductivity) is due to the thermal phonons.

Let us write the phonon distribution function in the standard form [14, 15, 18]:

$$N_{\mathbf{q}}^{\lambda} = N_{q\lambda}^{0} + g_{\lambda}(\mathbf{q}), \tag{3}$$

where $g_\lambda(\mathbf{q})$ is the nonequilibrium additive to this function. Substituting expressions (2) and (3) into (1), we obtain (by analogy with [18]) an expression for the phonon distribution function $g_\lambda(\mathbf{q})$:

$$g_\lambda(\mathbf{q}) = -\frac{N_{q\lambda}^0(N_{q\lambda}^0 + 1)}{\nu_{ph}^\lambda(q)} \frac{\hbar\omega_{q\lambda}}{k_B T^2}(\mathbf{v}_\mathbf{q}^\lambda \nabla T) + \frac{\hbar\mathbf{q}\cdot\mathbf{u}_\lambda}{k_B T} N_{q\lambda}^0(N_{q\lambda}^0 + 1)\frac{\nu_{phN}^\lambda(q)}{\nu_{ph}^\lambda(q)}. \tag{4}$$

Here, $\nu_{ph}^\lambda(q) = \nu_{phN}^\lambda(q) + \nu_{phR}^\lambda(q)$ is the total relaxation frequency of phonons with the wavevector $\mathbf{q}$ and the polarization λ, while

$$\nu_{phR}^\lambda(q) = \nu_{phi}(q) + \nu_{phL}^\lambda(q) + \nu_{phe}^\lambda(q) + \nu_{phU}^\lambda(q)$$

is the resistive relaxation frequency. The first term in the right-hand part of Eq. (4) is determined by the diffusion motion of phonons under the action of a temperature gradient. The second term takes into account the phonon drift related to the N process. Thus, we describe relaxation of the phonon system by six parameters (four relaxation frequencies, $\nu_{phR}^\lambda(q)$, and $\nu_{phN}^\lambda(q)$) and the average drift velocity $\mathbf{u}_\lambda$ (that can be different for the phonons with different polarizations), rather than by two parameters $\nu_{ph}^\lambda(q)$ representing the total relaxation frequencies as in [13]. For determining $\mathbf{u}_\lambda$, the system of kinetic equations (1) must be supplemented by the equation of the phonon momentum balance, which is obtained through multiplying Eq. (1) by the phonon momentum vector $\hbar\mathbf{q}$ and summing over all vectors $\mathbf{q}$. We will also take into account that the total momentum of the phonon subsystem does not change in the course of the N process:

$$\frac{1}{V}\sum_{\mathbf{q},\lambda}\hbar\mathbf{q}\nu_{phN}^\lambda(q)(N_\mathbf{q}^\lambda - N(\mathbf{q}, \mathbf{u}_\lambda)) = \frac{1}{V}\sum_{\mathbf{q},\lambda}\hbar\mathbf{q}\nu_{phN}^\lambda(q) \times\left[g_\lambda(\mathbf{q}) - \frac{\hbar\mathbf{q}\cdot\mathbf{u}_\lambda}{k_B T}N_{q\lambda}^0(N_{q\lambda}^0 + 1)\right] = 0. \tag{5}$$

The two known mechanisms of the normal three-phonon scattering processes include the Herring mechanism [17] and the Simons mechanism [16]. According to the Herring mechanism, the relaxation frequency of the transverse phonons is determined by the three-phonon scattering process involving one transverse and two longitudinal phonons ($t + L \longleftrightarrow L$):

$$\nu^{t}_{phN} \approx B_t T^4 \omega_t. \tag{6a}$$

The relaxation frequency of the longitudinal phonons in the anisotropic continuum model [15, 17] is determined by the three-phonon processes, whereby either a longitudinal phonon decays into two transverse phonons or two transverse phonons combine to form a longitudinal phonon ($L \longleftrightarrow t_1 + t_2$):

$$\nu^{L}_{phN} \approx B_L T^3 \omega_L^2. \tag{6b}$$

Thus, the Herring N-process relaxation mechanism involves the phonons possessing various polarizations. This relaxation mechanism provides for a redistribution of the drift momentum between longitudinal and transverse phonons. The Herring three-phonon processes in a nonequilibrium phonon system tend to establish a local equilibrium distribution at an average drift velocity equal for the phonons of both polarizations: $u_L = u_t = u_H$.

The Simons relaxation mechanism [16] involves phonons of the same polarization. In this scattering mechanism ($\nu^{\lambda}_{phN} \approx B_\lambda T^4 \omega_\lambda$), the law of the momentum conservation in the N process is valid for each branch of the phonon spectrum and, therefore, the drift velocities of the longitudinal and transverse phonons are generally different. The resistive relaxation frequency is also different for the longitudinal and transverse phonons. For this reason, we will consider two variants of the phonon momentum relaxation in the N process, whereby the phonon scattering redistributes the momentum (i) only inside each oscillation branch and (ii) predominantly between various oscillation branches. Now, we will demonstrate that a direct generalization of the Callaway formula for the lattice thermal conductivity [11, 14, 15] to the case of separated contributions of the phonons with different polarizations [2–10] corresponds to the Simons relaxation mechanism; this generalization is incorrect when the Herring mechanism dominates in the N process.

Using Eq. (4) and the equation of balance of the phonon momentum (5), we can express the average phonon drift velocity $\mathbf{u}_\lambda$ for the relaxation mechanisms according to Herring ($\mathbf{u}_H$) [17] and Simons ($\mathbf{u}_S$) [16]:

$$\begin{gathered}
\mathbf{u}_H = \frac{s_L^2(\Psi^L_N + 2S_*^3\Psi^t_N)}{k_B T(\Psi^L_{NR} + 2S_*^5\Psi^t_{NR})}(-k_B\nabla T), \\
\mathbf{u}_S = \frac{s_\lambda^2 \Psi^\lambda_N}{k_B T \Psi^\lambda_{NR}}(-k_B \nabla T).
\end{gathered} \tag{7}$$

Here,

$$\begin{gathered}
S_* = \frac{s_L}{s_t}, \quad z^{\lambda}_{q} = \frac{\hbar\omega_p\lambda}{k_B T} = \frac{q}{q_{T\lambda}}, \\
q_{T\lambda} = \frac{k_B T}{\hbar s_\lambda}, \quad z^{\lambda}_{2k} = \frac{2k}{q_{T\lambda}}
\end{gathered} \tag{8}$$

(for germanium [2]: $s_L = 5.21 \times 10^5$ cm/s; $s_t = 3.16 \times 10^5$ cm/s; $S_* = 1.65$), and the other functions are determined by the following expressions:

$$\begin{gathered}
\Psi^{\lambda}_{N} = \left\langle \frac{\nu^{\lambda}_{phN}(q)}{\nu^{\lambda}_{ph}(q)} \right\rangle_{z_{d\lambda}} \\
\equiv \int_0^{z_{d\lambda}} dz^{\lambda}_{q}(z^{\lambda}_{q})^4 \frac{\nu^{\lambda}_{phN}(q)}{\nu^{\lambda}_{ph}(q)} N^0_{q\lambda}(N^0_{q\lambda} + 1), \\
\Psi^{\lambda}_{NR} = \left\langle \frac{\nu^{\lambda}_{phR}(q)\nu^{\lambda}_{phN}(q)}{\nu^{\lambda}_{ph}(q)} \right\rangle_{z_{d\lambda}},
\end{gathered} \tag{9}$$

where $z_{d\lambda} = \hbar\omega_{d\lambda}/k_B T$ and $\omega_{d\lambda}$ is the Debye frequency of the phonons with the polarization λ. Substituting formulas (7) into expression (4) for the phonon distribution function $g_\lambda(\mathbf{q})$, we obtain

$$\begin{gathered}
g_\lambda(\mathbf{q}) = -\frac{N^0_{q\lambda}(N^0_{q\lambda} + 1)}{\tilde{\nu}^{\lambda}_{ph}(q)} \frac{\hbar\omega_q\lambda}{k_B T^2}(\mathbf{v}^{\lambda}_{q} \cdot \nabla T), \\
\tilde{\nu}^{\lambda}_{ph}(q) = \frac{\nu^{\lambda}_{ph}(q)}{1 + \nu^{\lambda}_{phN}(q)\beta_{(S,\,H)}},
\end{gathered} \tag{10}$$

where

$$\beta_S = \frac{\Psi^{\lambda}_{N}}{\Psi^{\lambda}_{NR}}, \quad \beta_H = \frac{\Psi^L_N + 2S_*^3\Psi^t_N}{\Psi^L_{NR} + 2S_*^5\Psi^t_{NR}}\left(\frac{s_L}{s_\lambda}\right)^2 \tag{11}$$

for the Simons and Herring relaxation mechanisms, respectively, and $\tilde{\nu}^{\lambda}_{ph}(q)$ is the effective relaxation frequency of the phonon momentum, renormalized by the N processes. As is well known [11, 14, 15], the lattice thermal conductivity (described with an allowance for the N processes) is determined by the effective relaxation frequency of the phonon momentum. We demonstrated that this renormalization may be different for the intra- and interbranch variants of the phonon momentum redistribution in the N process. The solution (10) transforms into the Callaway results [11] only for the Simons mechanism [16]. Evidently, under the condition $\nu^{\lambda}_{phN}(q) \ll \nu^{\lambda}_{phR}(q)$, the total frequency is $\tilde{\nu}^{\lambda}_{ph}(q) \approx \nu^{\lambda}_{phR}(q)$ and the contribution due to the phonon drift in the function $g_\lambda(\mathbf{q})$ can be ignored. In this limiting case, the Holland approximation [13] is justified. In the opposite limiting case of $\nu^{\lambda}_{phN}(q) \gg \nu^{\lambda}_{phR}(q)$, the influence of the phonon drift on the momentum exchange in the nonequilibrium phonon system has to be taken into account. Thus, in cases when the *N* processes play a significant role in the phonon momentum redistribution, the phonon subsystem should be described in an extended basis.

3. LATTICE THERMAL CONDUCTIVITY

An expression for the lattice thermal conductivity with the separated contributions from various branches of the phonon spectrum is as follows:

$$\kappa(T) = \sum_{\lambda} \frac{k_B s_{\lambda}^2 q_{T\lambda}^3}{6\pi^2} \times \int_0^{z_{d\lambda}} dz_q^{\lambda} \frac{(z_q^{\lambda})^4}{\tilde{\nu}_{ph}^{\lambda}(q)} N_{q\lambda}^0 (N_{q\lambda}^0 + 1). \tag{12}$$

Here, an allowance for the *N* processes reduces, as well as in the Callaway theory [11, 14, 15], to renormalization of the phonon relaxation frequency entering into the lattice conductivity. However, this renormalization is different in the cases when the *N* process redistributes the phonon momentum inside each branch of the phonon spectrum (Simons mechanism [16]) or predominantly between various oscillation branches (Herring mechanism [17]).

For the Herring relaxation mechanism ($\mathbf{u}_t = \mathbf{u}_L = \mathbf{u}_{ph}$),

$$\kappa_L(T) = \frac{k_B}{6\pi^2 s_L}\left(\frac{k_B T}{\hbar}\right)^3 \int_0^{z_{d\lambda}} dz \frac{z^4 e^z}{(e^z - 1)^2 \nu_{ph}^L(q)} \times \left(1 + \nu_{phN}^L \frac{\Psi_N^L + 2S_*^3 \Psi_N^t}{\Psi_{NR}^L + 2S_*^5 \Psi_{NR}^t}\right), \tag{13}$$

$$\kappa_t(T) = \frac{k_B}{3\pi^2 s_t}\left(\frac{k_B T}{\hbar}\right)^3 \int_0^{z_{d\lambda}} dz \frac{z^4 e^z}{(e^z - 1)^2 \nu_{ph}^t(q)} \times \left(1 + S_*^3 \nu_{phN}^t \frac{\Psi_N^L + 2S_*^3 \Psi_N^t}{\Psi_{NR}^L + 2S_*^5 \Psi_{NR}^t}\right). \tag{14}$$

As is seen from these expressions, the contribution from the transverse phonons to the drift terms significantly increases due to the factor S_* (for germanium, $S_* = 1.65$). If the momenta of both longitudinal and transverse phonons relax in the *N* process only inside each oscillation branch (Simons mechanism),

$$\kappa(T) = \sum_{\lambda} \frac{k_B}{6\pi^2 s_{\lambda}}\left(\frac{k_B T}{\hbar}\right)^3 \int_0^{z_{d\lambda}} dz \frac{z^4 e^z}{(e^z - 1)^2 \nu_{ph}^{\lambda}(q)} \times \left(1 + \nu_{phN}^{\lambda}(q) \frac{\Psi_N^{\lambda}}{\Psi_{NR}^{\lambda}}\right). \tag{15}$$

This expression corresponds to the generalized Callaway model [11] widely used in the calculations of thermal conductivity for isotope-enriched crystals of germanium, silicon, and diamond [2, 5, 6, 8–10].

In the Ziman limit [20], when $\nu_{phN}^{\lambda}(q) \gg \nu_{phR}^{\lambda}(q)$, we have $\nu_{ph}^{\lambda}(q) \approx \nu_{phN}^{\lambda}(q)$ and the total thermal conductivity can be expressed through the effective phonon relaxation frequency ν^*:

$$\kappa_{H,S}(T) = \frac{1}{3}\frac{C_V s_L^2}{\nu_{H,S}^*}, \tag{16}$$

where C_V is the isochoric heat capacity

$$C_V = \frac{k_B}{2\pi^2} q_{TL}^3 (J_L^{(4)} + 2S_*^3 J_t^{(4)}), \qquad J_{\lambda}^{(4)} = \int_0^{z_{d\lambda}} \frac{dz z^4 e^z}{(e^z - 1)^2}. \tag{17}$$

The effective phonon relaxation frequencies for the Herring and Simons mechanisms are different:

$$\nu_H^* = \frac{\langle \nu_{phR}^L \rangle + 2S_*^5 \langle \nu_{phR}^t \rangle}{J_L^{(4)} + 2S_*^3 J_t^{(4)}}, \qquad \nu_S^* = C_V \left[\sum_{\lambda} \left(\frac{s_{\lambda}}{s_L}\right)^2 \frac{C_{V\lambda} J_{\lambda}^{(4)}}{\langle \nu_{phR}^{\lambda} \rangle}\right]^{-1}. \tag{18}$$

Expression (15) for the additive lattice thermal conductivity representing a sum of contributions from the phonons with different polarizations can be directly obtained as a generalization of the Callaway formula [11]. This expression is valid when the *N* processes redistribute the phonon momentum only inside each oscillation branch (Simons mechanism). However, a dominating mechanism in the *N* process taking place in germanium with isotope disorder is the Herring relaxation process that leads to the momentum redistribution between various oscillation branches. Therefore, the lattice thermal conductivity of germanium samples with various degrees of isotope disorder should be calculated using formulas (13) and (14). This calculation refines the results of analysis performed in [2, 10] within the framework of the generalized Callaway model.

For the fitting parameters used in []2, the relaxation frequency of the transverse phonons in the N process is lower by three order of magnitude as compared to the value for the longitudinal phonons:

$$\nu^t_{phN}[\mathrm{s}^{-1}] \approx 2.6 \times 10^3 \left(\frac{T}{10}\right)^5 z_t,$$
$$\nu^L_{phN}[\mathrm{s}^{-1}] \approx 3.4 \times 10^6 \left(\frac{T}{10}\right)^5 z_L^2. \tag{19}$$

For the phonon scattering on the isotope disorder, the relaxation frequency of the transverse phonons is almost five times greater ($S_*^3 \approx 4.5$) than that of the longitudinal phonons:

$$\nu^t_{phi}[\mathrm{s}^{-1}] \approx 1.67 \times 10^8 g \left(\frac{T}{10}\right)^4 z_t^4,$$
$$\nu^L_{phi}[\mathrm{s}^{-1}] \approx \frac{1.67 \times 10^8 g}{S_*^3} \left(\frac{T}{10}\right)^4 z_L^4, \tag{20}$$

where $g = 0.816 \times 10^{-7}$ for ^{70}Ge (99.99%), 7.57×10^{-5} for ^{70}Ge (96.3%), and 5.87×10^{-4} for $^{\mathrm{nat}}$Ge with the natural isotope composition. For the phonon scattering on the sample boundaries,

$$\nu^\lambda_{phB}[\mathrm{s}^{-1}] = \frac{s_\lambda}{L_C}\left\{\frac{1-P}{1+P} + \frac{L_C}{l}\right\} = C_{B\lambda} \times 10^6, \tag{21}$$
$$C_{BL} = C_{Bt} S_*,$$

where L_C is the Casimir length, l is the sample length, and P is the mirror reflection probability for phonons. In the germanium crystals studied [2–4], the ν^L_{phB} value was typically about $(1–2) \times 10^6$ s^{-1}. The phonon relaxation rate in the umklapp processes was characterized by the parameters obtained in [2]

$$\nu^\lambda_{phU} = A_\lambda \left(\frac{T}{10}\right)^3 \exp\left(-\frac{C_\lambda}{T}\right) z_\lambda^2,$$

where $A_L = 1.72 \times 10^6$ s^{-1}, $A_t = 8.6 \times 10^6$ s^{-1}, $C_L = 180$ K, and $C_t = 55$ K.

As can be readily seen, the transverse phonons in the whole temperature interval obey the inequality $\nu^t_{phN}(q) \ll \nu^t_{phR}(q)$, so that the contribution of the transverse phonons to the thermal conductivity is determined predominantly by the diffusion motions. However, the longitudinal phonons are characterized by the ratio

$$\frac{\nu^L_{phN}}{\nu^L_{phR}} = 3.4\left(\frac{T}{10}\right)^5 z_L^2 \left[C_{BL} + 167 S_*^{-3} q \left(\frac{T}{10}\right)^4 z_L^4 + 1.72\left(\frac{T}{10}\right)^3 \exp\left(-\frac{180}{T}\right) z_L^2\right]^{-1}, \tag{22}$$

which is greater than unity for $T > 10$ K and is significantly greater than unity in the interval $20\ \mathrm{K} < T < 100$ K. From this, we infer that the contribution of the longitudinal phonons to the thermal conductivity is determined to a considerable extent by the drift motion of phonons in the isotope-enriched germanium crystals. Taking into account the above inequalities, the thermal conductivity can be expressed as

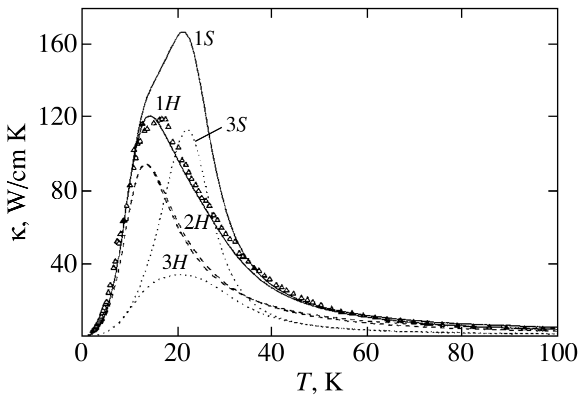


Fig. 2. Plots of the (*1*) total thermal conductivity and the contributions of (*2*) transverse and (*3*) longitudinal phonons versus temperature for a polished 99.99% ^{70}Ge crystal [2] ($C_{BL} = 1.6$; $P = 0.15$) calculated using the Herring relaxation mechanism (*H*) and the generalized Callaway model (*S*).

$$\kappa_L T = \frac{1}{3}\frac{C_{VL} s_L^2}{\langle \nu_{phR}\rangle_L},$$
$$\langle \nu_{phR}\rangle_L = \frac{\langle \nu^L_{phR}\rangle + 2S_*^5 \langle \nu^t_{phN}\rangle}{J_L^{(4)}}. \tag{23}$$

As can be seen from this formula, the effective relaxation frequency of the longitudinal phonons according to the Herring mechanism (in contrast to the generalized Callaway model [2]) contains the contributions from both the resistive processes of longitudinal phonon scattering and the normal scattering of transverse phonons. We will demonstrate that this leads to a significant suppression of the drift motions of the transverse phonons and, accordingly, to a decrease of the corresponding contribution to the thermal conductivity within the framework of the Herring relaxation mechanism in isotopically pure germanium crystals.

Below, we present the results of calculating the thermal conductivity $\kappa(T)$ for the germanium samples with various degrees of isotope disorder calculated using two variants of the phonon momentum relaxation in the N process described by formulas (13)–(15). All parameters of the theory were taken from [2]. The fitting with the aid of a single parameter C_{BL} characterizing the scattering of the longitudinal phonons on the sample

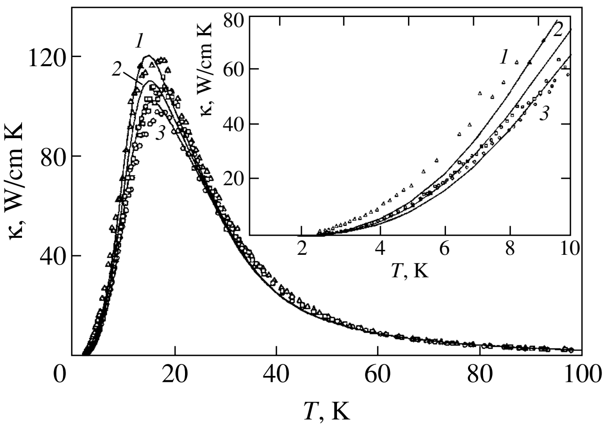


Fig. 3. Temperature dependence of the total thermal conductivity for isotopically pure germanium crystals (99.99% ^{70}Ge) with various degrees of surface finish: (*1*) ($C_{BL} = 1.6$, $P = 0.15$; (*2*) $C_{BL} = 1.9$, $P = 0.07$; (*3*) $C_{BL} = 2.33$, $P = 0$. Symbols represent the experimental data from [2, 3]; solid curves show the results of calculations for the Herring mechanism.

boundaries (for the transverse phonons, $C_{Bt} = C_{BL}/S_*$) was performed so as to provide for the best coincidence of the theoretical κ(*T*) value and the experimental data both in the region of the κ(*T*) maximum and at lower temperatures.

Figure 2 shows the temperature dependence of the thermal conductivity of a polished germanium crystal (99.99% ^{70}Ge) [2] and the contributions due to the transverse and longitudinal phonons calculated for two variants of the phonon momentum relaxation, representing the Herring mechanism and the generalized Callaway model (Simons mechanism). All parameters of the theory were taken from [2], $C_{BL} = 1.6$, and $P = 0.15$. For the generalized Callaway model ("model 1" in [2]), $\kappa^L_{\max}$ exceeds $\kappa^t_{\max}$ by a factor of 1.3 and the maximum values of the total thermal conductivity are determined predominantly by the longitudinal phonons. This result also agrees with the calculations performed in [10]. For the same parameters used with the Herring mechanism, the $\kappa^L_{\max}$ value drops by a factor of 3.4 to become 2.8 times smaller than $\kappa^t_{\max}$ (the latter remains almost unchanged). For this reason, the Herring mechanism corresponds to the dominating contribution (reaching about 80% of the total thermal conductivity) of the transverse phonons in the region of maximum.

Thus, the two variants of the phonon momentum relaxation in the *N* process lead to qualitatively different results. As can be seen from Fig. 3, the κ(*T*) values calculated for germanium samples with different surface finish containing 99.99% ^{70}Ge agree well with the experimental data both in the region of maximum and at higher temperatures. In the region of $T < 10$ K, the calculated values of thermal conductivity are noticeably lower than the experimental values measured for a polished sample (see the inset in Fig. 3). In [2, 10], this discrepancy was explained within the framework of the Ziman–Soffer theory [20, 21], according to which the calculation of thermal conductivity in the low-temperature region must take into account dependence of the probability *P* of mirror scattering (and, hence, of the relaxation frequency $\nu_{phB}(q)$) on the phonon wavelength [2, 10, 20, 21].

We may note another effect that may also lead to a discrepancy between calculated and measured values at low temperatures. As is seen from Fig. 4, the bulk value of the mean free path (MFP) of the phonons (Fig. 4b) at low temperatures is significantly greater than the effective MFP for the phonons of both polarizations (Fig. 4a) and the sample dimensions. For example, in the isotopically pure germanium sample (99.99% ^{70}Ge) (curves *1*, Fig. 4b), the bulk MFP of the transverse phonons (curve 1*t*) already at $T = 10$ K (4.8 cm) exceeds the sample length, while at $T = 5$ K this MFP is 250 times the sam-

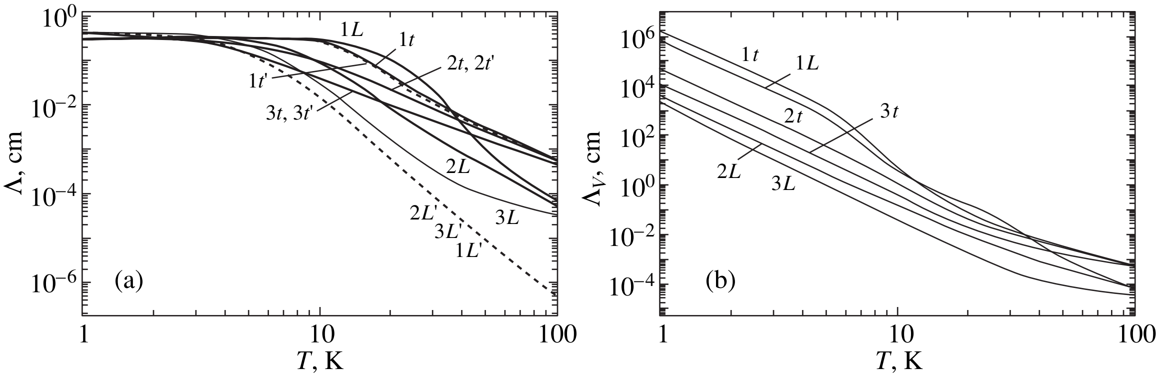


Fig. 4. Temperature dependence of (a) the effective mean free path of longitudinal (solid curves *L*) and transverse (solid lines *t*) phonons and the diffusion mean free path of phonons (dashed curves) and (b) bulk mean free path of phonons with longitudinal (*L*) and transverse (*t*) polarizations for germanium samples with different content of ^{70}Ge isotope: (*1*) 99.99%; (*2*) 96.5%; (*3*) ^{nat}Ge.

ple length. For this reason, a ballistic transport of phonons is possible at $T \leq 8$ K, which may also contribute to the thermal conductivity. The diffusion MFP of the longitudinal phonons is determined predominantly by the N processes and is not very sensitive to the degree of isotope disorder. However, since the probability of the N processes for the transverse phonons is small, the contribution of these phonons to the thermal conductivity is determined primarily by the diffusion motion. Therefore, the diffusion MFP of these phonons virtually coincides with the effective value and significantly depends on the isotope disorder.

Figure 5 shows the results of calculations of the thermal conductivity $\kappa(T)$ performed using the same parameters for the Herring relaxation mechanism. It is seen that the theory agrees well with the experimental data for germanium crystals with different degrees of isotope disorder [2, 3, 12]. Obviously, an increase in the isotope disorder leads to a decrease in the total thermal conductivity $\kappa(T)$. As a result, the difference between the variants of the interbranch (Herring mechanism) and intrabranch redistribution of the phonon momentum decreases to become less than 1% for the samples of germanium with the natural isotope composition. However, the contribution due to the drift motions of the longitudinal phonons to the thermal conductivity $\kappa_L(T)$ at $T = 15$ K exceeds the diffusion contribution by a factor of 1.5 even for the natural isotope composition. We have analyzed the contribution of the transverse phonons to the total thermal conductivity in the samples of germanium with various degrees of isotope disorder. As can be seen from Fig. 6, the transverse phonons make the main contribution to the thermal conductivity. This contribution increases with the degree of isotope disorder and accounts for more than 90% of the total thermal conductivity in the samples with the natural isotope composition ($^{\text{nat}}$Ge) in the temperature interval from 15 to 60 K. A decrease in the isotope disorder is accompanied by an increase in the phonon drift velocity. In the isotopically pure samples (99.99% ^{70}Ge), the contribution of the longitudinal phonons to the thermal conductivity increases, while that of the transverse decreases to 55% at $T = 29$ K for the Herring mechanism and to 30% at $T = 23$ K for the generalized Callaway model.

We must also note an interesting fact related to the influence of the Herring N-process relaxation on the thermal conductivity of germanium samples with various degrees of isotope disorder. For the phonon scattering on the isotope disorder, the relaxation frequency of the transverse phonons (see Eq. (20)) is approximately five times greater than that of the longitudinal phonons. It might seem that the growth of the isotope disorder must be accompanied by an increase in the contribution of longitudinal phonons to the thermal conductivity. However, the disorder growth leads to a sharp drop in the drift velocity of the longitudinal phonons, the latter value decreasing by almost two orders of magnitude in the composition range studied (see below). This circumstance significantly decreases the contribution of the longitudinal phonons to the thermal conductivity in $^{\text{nat}}$Ge as compared to that in 99.99% ^{70}Ge. The above conditions account for the successful application of the Holland theory [13] to interpretation of the thermal conductivity of germanium crystals with the natural isotope composition. Our results agree with the analysis performed in [22] using the variation approach, where the probabilities of phonon scattering were determined using the experimental values of elastic constants and a good coincidence was obtained with the data reported in [12, 23]. According to [22], 80–90% of

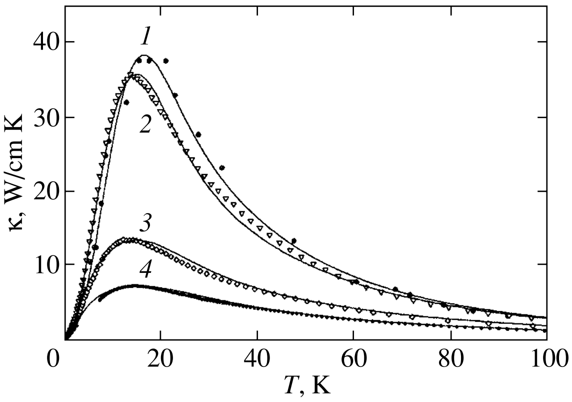


Fig. 5. Temperature dependence of the thermal conductivity for germanium crystals with various isotope compositions: (*1*) 95.8% ^{74}Ge, $g = 3.6 \times 10^{-5}$, $C_{BL} = 3.43$; (*2*) 96.3% ^{70}Ge, $g = 7.57 \times 10^{-5}$, $C_{BL} = 1.6$; (*3*) $^{\text{nat}}$Ge, $g = 5.87 \times 10^{-4}$, $C_{BL} = 1.2$; (*4*) $^{70/76}$Ge, $g = 1.53 \times 10^{-3}$, $C_{BL} = 1.35$. Symbols represent the experimental data from [2, 3, 12].

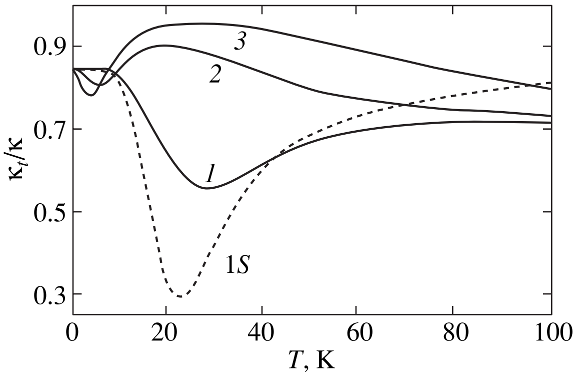


Fig. 6. Temperature variation of the contributions of transverse phonons to the total thermal conductivity for germanium crystals with various degrees of isotope disorder: (*1*) 99.99% ^{70}Ge; (*2*) 96.3% ^{70}Ge; (*3*) $^{\text{nat}}$Ge. The calculations were performed for the Herring mechanism (solid curves) and the generalized Callaway model (dashed curve *S*, 99.99% ^{70}Ge).

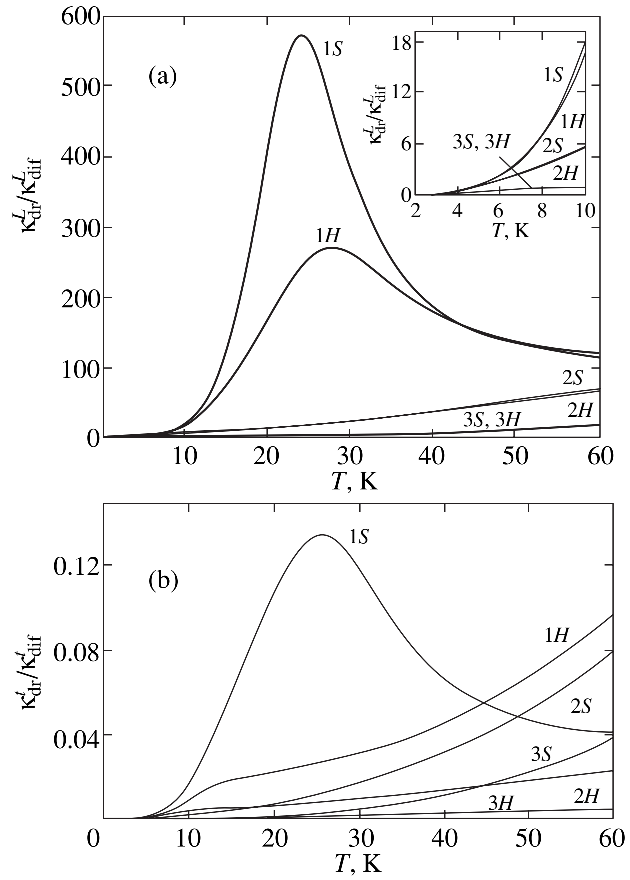


Fig. 7. Plots of the ratio $\kappa_{dr}^{\lambda}/\kappa_{dif}^{\lambda}$ of the drift and diffusion contributions to the total conductivity versus temperature for (a) longitudinal and (b) transverse phonons in germanium crystals with various degrees of isotope disorder: (*1*) 99.99% ^{70}Ge; (*2*) 96.3% ^{70}Ge; (*3*) ^{nat}Ge. The calculations were performed for the Herring mechanism (*H*) and the generalized Callaway model (Simons mechanism) (*S*).

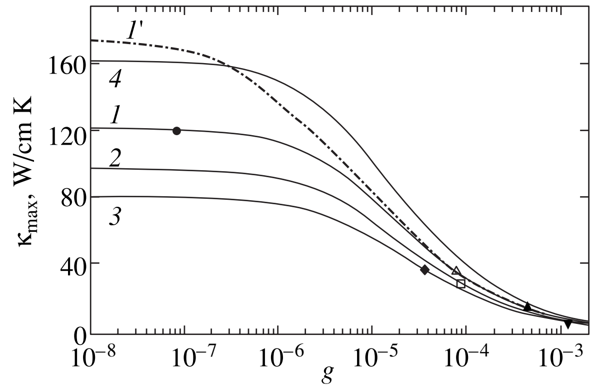


Fig. 8. Plots of the maximum thermal conductivity κ_{max} versus the degree of isotope disorder *g*. Solid curves were calculated for the Herring mechanism with C_{BL} = 1.6 (*1*), 2.48 (*2*), 3.43 (*3*), and 0.8 (*4*); dashed curve (*1*') refers to the generalized Callaway model with C_{BL} = 1.6. Symbols represent the experimental data from [2, 3, 12]: (●) $g = 8.16 \times 10^{-8}$, $C_{BL} = 1.6$; (◆) $g = 3.6 \times 10^{-5}$, $C_{BL} = 3.43$; (△) $g = 7.57 \times 10^{-5}$, $C_{BL} = 1.6$; (□) $g = 8.7 \times 10^{-5}$, $C_{BL} = 2.48$; (▲) $g = 5.87 \times 10^{-4}$, $C_{BL} = 1.2$; (▼) $g = 1.53 \times 10^{-3}$, $C_{BL} = 1.35$.

the heat in germanium crystals in the temperature interval from 2 to 300 K is transferred by the transverse phonons.

Let us consider the ratio $\kappa_{dr}^{\lambda}/\kappa_{dif}^{\lambda}$ of the drift and diffusion contributions for both branches of the phonon spectrum and the role of this factor in the thermal conductivity of germanium with various degrees of the isotope disorder. As can be seen from Fig, 7, the contribution of the drift motion of the longitudinal phonons to the thermal conductivity of 99.99% ^{70}Ge in the region of maximum decreases in the Herring mechanism approximately by half as compared to the value obtained for the generalized Callaway model (with all parameters taken from [2]). However, even in this case, the drift contribution of the longitudinal phonons is higher by two orders of magnitude than their diffusion contribution. Note that the contribution of the drift of longitudinal phonons to the thermal conductivity of 99.99% ^{70}Ge according to the Herring mechanism reaches a maximum at 28 K, where this contribution is 270 times greater than the diffusion contribution. On the other hand, the contribution due to the drift motion of the transverse phonons in the region of the thermal conductivity maximum amounts to 6% of the total value κ^t for the generalized Callaway model (Simons mechanism) and is less than 2% for the Herring mechanism. As the degree of the isotope disorder increases, both the $\kappa_{dr}^{\lambda}/\kappa_{dif}^{\lambda}$ ratio for the two branches of the phonon spectrum and the difference between the Herring and Simons variants become negligibly small.

Figure 8 illustrates the dependence of the thermal conductivity at maximum (κ_{max}) on the degree of the isotope disorder *g* in germanium crystals with various phonon relaxation frequencies at the sample boundaries (C_{BL}). Apparently, the proposed model of the interbranch redistribution of the phonon moments by the Herring *N*-process relaxation mechanism (curve *1*) provides for a much better description of the experimental data [2–4, 12] than does the generalized Callaway model (curve *1*') [2]. As can be seen from this figure for $g < 10^{-6}$ (which corresponds to 99.9% ^{70}Ge), the κ_{max} reaches saturation with respect to the isotopic purity of germanium crystals, but still significantly depends on the magnitude and character of the boundary phonon scattering. A comparison between curves *1* and *4* shows that a twofold increase in the sample cross section area must increase the κ_{max} value by one-third for 99.9% ^{70}Ge as compared to the maximum value of thermal conductivity reached in [2] for 99.99% ^{70}Ge. Since a dominating contribution to the thermal conductivity of germanium is due to the transverse phonons, it would be of interest to study how sensitive the thermal conductivity is in the region of maximum to the Debye spectrum of the transverse phonons ($w_{qt} = s_t q$). However, these calculations are outside the scope of this paper and require a special consideration.

4. CONCLUSION

We have considered the effect of the normal phonon–phonon scattering (*N* process) on the thermal conductivity of germanium crystals with various degrees of isotope disorder. The contributions of the drift motion of longitudinal and transverse phonons to the thermal conductivity were thoroughly analyzed for two variants of the phonon relaxation in the *N* process: (i) the *N* process redistributes the phonon momentum only inside each oscillation branch (Simons mechanism), and (ii) the momentum redistribution takes place predominantly between various oscillation branches (Herring mechanism). It is shown that the *N* process plays an important role in isotope-enriched germanium crystals and the results of calculations significantly depend on the mechanism of momentum relaxation in the *N* process. For the generalized Callaway model, the drift contribution of the longitudinal phonons to the thermal conductivity is almost three orders of magnitude greater than the diffusion contribution and is the main factor determining the total thermal conductivity in the region of maximum. This result agrees with the calculations performed in [10]. In this case, the contribution of the transverse phonons to the thermal conductivity is determined by their diffusion motions.

However, the main mechanism of the *N* process in germanium crystals is the Herring mechanism [2, 11, 13, 14, 17, 22]. An analysis of the thermal conductivity of germanium crystals with various degrees of isotope disorder once again confirmed this fact. We succeeded in matching the results of calculations of the thermal conductivity of germanium to experiment for the same parameters as in [2], without recourse to an additional mechanism of the phonon scattering on dislocations (the concentration of which [24] is four orders of magnitude lower than required in [2] for fitting the calculated values to the results of measurement).

It must be noted that the results of calculation of the thermal conductivity of germanium crystals for two variants of the phonon momentum relaxation in the *N* process differ both quantitatively and qualitatively. According to the generalized Callaway model (Simons mechanism), the thermal conductivity of germanium in the region of maximum is determined mostly by the longitudinal phonons, while in the Herring mechanism the major role belongs to the transverse phonons. We demonstrated that redistribution of the phonon momentum between the longitudinal and transverse phonons in the Herring *N*-process relaxation leads to a significant decrease in the drift motions of the longitudinal phonons and, accordingly, in the contribution of these phonons to the thermal conductivity.

We hope that the approach developed in this paper can be useful for studying the thermal conductivity in other systems with isotope disorder. Apparently, an increase in the degree of isotope disorder in crystals such as silicon and diamond (where the Debye frequencies of the transverse phonons are significantly higher than in germanium), would lead to a stringent dependence of the thermal conductivity on the isotope composition as a result of the “earlier” inactivation of the phonon–phonon umklapp processes.

The theory proposed in this study is based on the features of three-phonon Herring and Simons relaxation processes, while developing the general concepts introduced by Callaway [1]. The further development of the theory of lattice thermal conductivity requires analysis of the microscopic kinetic equation for the nonequilibrium phonon distribution function, with separation of the matrix of the relaxation frequencies $\nu_{phN}^{\lambda\lambda'}(q)$, describing redistribution of the phonon momentum both inside and between the oscillation branches, from the three-phonon collision integral. Extensive investigations into the physical properties of isotopically pure crystals [1–10] indicate that the solution of this problem is a currently important task in solid state physics.

ACKNOWLEDGMENTS

The authors are grateful to A.V. Inyushkin and A.N. Taldenkov for kindly providing the experimental data and valuable remarks in the course of the study, and to A.P. Tankeev and V.E. Naish for fruitful discussions and critical remarks.

This study was supported by the Russian Foundation for Basic Research, project nos. 00-02-16299 and 01-02-06238.

REFERENCES

1. V. G. Plekhanov, Usp. Fiz. Nauk **170**, 1245 (2000).
2. M. Asen-Palmer, K. Bartovski, E. Gmelin, *et al.*, Phys. Rev. B **56**, 9431 (1997).
3. V. I. Ozhogin, A. V. Inyushkin, A. N. Taldenkov, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **63**, 463 (1996) [JETP Lett. **63**, 490 (1996)].
4. A. N. Taldenkov, A. V. Inyushkin, V. I. Ozhogin, *et al.*, in *Proceedings of the IV Conference “Physicochemical Processes during Selection of Atoms and Molecules”, Zvenigorod, 1999* (Nauka, Moscow, 1999).
5. R. Berman, Phys. Rev. B **45**, 5726 (1992).
6. W. S. Capinski, H. J. Maris, E. Bauser, *et al.*, Appl. Phys. Lett. **71**, 2109 (1997).
7. Lanhua Wei, P. K. Kuo, R. L. Thomas, *et al.*, Phys. Rev. Lett. **70**, 3764 (1993).
8. J. E. Graebner, M. E. Reiss, L. Seibles, *et al.*, Phys. Rev. B **50**, 3702 (1994).
9. J. R. Olson, R. O. Pohl, J. W. Vandersande, *et al.*, Phys. Rev. B **47**, 14850 (1993).

10. A. P. Zhernov and D. A. Zhernov, Zh. Éksp. Teor. Fiz. **114**, 1757 (1998) [JETP **87**, 952 (1998)]; A. P. Zhernov, Fiz. Tverd. Tela (St. Petersburg) **41**, 1185 (1999) [Phys. Solid State **41**, 1079 (1999)].
11. J. Callaway, Phys. Rev. **113**, 1046 (1959).
12. T. H. Geballe and G. W. Hull, Phys. Rev. **110**, 1773 (1958).
13. M. G. Holland, Phys. Rev. **132**, 2461 (1963).
14. R. Berman, *Thermal Conduction in Solids* (Clarendon Press, Oxford, 1976; Mir, Moscow, 1979).
15. B. M. Mogilevskiĭ and A. F. Chudnovskiĭ, *Thermal Conductivity of Semiconductors* (Nauka, Moscow, 1972).
16. S. Simons, Proc. Phys. Soc. London **82**, 401 (1963); **83**, 799 (1963).
17. C. Herring, Phys. Rev. **95**, 954 (1954).
18. I. G. Kuleev, Fiz. Tverd. Tela (St. Petersburg) **42**, 1952 (2000) [Phys. Solid State **42**, 2004 (2000)].
19. I. G. Kuleev, Fiz. Met. Metalloved. **90** (6), 18 (2000).
20. J. M. Ziman, *Electrons and Phonons* (Clarendon, Oxford, 1960; Inostrannaya Literatura, Moscow, 1962).
21. B. Soffer, J. Appl. Phys. **38**, 1710 (1967).
22. R. A. H. Hamilton and J. E. Parrot, Phys. Rev. **178**, 1284 (1969).
23. G. A. Slack and C. J. Glassbrenner, Phys. Rev. **120**, 782 (1960).
24. K. Itoh, PhD Thesis (Univ. of California, Berkeley, 1994).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 579–585.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 661–667.
Original Russian Text Copyright © 2001 by Sannikov.

SOLIDS
Structure

A Phenomenological Theory of Possible Sequences of Ferrotoroidal Phase Transitions in Boracites

D. G. Sannikov*
Shubnikov Institute of Crystallography, Russian Academy of Sciences,
Leninskiĭ pr. 59, Moscow, 117333 Russia
**e-mail: baranov@ns.crys.ras.ru*

Received April 12, 2001

Abstract—A phenomenological theory of the sequence of two second-order phase transitions with close temperatures is considered; such transitions occur in the Ni–Br boracite. The thermodynamic potential is written as a function of polarization P_i, magnetization M_i, and toroidal moment T_i vectors and fields E_i and H_i; T_i is treated as an order parameter. It is assumed that only one coefficient of T_i^2 passes through zero as T decreases. The possibility of a sequence of two proper ferrotoroidal phase transitions along the T_1 and T_2 components is demonstrated. Spontaneous T_i, P_i, and M_i vector values and equations for susceptibility tensors (dielectric $\chi_{ij} = dP_i/dE_j$, magnetic $k_{ij} = dM_i/dH_j$, and magnetoelectric $\alpha_{ij} = dP_i/dH_j = dM_j/dE_i$) were obtained for three phases. Some of these values have well-defined anomalies in the vicinity of transitions. All possible sequences of ferrotoroidal phase transitions in boracites are considered. Depending on two potential coefficient values, these sequences may consist of one, two, or three such transitions. © *2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

A sequence of phase transitions is observed in the Ni–Br boracite ($Ni_3B_7O_{13}Br$); this sequence can conveniently be denoted by $G_0 \longrightarrow G_1 \longrightarrow G_2 \longrightarrow G_3$. The high-temperature first-order phase transition occurs at $T = 398$ K from group $T'_d = \bar{4}3m1'$ of the G_0 cubic phase to group $C'_{2v} = mm21'$ of the G_1 orthorhombic phase. Further, a sequence of two low-temperature second-order phase transitions occurs, from the G_1 phase to another orthorhombic phase G_2, symmetry group $C_{2v}(C_s) = m'm2'$, at $T = 30$ K and from G_2 to the G_3 phase, symmetry group $C_1 = 1$, at $T = 21$ K [1]. Note that the last result (group $C_1 = 1$) is questionable (see below).

In several other boracites (Co–Br [2], Co–I [3], and Ni–Cl [4]), one low-temperature second-order phase transition $G_1 \longrightarrow G_2$ was observed with the same change in symmetry groups. This transition was identified [5] as a proper ferrotoroidal phase transition (the order parameter was the T_1 component of the T_i toroidal moment), which explained the narrow temperature peak of the α_{32} component of the α_{ij} magnetoelectric tensor in the vicinity of the $G_1 \longrightarrow G_2$ phase transition in the G_2 phase (and the absence of a similar peak for the α_{23} component). Note that phase transitions with the T_i order parameter were for the first time described as a separate class of transitions in [6] (also see [7]). A phenomenological approach to the phase transition in the Ni–I boracite as a ferrotoroidal (or toroidal, there is no settled terminology yet) transition was for the first time considered in [8]. The existence of two transitions close in temperature in the Ni–Br boracite allows us to suggest that these transitions occur by the same mechanism. In other words, they can be described by a unified thermodynamic potential in which only one coefficient A of T_i^2 depends on temperature and passes through zero as temperature decreases. This means that both transitions are assumed to be proper ferrotoroidal phase transitions: the first is the $G_1 \longrightarrow G_2$ transition along the T_1 component, and the second is the $G_2 \longrightarrow G_3$ transition along the T_2 component of the T_i vector. In addition, we will consider all other possible sequences of ferrotoroidal phase transitions in boracites.

Note that the use of the same symbol, T, to denote temperature and the T_i toroidal moments should not cause confusion because T_i are always written with indices.

2. THERMODYNAMIC POTENTIAL

We will start with writing the thermodynamic potential. The x_1, x_2, x_3 coordinate system of the orthorhombic G_1 phase is convenient from the point of view of experiment. This system is rotated through 45° about the z axis of the x, y, z coordinate system of the G_0 cubic phase. The coordinates are therefore related as

$$x_1 = \frac{1}{\sqrt{2}}(x - y), \quad x_2 = \frac{1}{\sqrt{2}}(x + y), \quad x_3 = z. \quad (1)$$

1063-7761/01/9303-0579$21.00 © 2001 MAIK “Nauka/Interperiodica”

The potential will be written based on cubic G_0 phase symmetry. In this way, it is simpler to construct invariants and further use their coefficients to compare the orders of values in the obtained equations for the P_i, M_i, and T_i spontaneous vectors and the χ_{ij}, k_{ij}, and α_{ij} susceptibility tensors in the G_0, G_1, G_2, and G_3 phases.

The thermodynamic potential will be written in the form

$$\begin{aligned}\Phi = {}& \alpha R^2 - \frac{1}{2}\bar{\beta}R^4 + \frac{1}{3}\gamma R^6 + \frac{1}{2}AT_i^2 + \frac{1}{4}CT_i^4 \\ & - \bar{D}R^2T_i^2 + \frac{1}{2}\kappa P_i^2 - \sigma P_3R^2 + \frac{1}{2}BM_i^2 \\ & + cI_1 - dI_2 + aI_3 - bI_4 - P_iE_i - M_iH_i.\end{aligned} \tag{2}$$

Here, R^2 is the square of the order parameter for the $G_0 \longrightarrow G_1$ transition. This six-component R_i parameter transforms under the six-dimensional representation of the $T_d^5 = F\bar{4}3c$ space group of the G_0 cubic phase [9]. A mixed invariant proportional to P_3R^2 (and certain invariants fourth- and sixth-order in R_i) should contain angular variables in the R_i space. We assume that the minimization with respect to the angular variables have already been performed. Strictly speaking, we must take into account the R^6 invariant, because the $G_0 \longrightarrow G_1$ transition is first-order. It is assumed that spontaneous polarization in the G_1 phase is directed along the x_3 axis and that $P_3 > 0$ (the crystal is single-domain). The σ, β, and γ coefficients are positive, $\sigma > 0$, $\bar{\beta} > 0$, and $\gamma > 0$. The structural invariants have coefficients denoted by Greek letters, the coefficients of exchange invariants are denoted by capital Latin letters, and relativistic invariants, by small Latin letters. The relativistic invariants in (2) have the form

$$\begin{aligned}I_1 &= \frac{1}{4}(T_1^2 - T_2^2)^2 + (T_1^2 + T_2^2)T_3^2, \\ I_2 &= (P_1T_1 - P_2T_2)T_3 + \frac{1}{2}P_3(T_1^2 - T_2^2), \\ I_3 &= (P_2M_3 - P_3M_2)T_1 + (P_3M_1 - P_1M_3)T_2 \\ &\quad + (P_1M_2 - P_2M_1)T_3, \\ I_4 &= (M_2T_1 - M_1T_2)(T_1^2 - T_2^2) \\ &\quad - 2(M_2T_1 + M_1T_2)T_3^2 + 4M_3T_1T_2T_3.\end{aligned} \tag{3}$$

More details concerning the selection of invariants and the form of the thermodynamic potential are given in [5].

Let us eliminate the R^2 variable, which is of no interest for the further analysis, from the Φ potential to simplify it. The α coefficient, which depends on temperature T, will be replaced by a value known from experiments; for this purpose, spontaneous polarization can conveniently be used. Varying potential (2) with respect to the R^2 and P_3 variables, using the $P_3 = P_0(T)$ solution for the G_1 phase, and applying expansions usual for first-order transitions, we obtain

$$\begin{aligned}\Phi = {}& \Phi_0 + \frac{1}{2}\bar{A}T_i^2 + \frac{1}{4}\bar{C}T_i^4 + \frac{1}{2}\kappa(P_1^2 + P_2^2) \\ & + \frac{1}{2}\tilde{\kappa}P_3^2 - \tilde{\kappa}P_0P_3 - \tilde{\kappa}DP_3T_i^2 + \frac{1}{2}BM_i^2 \\ & + cI_1 - dI_2 + aI_3 - bI_4 - P_iE_i - M_iH_i.\end{aligned} \tag{4}$$

Here and throughout, we use the notation

$$\begin{aligned}& \Phi_0 = \frac{\tilde{\beta}^3}{3\gamma^2} - \frac{\beta\tilde{\kappa}^2}{2\sigma^2}P_0^2, \\ & \tilde{\beta} = \bar{\beta} + \frac{\sigma^2}{\kappa}, \quad \beta = \tilde{\beta} + \frac{\sigma^2}{\kappa}, \\ & \tilde{\kappa} = \kappa - \frac{\sigma^2}{\beta}, \quad \beta\tilde{\kappa} = \tilde{\beta}\kappa, \quad D = \frac{\sigma\bar{D}}{\beta\tilde{\kappa}}, \\ & \bar{A} = A - 2\frac{\tilde{\kappa}D}{\sigma}P_0, \quad \tilde{A} = \bar{A} - 2\tilde{\kappa}DP_0 - dP_0, \\ & \bar{C} = C - 2\frac{\bar{D}^2}{\beta}, \quad \tilde{C} = \bar{C} - 2\tilde{\kappa}D^2 + c - 2Dd, \\ & c' = c - Dd, \quad c'' = c + Dd.\end{aligned} \tag{5}$$

Note that this result [Eqs. (4), (5)] only differs from that obtained in [5], where the γR^6 invariant was not taken into account, in the expression for Φ_0, which does not enter into any of the subsequent formulas and is therefore of little significance. Equation (4) was derived using the solution for the G_1 phase, and potential (4), unlike (2), is therefore not valid for the G_0 phase. To simplify the further equations, the D coefficient is replaced in this work (as distinguished from [5]) by $\bar{D}$, and a new D coefficient expressed through $\bar{D}$ is introduced.

For the $G_1 \longrightarrow G_2$ ferrotoroidal phase transition to occur, it is necessary that coefficient $\bar{A}$ [more exactly, $\tilde{A}$, see (5)] change sign as temperature T decreases. As usual, it is assumed that this coefficient depends linearly on T,

$$\tilde{A} = A_T(T - \theta_1). \tag{6}$$

The other coefficients are assumed to be independent of T (although their temperature dependence may follow from experimental data, and this dependence should then be taken into account). The $P_0(T)$ dependence is determined experimentally for the G_1 phase (and by extrapolation for the G_2 and G_3 phases). It follows from potentials (2) and (4) [also see (5) and (6)] that $\beta > 0$, $\tilde{\beta} > 0$, $\kappa > 0$, $\tilde{\kappa} > 0$, $\tilde{C} > 0$, $B > 0$, and $\alpha_T > 0$. The D, a, b, c, and d coefficients may have arbitrary signs.

3. PHASE DIAGRAM

An analysis of potential (4) shows that, if condition (6) is met, various sequences of ferrotoroidal phase transitions are possible depending on the signs of the c and Dd coefficients. The diagram given in Fig. 1 visually represents four such possibilities corresponding to regions I–IV in the c, Dd plane. Consider them one by one. In region I, $c > 0$, $c - Dd > 0$, and the $G_1 \longrightarrow G_2 \longrightarrow G_3$ sequence of two transitions is possible, at $T = \theta_1$ along T_1 and at $T = \theta_2$ along T_2. Note that here and throughout, we say "is possible" rather than "occurs" because the θ_1, θ_2, θ_2', and θ_3 transition temperatures may formally be lower than absolute zero.

In region II, $c + Dd < 0$, $c - 2Dd < 0$, and the $G_1 \longrightarrow G_2 \longrightarrow G_3'$ sequence of two transitions is possible, at $T = \theta_1$ along T_1 and at $T = \theta_2'$ along T_3. In region III, $c < 0$, $c - 2Dd > 0$, and a sequence of three transitions, $G_1 \longrightarrow G_2 \longrightarrow G_3' \longrightarrow G_4$ is possible, at $T = \theta_1$ along T_1, at $T = \theta_2'$ along T_3, and at $T = \theta_3$ along T_2. In region IV, $c - Dd < 0$, $c + Dd > 0$, and only one $G_1 \longrightarrow G_2$ transition at $T = \theta_1$ along T_1 is possible. This situation has already been considered in [5]. In this work, a detailed analysis of cases I and II is performed.

Throughout this paper, it is assumed that P_0 is comparatively small (although the P_0 values are considered far from the $G_0 \longrightarrow G_1$ transition in which P_0 arises, P_0 grows insignificantly as temperature decreases). We also assume spontaneous T_i values to be small in the whole region of ferrotoroidal transitions. The a, b, c, and d coefficients of relativistic invariants are also small compared with the B, C, and D coefficients of exchange invariants. Expansions in all these P_0, T_i^2, a, b, c, and d values will be performed and only the highest expansion terms will as a rule be taken into account. Note that the expansion in P_0 and T_i^2 is inherent already in potentials (2) and (4), where the invariants containing the higher powers of P_0 and T_i^2 are not taken into account. Any coefficient in all expressions considered below can be replaced, for instance, coefficient κ by $\kappa + \kappa' P_0 + \kappa'' T_i^2$ (if spontaneous P_0 and T_i^2 do exist in the phase under consideration).

4. SEQUENCE OF TWO TRANSITIONS ALONG T_1 AND T_2

Consider case I. Varying thermodynamic potential (4) with respect to P_i, M_i, and T_i at $E_i = H_i = 0$ and solving the resulting equations, we obtain the following expressions for the P_i, M_i, and T_i spontaneous values in the G_1, G_2, and G_3 phases (everywhere, only nonzero components are given).

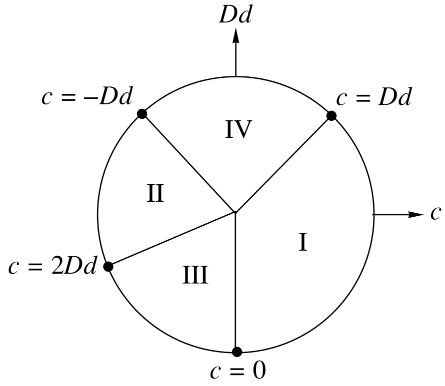


Fig. 1. Diagram in the plane of c and Dd coefficients representing four possible sequences of ferrotoroidal phase transitions in boracites.

For the G_1 phase,

$$P_3 = P_0. \tag{7}$$

For the G_2 phase,

$$\begin{gathered} T_1^2 = -\frac{\tilde{A}}{\tilde{C}} = \frac{A_T}{\tilde{C}}(\theta_1 - T), \quad P_3 = P_0 + DT_1^2, \\ M_2 = \frac{aP_0}{B}T_1 + \frac{Da + b}{B}T_1^3. \end{gathered} \tag{8}$$

It follows from (8) that the T_1 component takes on a spontaneous value in the $G_1 \longrightarrow G_2$ proper phase transition. According to (6), its square linearly depends on temperature T. Spontaneous polarization P_3 acquires a new term in the G_2 phase compared with G_1; this term is proportional to T_1^2 and arises because potential (4) includes a mixed invariant with coefficient D. The $G_1 \longrightarrow G_2$ transition is therefore simultaneously an improper ferroelectric transition (with an $f = 2$ weakness index). A spontaneous M_2 magnetization value also arises in the G_2 phase. The second term in the expression for M_2 is caused by the $bM_2T_1^3$, $\tilde{\kappa} DP_3T_1^2$, and $aM_2P_3T_1$ mixed invariants, which linearly relate M_2 to T_1^3. The $G_1 \longrightarrow G_2$ transition is therefore improper ferromagnetic (with an $f = 3$ weakness index; latent ferromagnetism according to the terminology of [10]). The first M_2 term in (8) is caused by a linear relation between M_2 and T_1 in potential (4) because of the $aP_3M_2T_1$ invariant and the $P_3 = P_0$ spontaneous value in the G_1 phase. The $G_1 \longrightarrow G_2$ transition is therefore a weak ferromagnetic transition (the aP_0 coupling constant is small). The $G_1 \longrightarrow G_2$ transition can be treated as weak ferromagnetic because of the preceding $G_0 \longrightarrow G_1$ transition, in which P_0 takes on a spontaneous value.

In the G_3 phase,

$$\begin{gathered}
T_1^2 = \frac{d}{c'}P_0\left[1+\frac{c''}{dP_0}T_2^2\right], \quad T_2^2 = \frac{c'}{2c}\frac{A_T}{\tilde{C}}(\theta_2 - T), \\
\theta_1 - \theta_2 = \frac{\tilde{C}}{A_T}\frac{d}{c'}P_0, \quad P_3 = \frac{c}{c'}P_0\left[1+\frac{2D}{P_0}T_2^2\right], \\
M_1 = -\frac{ac+bd}{Bc'}P_0T_2\left[1+\frac{2D}{P_0}T_2^2\right], \\
M_2 = \frac{ac+bd}{Bc'}P_0T_1\left[1+\frac{2D}{P_0}T_2^2\right].
\end{gathered} \tag{9}$$

According to (9), the $G_2 \longrightarrow G_3$ transition results in the appearance of a spontaneous T_2 component value, which is the order parameter for this transition. The transition is therefore a proper toroidal phase transition. Like the $G_1 \longrightarrow G_2$ transition considered above, the $G_2 \longrightarrow G_3$ transition is simultaneously improper ferroelectric ($f = 2$) with respect to the P_3 component, improper ferromagnetic ($f = 2$) with respect to M_2, and weak ferromagnetic with respect to the M_1 component, which takes on a spontaneous value in the G_3 phase along with and because of the arising of a T_2 spontaneous value. The $\theta_1 - \theta_2$ temperature difference between the two ferrotoroidal transitions is small because of the smallness of P_0 [see (9)]. For simplicity, the a, b, c, and d coefficients will be assumed to be of the same order of smallness (although this may not be the case).

5. SUSCEPTIBILITIES

We continue our analysis of case I. Let us calculate the dielectric $\chi_{ij} = dP_i/dE_j$, magnetic $k_{ij} = dM_i/dH_j$, and magnetoelectric $\alpha_{ij} = dP_i/dH_j = dM_j/dE_i$ susceptibilities. For this purpose, we will vary potential (4) twice, first with respect to P_i, M_i, and T_i and then with respect to fields E_i and H_i. Solving the equations obtained yields the following results.

For the G_0 phase,

$$\chi_{ii} = \frac{1}{\kappa}, \quad k_{ii} = \frac{1}{B}. \tag{10}$$

This result is obtained from potential (2).

For the G_1 phase,

$$\begin{gathered}
\chi_{11} = \chi_{22} = \frac{1}{\kappa}, \quad \chi_{33} = \frac{1}{\tilde{\kappa}}, \\
k_{11} = k_{33} = \frac{1}{B}, \quad k_{22} = \frac{1}{B} + \frac{a^2P_0^2}{B^2A_T(T-\theta_1)}.
\end{gathered} \tag{11}$$

It follows from a comparison of (10) and (11) that the χ_{33} component jumps up as a result of the $G_0 \longrightarrow G_1$ transition, as should be in an improper transition ($f = 2$) ferroelectric with respect to the P_3 component. The k_{22} component obeys the Curie–Weiss law in the vicinity of the $G_1 \longrightarrow G_2$ transition in the G_2 phase with a very low Curie constant proportional to $a^2P_0^2$ and, similarly, in the G_2 phase with a two times lower constant (see below). For this reason, the temperature peak of k_{22} should be very narrow in the vicinity of the $G_1 \longrightarrow G_2$ transition (Fig. 2).

For the G_2 phase, we have

$$\begin{gathered}
\chi_{11} = \chi_{22} = \frac{1}{\kappa}, \quad \chi_{33} = \frac{1}{\tilde{\kappa}} + \frac{2D^2}{\tilde{C}}, \\
k_{11} = \frac{1}{B} + \frac{(ac+bd)^2P_0}{2B^2c'^2d}\frac{\theta_1-\theta_2}{T-\theta_2}, \\
k_{22} = \frac{1}{B} + \frac{a^2P_0^2}{2B^2A_T(\theta_1 - T)}, \quad k_{33} = \frac{1}{B}, \\
\alpha_{23} = -\frac{a}{\kappa B}T_1, \\
\alpha_{32} = \frac{DaP_0}{B\tilde{C}}\frac{1}{T_1}\left\{1+\left[\frac{3(Da+b)}{\tilde{C}a}+\frac{1}{\tilde{\kappa}D}\right]\frac{\tilde{C}}{P_0}T_1^2\right\}.
\end{gathered} \tag{12}$$

Comparing (11) and (12) shows that, as expected, the χ_{33} component jumps up in the $G_1 \longrightarrow G_2$ transition (the jump does not contain small values). The k_{11} component obeys the Curie–Weiss law in the vicinity of the $G_2 \longrightarrow G_3$ transition, which causes the appearance of an $M_1 \propto T_2$ spontaneous value (see (9)), with a small Curie constant proportional to cP_0^2 and, similarly, in the G_3 phase with a two times smaller Curie constant (see below). The k_{11} temperature peak in the vicinity of the $G_2 \longrightarrow G_3$ transition should therefore be narrow (Fig. 2).

The $G_1 \longrightarrow G_2$ transition causes the appearance of the α_{23} and α_{32} off-diagonal components of the α_{ij} magnetoelectric effect tensor, which is a distinguishing feature of ferrotoroidal phase transitions. This property is related to the presence of the $aI_3 = a[\mathrm{P} \times \mathrm{M}]\mathrm{T}$ invariant, which always exists because of the transformation properties of the P_i, M_i, and T_i vectors. Unlike α_{23}, the α_{32} component has a narrow ($\propto aP_0$) temperature peak ($\propto T_1^{-1}$) in the vicinity of the $G_1 \longrightarrow G_2$ transition in the G_2 phase (Fig. 2). The difference between the temperature dependences of the α_{32} and α_{23} components can be explained as follows. According to (4), the equation for P_3 contains the DT_1^2 term (no such term is present in the equation for P_2), and the equation for T_1 contains the $CT_1^3 - aP_3M_2 + DP_3T_1$ sum. For this reason, an additional term proportional to $1/T_1$ appears as a result of the differentiation of P_3 with respect to H_2 (as distinguished from the differentiation of P_2 with respect to H_3). Similar conclusions follow from a comparison of the results obtained in differentiating M_2 and M_3 with respect to E_3 and E_2, respectively.

Narrow α_{32} component temperature peaks were observed for the $G_1 \longrightarrow G_2$ transition in several boracites (Co–Br, Co–I, and Ni–Cl [2–4]), which is evidence in favor of treating this transition as ferrotoroidal [5]. It is more difficult to explain strong changes in α_{32} (this component passes through zero two times and then increases sharply in the Co–I boracite as temperature decreases) and α_{23} (passes through zero and increases sharply in the Co–Br boracite under the same conditions). These changes may be related to sign reversal of one of the D, a, b, and c coefficients or to anomalously large values of some coefficients of the κ', κ'' type (see above). An attempt at explaining this behavior of the α_{32} and α_{23} components taking into account invariants of higher orders in P_i and T_i, which was undertaken in [5], should be considered unsuccessful.

For the G_3 phase, we have

$$
\begin{aligned}
\chi_{11} &= \chi_{22} = \frac{1}{\kappa},\\
\chi_{33} &= \frac{1}{\tilde{\kappa}} + \frac{2D^2}{\tilde{C}} + \frac{1}{\tilde{\kappa}}\left(\frac{1}{2\tilde{\kappa}} + \frac{D^2}{\tilde{C}}\right)\frac{d^2}{c},\\
\chi_{12} &= \chi_{21} = -\frac{d^2}{2\kappa^2 c}\frac{T_2}{T_1}\left[1 - \frac{c'}{dP_0}T_2^2\right],\\
k_{11} &= \frac{1}{B} + \frac{(ac+bd)^2 P_0}{4B^2 c'^2 d}\frac{\theta_1 - \theta_2}{\theta_2 - T},\\
k_{22} &= k_{33} = \frac{1}{B},\\
k_{12} &= k_{21} = \frac{(ac+bd)(ac+5bd)P_0}{8B^2 cc'd}\frac{T_1}{T_2},\\
\alpha_{13} &= \frac{ac+2bd}{\kappa Bc}T_2,\\
\alpha_{23} &= -\frac{a}{\kappa B}T_1\left[1 + \frac{2bc'}{acP_0}T_2^2\right],\\
\alpha_{31} &= \frac{(ac+bd)P_0}{2Bc}\frac{1}{T_2}\left(-\frac{D}{\tilde{C}} + \frac{d}{2\tilde{\kappa}c'}\right)\left[1 + o\left(\frac{1}{P_0}T_2^2\right)\right],\\
\alpha_{32} &= \frac{(ac+bd)}{2Bcd}T_1\left[\frac{D(c+5Dd)}{\tilde{C}} + \frac{5d}{2\tilde{\kappa}}\right]\\
&\quad\times\left[1 + o\left(\frac{1}{P_0}T_2^2\right)\right].
\end{aligned}
\tag{13}
$$

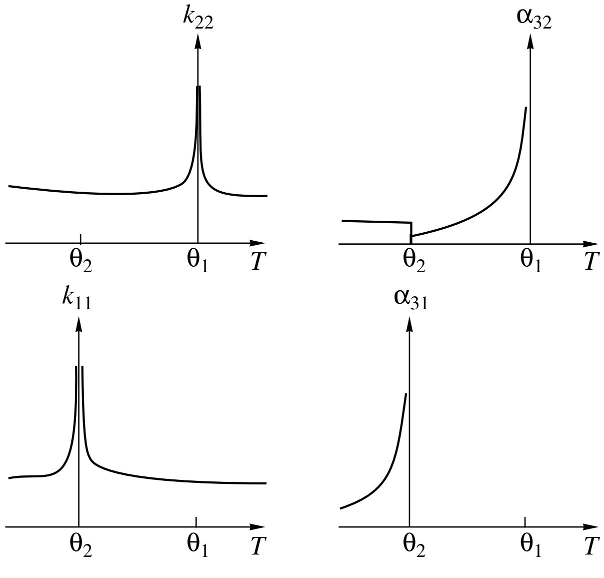


Fig. 2. Schematic drawing of temperature T dependences of k_{22}, α_{32}, k_{11}, and α_{31} in the vicinity of the $G_1 \longrightarrow G_2$ ($T = \theta_1$) and $G_2 \longrightarrow G_3$ ($T = \theta_2$) phase transitions.

Note that the terms proportional to $P_0^{-1}T_2^2$ are too cumbersome to be given; in addition, they are of no significance experimentally. A comparison of (12) and (13) shows that the $G_2 \longrightarrow G_3$ transition causes an additional, though small, jump of χ_{33} proportional to d (improper transition with $f = 2$, ferroelectric with respect to P_3) and the appearance of new $\chi_{12} = \chi_{21} \propto T_2$ components. We also have new $k_{12} = k_{21}$ components. They have a narrow temperature peak proportional to T_2^{-1} with a small constant proportional to aP_0T_1 in the G_3 phase. In addition, new α_{13} and α_{31} components arise, and α_{31}, unlike α_{13}, has a narrow temperature peak proportional to T_2^{-1} with a small constant proportional to aP_0 (Fig. 2). Comparing (12) and (13) shows that the α_{23} and α_{32} components have no singularities caused by the $G_2 \longrightarrow G_3$ transition.

As the $G_2 \longrightarrow G_3$ phase transition is a proper ferrotoroidal transition with respect to the T_2 component, the symmetry group of the G_3 phase should be $C_2(C_1) = 2'$. It follows from a comparison with case II (see below) that case I is closer than II to the experimental data on the Ni–Br boracite [1]. The observation of the well-defined temperature anomalies of the k_{11}, $k_{12} = k_{21}$, and α_{31} component predicted to accompany the $G_2 \longrightarrow G_3$ transition might serve as a test of the correctness of the suggested theory. The C_1 group of the G_3 phase determined experimentally [1] is not the closest subgroup of the $C_{2v}(C_s) = m'm2'$ group of the G_2 phase. Strictly speaking, the $G_2 \longrightarrow G_3$ second-order phase transition cannot be in phase with the $C_1 = 1$ symmetry group. This also requires experimental verification.

6. A SEQUENCE OF TWO TRANSITIONS ALONG T_1 AND T_3

We will consider case II only going into details of its differences from case I. Let the sequence of transitions be denoted by $G_0 \longrightarrow G_1 \longrightarrow G_2 \longrightarrow G_3'$, and let the corresponding temperatures be θ_0, θ_1, and θ_2'.

Instead of (9) for the G_3 phase, we have for G_3'

$$T_1^2 = -\frac{d}{c''}P_0\left[1 - \frac{3c}{dP_0}T_3^2\right],$$

$$T_3^2 = \frac{c''}{3c}\frac{A_T}{\tilde{C}}(\theta_2' - T),$$

$$P_1 = \frac{d}{\kappa}T_1T_3, \quad P_3 = \frac{c}{c''}P_0\left[1 + \frac{3D}{P_0}T_3^2\right], \tag{14}$$

$$M_2 = \frac{ac - bd}{Bc''}P_0T_1\left[1 + \frac{3D}{P_0}T_3^2\right],$$

$$\theta_1 - \theta_2' = -\frac{\tilde{C}}{A_T}\frac{dP_0}{c''}.$$

Instead of spontaneous T_2, spontaneous T_3 appears (T_3 is the order parameter of the $G_2 \longrightarrow G_3'$ phase transition). In addition, we have spontaneous P_1 but not spontaneous M_1. The $G_2 \longrightarrow G_3'$ phase transition, like $G_2 \longrightarrow G_3$, is improper ferroelectric with respect to the P_3 component ($f = 2$) and improper ferromagnetic with respect to M_2 ($f = 2$). The $\theta_1 - \theta_2'$ difference of the transition temperatures is small because of smallness of P_0.

The expressions for the χ_{ij}, k_{ij}, and α_{ij} susceptibilities differ from those obtained in case I already in the G_2 phase [cf. (8)],

$$\chi_{11} = \frac{1}{\kappa} - \frac{d^2}{\kappa^2 c''}\frac{\theta_1 - \theta_2'}{T - \theta_2'}, \quad k_{11} = \frac{1}{B}. \tag{15}$$

The expressions for the other components coincide with (8). The χ_{11} susceptibility obeys the Curie–Weiss law with a small Curie constant proportional to dP_0 in the vicinity of the $G_2 \longrightarrow G_3'$ phase transition in the G_2 phase and similarly in the G_3' phase with a two times smaller constant (see below). At the same time, the k_{11} component remains unchanged.

For the G_3' phase, we have

$$\chi_{11} = \frac{1}{\kappa} - \frac{d^2}{2\kappa^2 c''}\frac{\theta_1 - \theta_2'}{\theta_2' - T}, \quad \chi_{22} = \frac{1}{\kappa},$$

$$\chi_{33} = \frac{1}{\tilde{\kappa}} + \frac{2D^2}{\tilde{C}} + o(c),$$

$$\chi_{13} = \chi_{31} = \frac{d}{2\kappa c}\frac{T_1}{T_3}\left(\frac{d}{2\tilde{\kappa}} + \frac{Dc''}{\tilde{C}}\right)\left[1 + o\left(\frac{1}{P_0}T_3^2\right)\right],$$

$$k_{11} = k_{22} = k_{33} = \frac{1}{B},$$

$$k_{13} = k_{31} = -\frac{(ac - bd)b}{B^2cd}T_1T_3, \tag{16}$$

$$\alpha_{12} = \frac{(ac - 7bd)dP_0}{6\kappa Bcc''}\frac{1}{T_3}\left[1 + o\left(\frac{1}{P_0}T_3^2\right)\right],$$

$$\alpha_{21} = \frac{5ac - bd}{4\kappa Bc}T_3,$$

$$\alpha_{23} = -\frac{a}{\kappa B}T_1\left[1 + \frac{bc''}{ac}\frac{1}{P_0}T_3^2\right],$$

$$\alpha_{32} = \frac{ac - bd}{3Bcd}T_1\left[\frac{7d}{2\tilde{\kappa}} + \frac{D(ac + 5Dd)}{\tilde{C}}\right]\left[1 + o\left(\frac{1}{P_0}T_3^2\right)\right].$$

Again, the terms proportional to $P_0^{-1}T_3^2$ are not given because they are too cumbersome and useless experimentally. The $\chi_{13} = \chi_{31}$ components appear; they have a temperature peak proportional to T_3^{-1} with a constant proportional to T_1. The k_{11} component does not obey the Curie–Weiss law [cf. (8)]. The α_{12} and α_{21} components replace α_{13} and α_{31}, and α_{12} (like α_{31} in the $G_2 \longrightarrow G_3$ transition) has a temperature peak proportional to T_3^{-1} with a constant proportional to dP_0. The α_{23} and α_{32} components exhibit regular behavior, as in (8). The $G_2 \longrightarrow G_3'$ phase transition is a proper ferrotoroidal transition with respect to the T_3 component, and the symmetry group of the G_3' phase should therefore be $C_s = m$.

7. CONCLUSION

We will not consider case III because, experimentally, the $G_0 \longrightarrow G_1 \longrightarrow G_2 \longrightarrow G_3' \longrightarrow G_4$ sequence was not observed in any of the boracites. Note only that

$$\theta_2' - \theta_3 = \frac{3cdP_0\tilde{C}}{(c - 2Dd)c''A_T}, \tag{17}$$

where θ_3 is the temperature of the $G_3' \longrightarrow G_4$ proper transition ferrotoroidal with respect to the T_2 component. The $\theta_2' - \theta_3$ difference is small in so far as P_0 is small. The $G_3' \longrightarrow G_4$ transition causes the appearance of all P_i, M_i, and T_i vectors and χ_{ij}, k_{ij}, and α_{ij} tensor components lacking in the G_3' phase. The symmetry group of the G_4 phase is $C_1 = 1$.

Case IV was considered in [5] and virtually reproduced in this work (see (7), (8), and (10)–(12)). It was stressed in [5] that a closed system of spins formed a structure with a spontaneous toroidal moment normal to the plane of spins and that, consequently, such a structure of spins should be sought in boracites. However, a system of antiparallel spins should also have a toroidal moment normal to their plane if the distances between oppositely directed spins are not equal. Note that the explanation of the α_{32} temperature peak in boracites was given earlier [11]. The $G_1 \longrightarrow G_2$ phase transition was treated in [11] as an antiferromagnetic transition. We would like to stress that, from the point of view of the phenomenological theory of phase transitions [12], an important thing is the transformation properties rather than the physical meaning of the order parameter that is, the representation of the initial symmetry group of the crystal that governs order parameter transformations. For instance, consider the highest magnetic symmetry group O_h'. The M_i magnetization vector transforms according to one of four magnetic three-dimensional representations, and the L_i antiferromagnetic vectors transform according to the remaining three representations. In other words, there exist three antiferromagnetic vectors that differ from each other in their transformation properties. The T_i toroidal moment vector transforms as one of them. It follows that T_i can be said to be a special case of L_i in its transformation properties.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 00-02-17746).

REFERENCES

1. S.-Yu Mao, H. Schmid, G. Triscone, and J. Muller, J. Magn. Magn. Mater. **195**, 65 (1999).
2. M. Clin, J.-P. Rivera, and H. Schmid, Ferroelectrics **79**, 173 (1988).
3. M. Clin, J.-P. Rivera, and H. Schmid, Ferroelectrics **108**, 213 (1990).
4. J.-P. Rivera and H. Schmid, J. Appl. Phys. **70**, 6410 (1991).
5. D. G. Sannikov, Zh. Éksp. Teor. Fiz. **111**, 536 (1997) [JETP **84**, 293 (1997)].
6. V. L. Ginzburg, A. A. Gorbatsevich, Yu. V. Kopaev, and B. A. Volkov, Solid State Commun. **50**, 339 (1984).
7. B. A. Volkov, A. A. Gorbatsevich, Yu. V. Kopaev, and V. V. Tugushev, Zh. Éksp. Teor. Fiz. **85**, 729 (1981) [Sov. Phys. JETP **54**, 391 (1981)].
8. D. G. Sannikov and I. S. Zheludev, Fiz. Tverd. Tela (Leningrad) **27**, 1369 (1985) [Sov. Phys. Solid State **27**, 826 (1985)].
9. V. Dvorak and J. Petzelt, Czech. J. Phys. B **21**, 1141 (1971); V. Dvorak, Czech. J. Phys. B **21**, 1250 (1971).
10. I. E. Dzyaloshinskiĭ and V. I. Man'ko, Zh. Éksp. Teor. Fiz. **46**, 1352 (1964) [Sov. Phys. JETP **19**, 915 (1964)].
11. I. E. Chupis, Fiz. Nizk. Temp. **18**, 306 (1992) [Sov. J. Low Temp. Phys. **18**, 210 (1992)].
12. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1995; Pergamon, Oxford, 1980), Part 1.

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 586–595.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 668–677.
Original Russian Text Copyright © 2001 by Balagurov.

SOLIDS
Electronic Properties

On the Conductivity of Composites with Two-Dimensional Periodic Structure

B. Ya. Balagurov
Emanuel' Institute of Biochemical Physics, Russian Academy of Sciences, Moscow GSP-1, 119991 Russia
e-mail: balagurov@deom.chph.ras.ru

Received January 26, 2001

Abstract—A method is suggested of successive solution of the problem on the conductivity of two-dimensional periodic systems with inclusions of arbitrary shape. The complex potential outside of the inclusions is expressed in terms of the Weierstrass zeta function and its derivatives. The field induced on a separate inclusion is described using the matrix of multipole polarizabilities. The "joining" of potentials is performed at a distance ρ such that $R < \rho < a$, where R is the characteristic dimension (maximum "radius") of the inclusion and a is the half-period of the lattice. The approach suggested enables one to find exact virial expansions for the conductivity of other effective characteristics of similar systems as well. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The investigation of the conductivity and other electrophysical characteristics of inhomogeneous unordered systems (in particular, composites) involves certain mathematical difficulties [1]. Therefore, the results available for the electrical conductivity of such media have largely been obtained using model and numerical experiments. In the case of composites with a periodic structure, the situation is more favorable from the theoretical standpoint, especially, as regards two-dimensional systems. In this case, it is sufficient to find the potential within a single elementary cell, whereby the problem is simplified considerably (though still remaining fairly complex). Note that the study of the electrical properties of composites with a periodic structure is of considerable interest from both the generally physical (the problem of metal-dielectric phase transition) and applied (microelectronics) standpoints.

The analytical solution to the problem on conductivity is given for a number of two-dimensional doubly periodic systems [1–3] using the methods of the theory of functions of complex variables. Treated in [1–3] are various models of systems with dielectric or ideally conducting inclusions, whereby it is possible to restrict oneself to solving the external problem. A closed solution in the case of finite (nonzero) conductivity of both components is given in [1] for a model with a chessboard structure. In [4, 5], a scheme is developed for successively finding virial expansions for various effective characteristics of the model (first treated by Rayleigh [6]) of a thin film with circular inclusions that form a square (rectangular) lattice. Note that each of the methods used in [1–6] is valid for some single model with inclusions of a concrete shape. At the same time, no unified approach exists to this problem.

In this paper, I suggest a method of successively obtaining virial expansions for the conductivity and other effective characteristics of composites with a two-dimensional periodic structure formed by inclusions of arbitrary shape. The complex potential outside of the inclusions is expressed in terms of the Weierstrass zeta function [7, 8] and its derivatives. As to the properties of concrete inclusion, they enter the solution of the problem in the form of multipole polarizabilities of this inclusion, i.e., of respective coefficients in "responses" to various external fields.

This approach enables one to fully solve the "lattice" part of the problem by reducing the initial problem of "inclusion in elementary cell" to the problem of finding the response of solitary inclusion to an external field with the asymptotic forms preassigned at infinity. As to the problem of finding multipole polarizabilities, it must be solved as an independent problem in each concrete case using analytical or numerical methods. I found, by way of example, a complete matrix of polarizabilities for inclusions of elliptic shape. Note that I have treated the case of a square lattice that is the simplest for analysis, although the method suggested is valid for lattices of different symmetry as well. The treatment is performed within the framework of macroscopic electrodynamics, so that it is assumed, in particular, that the characteristic dimensions of inclusions are large compared with the free path of carriers.

2. MULTIPOLE POLARIZABILITIES

In what follows, we will require the solution of the problem on the response of separate inclusion (body) to the external electric field. If a uniform field of intensity $\mathbf{E}_0$ is applied, the electric potential φ in a dipole approximation has the following form (two-dimensional case) at large distances from the body:

$$\varphi(\mathbf{r}) = -\mathbf{E}_0 r + 2\frac{\mathbf{pr}}{r^2} + \ldots, \quad r \longrightarrow \infty. \tag{1}$$

Here,

$$\mathbf{p} = \hat{\Lambda}\mathbf{E}_0 \tag{2}$$

is the dipole moment of the inclusion, and $\hat{\Lambda}$ is the tensor of dipolar polarizability.

If $\mathbf{E}_0$ is directed along one of the principal axes of the tensor $\hat{\Lambda}$ (we will select this axis as the coordinate axis x), then

$$\varphi(\mathbf{r}) = -E_0\left\{x - 2\frac{x\Lambda^{(x)}}{r^2} + \ldots\right\}, \quad r \longrightarrow \infty, \tag{3}$$

where $\Lambda^{(x)}$ is the corresponding principal value of tensor $\hat{\Lambda}$. The value of $\Lambda^{(x)}$ (as well tensor Λ itself) is proportional to the area (volume in the 3D case) of inclusion s:

$$\Lambda^{(x)} = s\alpha^{(x)}, \tag{4}$$

where $\alpha^{(x)}$ is the dimensionless dipolar polarizability depending on the shape of the body and on the argument $h = \sigma_2/\sigma_1$, and σ_2 and σ_1 denote the conductivity of the inclusion and of the surrounding medium, respectively.

Below, it will be convenient to use the complex potential $\Phi(z)$, the derivative of which is related to the components of the electric field intensity $\mathbf{E}$ as

$$\Phi'(z) = -E_x + iE_y. \tag{5}$$

The real part of $\Phi(z)$ gives the electric potential $\varphi(\mathbf{r})$,

$$\varphi = \mathrm{Re}\Phi(z). \tag{6}$$

The complex potential corresponding to expression (3) has the form

$$\Phi(z) = -E_0\left\{z - \frac{2\Lambda^{(x)}}{z} + \ldots\right\}, \quad |z| \longrightarrow \infty, \tag{7}$$

with the real constant $\Lambda^{(x)}$.

When the higher order (multipole) moments are included, the expression for $\Phi(z)$ at high values of z takes the form

$$\Phi(z) = z + \sum_{m=0}^{\infty}\frac{\Lambda^{(x)}_{1,2m+1}}{z^{2m+1}}, \quad |z| \longrightarrow \infty. \tag{8}$$

In Eq. (8), the common factor is omitted; for simplifying the computation, it is assumed that the inclusion has a fairly symmetric shape, so that the complex potential is even with respect to z, and the quantities $\Lambda^{(x)}_{1,2m+1}$ are real. A comparison of Eq. (8) with (7) reveals that

$$\Lambda^{(x)}_{11} = -2\Lambda^{(x)} = -2s\alpha^{(x)}. \tag{9}$$

In what follows, we will further require the response of inclusion to a nonuniform external field of the form

$$\mathrm{Re}z^{2n+1} = r^{2n+1}\cos(2n+1)\theta,$$

where θ is the polar angle. In this case, we have, similarly to Eq. (8),

$$\Phi(z) = z^{2n+1} + \sum_{m=0}^{\infty}\frac{\Lambda^{(x)}_{2n+1,2m+1}}{z^{2m+1}}, \quad |z| \longrightarrow \infty, \tag{10}$$

with real constants $\Lambda^{(x)}_{2n+1,2m+1}$, which will be referred to as multipole polarizabilities. Note that equality (10) may also be represented as

$$\Phi(z) = z^{2n+1} + \sum_{m=0}^{\infty}\frac{\Lambda^{(x)}_{2n+1,2m+1}}{(2m)!}\left(\frac{d^{2m}}{dz^{2m}}\frac{1}{z}\right), \quad |z| \longrightarrow \infty. \tag{11}$$

The even–even multipole polarizabilities $\Lambda_{2n,2m}$ are introduced in a similar manner.

Note that Dykhne's symmetry transformation [9] enables one to relate the complex potentials of the initial and so-called reciprocal (differing from the initial one by the replacement $h \longrightarrow 1/h$) $\tilde{\Phi}(z)$ systems (compare with [3]),

$$\Phi^{(x)}(z) = i\tilde{\Phi}^{(y)}(z). \tag{12}$$

Here, the superscript $\nu = x$ at $\Phi^{(\nu)}(z)$ implies that, in the asymptotic forms of the electric potential $\varphi = \mathrm{Re}\Phi^{(\nu)}(z)$, the leading term has the form of $\mathrm{Re}z^{2n+1}$, and $\nu = y$ implies accordingly the form $\mathrm{Im}z^{2n+1}$ for this leading term. Here and below, the tilde indicates the quantities pertaining to the reciprocal system. For $\Phi^{(y)}(z)$, we have an expansion similar to that given by Eq. (10),

$$\Phi^{(y)}(z) = -i\left\{z^{2n+1} - \sum_{m=0}^{\infty}\frac{\Lambda^{(y)}_{2n+1,2m+1}}{z^{2m+1}}\right\}, \quad |z| \longrightarrow \infty. \tag{13}$$

The substitution of Eqs. (10) and (13) into (12) gives the relation

$$\tilde{\Lambda}^{(y)}_{2n+1,2m+1} = -\Lambda^{(x)}_{2n+1,2m+1}. \tag{14}$$

The even–even polarizabilities $\Lambda_{2n,2m}$ are related by the same relation.

From considerations of dimensionality, it follows from Eq. (10) that

$$\Lambda_{nm} = R^{n+m}\alpha_{nm}, \tag{15}$$

where R is the characteristic dimension (in the xy plane) of inclusion, and α_{nm} denotes dimensionless quantities dependent on the shape of inclusion and on the argument h.

3. ELECTRIC FIELD IN A MEDIUM

We will treat a two-dimensional system with identical (and identically oriented) inclusions forming a square lattice with a period of $2a$. We will assume that the principal axes of the tensors of polarizability of inclusions coincide with the axes of the lattice and with the x and y coordinate axes. Then, if the average (with respect to the elementary cell area) intensity of electric field $\langle \mathbf{E} \rangle$ is directed along the x axis, all quantities Λ_{nm} in Eqs. (8) and (10) are real. The problem of finding the potential is solved with the aid of expansion in terms of the formally small parameter R/a, where, in this case, R is the maximum "radius" of inclusion.

In zero approximation, the complex potential corresponding to the uniform external field applied along the x axis has the form

$$\Phi^{(0)}(z) = \beta z. \tag{16}$$

The response of the inclusion located at the origin of coordinates to the field given by Eq. (16) is given, according to Eq. (8), by the expression

$$\begin{aligned}\Phi_I^{(1)}(z) &= \beta \sum_{n=0}^{\infty} \frac{\Lambda_{1,2n+1}}{z^{2n+1}} \\ &= \beta \sum_{n=0}^{\infty} \frac{\Lambda_{1,2n+1}}{(2n)!}\left(\frac{d^{2n}}{dz^{2n}}\frac{1}{z}\right).\end{aligned} \tag{17}$$

Here and below, we omit the superscript x at $\Lambda_{nm}^{(x)}$. We add up all responses of the type given by Eq. (17) from all inclusions to derive the following expression for the first-approximation correction to Eq. (16):

$$\Phi^{(1)}(z) = \beta \sum_{n=0}^{\infty} B_{2n}^{(1)} \zeta^{(2n)}(z), \tag{18}$$

$$B_{2n}^{(1)} = \frac{1}{(2n)!}\Lambda_{1,2n+1}. \tag{19}$$

In equality (18),

$$\zeta(z) = \frac{1}{z} + \sum_{k,l}{}' \left[\frac{1}{z - z_{kl}} + \frac{1}{z_{kl}} + \frac{z}{(z_{kl})^2}\right] \tag{20}$$

is the Weierstrass zeta function [7, 8], $\zeta^{(2n)}(z)$ is the derivative of the order $2n$ of $\zeta(z)$, and $z_{kl} = 2(k + il)a$. The function $\zeta(z)$ (the term with $n = 0$ in Eq. (18)) arises as a result of summation of the dipole potentials induced by the external field. In this case, as in [4, 5], the respective sum is regularized, which provides for its convergence. The terms with $n \geq 1$ in Eq. (18) correspond to higher order multipoles.

In the next approximation, the external (with respect to the selected inclusion) potential is provided by the quantity $\Phi^{(1)}(z)$ from Eq. (18) less the eigenfield, i.e., the potential $\Phi_I^{(1)}(z)$ from Eq. (17),

$$\begin{aligned}&\Phi^{(1)}(z) - \Phi_I^{(1)}(z) \\ &= \beta \sum_{n=0}^{\infty} B_{2n}^{(1)}\left\{\zeta^{(2n)}(z) - \frac{(2n)!}{z^{2n+1}}\right\}.\end{aligned} \tag{21}$$

According to [7], the expansion

$$\zeta(z) = \frac{1}{z} - \sum_{m=2}^{\infty} \frac{c_m}{2m-1} z^{2m-1} \tag{22}$$

holds for the zeta function.

Expressions for the coefficients c_m are given, for example, in [7] (see also [4, 5]). In particular, for a square lattice, all coefficients c_m with odd subscripts are zero. Equation (22) yields, for $\zeta^{(2n)}(z)$,

$$\begin{aligned}\zeta^{(2n)}(z) &= \frac{(2n)!}{z^{2n+1}} \\ &- \sum_{k=0}^{\infty} \frac{(2n+2k)!}{(2k+1)!} c_{n+k+1} z^{2k+1},\end{aligned} \tag{23}$$

so that Eq. (21) takes the form

$$\begin{aligned}&\Phi^{(1)}(z) - \Phi_I^{(1)}(z) \\ &= -\beta \sum_{k=0}^{\infty}\left\{\sum_{m=0}^{\infty} B_{2m}^{(1)} \frac{(2k+2m)!}{(2k+1)!} c_{k+m+1}\right\} z^{2k+1}.\end{aligned} \tag{24}$$

According to Eq. (11), the response of selected inclusion to the external stimulation given by Eq. (24) is

$$\begin{aligned}\Phi_I^{(2)}(z) &= -\beta \sum_{m=0}^{\infty} B_{2m}^{(1)} \sum_{k=0}^{\infty} \frac{(2k+2m)!}{(2k+1)!} c_{k+m+1} \\ &\times \sum_{n=0}^{\infty} \frac{\Lambda_{2k+1,2n+1}}{(2n)!}\left(\frac{d^{2n}}{dz^{2n}}\frac{1}{z}\right).\end{aligned} \tag{25}$$

The summation of responses of the form of (25) from all inclusions gives the contribution of the second approximation to the total potential

$$\Phi^{(2)}(z) = \beta \sum_{n=0}^{\infty} B_{2n}^{(2)} \zeta^{(2n)}(z), \tag{26}$$

where

$$B_{2n}^{(2)} = -\sum_{m=0}^{\infty} P_{nm} B_{2m}^{(1)} \tag{27}$$

with $B_{2n}^{(1)}$ from Eq. (19) and

$$P_{nm} = \sum_{k=0}^{\infty} \frac{(2k+2m)!}{(2k+1)!} c_{k+m+1} \frac{\Lambda_{2k+1,2n+1}}{(2n)!}. \tag{28}$$

We continue this procedure to conclude that the total potential outside of the inclusions has exactly the same form as in the case of circular inclusions [4, 5],

$$\Phi(z) = \beta\left\{z + \sum_{n=0}^{\infty} B_{2n}\zeta^{(2n)}(z)\right\}. \tag{29}$$

Here, the corrections to the coefficients B_{2n} of the $(N+1)$th and Nth approximations are related by

$$B_{2n}^{(N+1)} = -\sum_{m=0}^{\infty} P_{nm}B_{2m}^{(N)}, \tag{30}$$

so that

$$\begin{aligned} B_{2n} &= \sum_{N=0}^{\infty} B_{2n}^{(N+1)} \\ &= \sum_{N=0}^{\infty}(-1)^N \sum_{m=0}^{\infty}(\hat{P}^N)_{nm}B_{2m}^{(1)} \end{aligned} \tag{31}$$

or

$$\begin{aligned} B_{2n} &= B_{2n}^{(1)} - \sum_{k} P_{nk}B_{2k}^{(1)} + \sum_{k,l} P_{nk}P_{kl}B_{2l}^{(1)} \\ &\quad - \sum_{k,l,m} P_{nk}P_{kl}P_{lm}B_{2m}^{(1)} + \ldots. \end{aligned} \tag{32}$$

It is easy to check that the quantities B_{2n} from Eq. (32) satisfy the equation

$$B_{2n} + \sum_{m=0}^{\infty} P_{nm}B_{2m} = B_{2n}^{(1)} \tag{33}$$

with $B_{2n}^{(1)}$ from Eq. (19) and P_{nm} from Eq. (28). Expressions (29) and (33) with (28) give a formal solution to the problem set.

The case when $\langle\mathbf{E}\rangle$ is directed along the y axis is treated analogously. The appropriate quantities will be provided with an overscribed bar, so that $\Phi^{(y)}(z) = \bar{\Phi}(z)$, $\Lambda_{nm}^{(y)} = \bar{\Lambda}_{nm}$, and so on. We use formula (13) for searching for the response and derive the following expression for the potential outside of the inclusions:

$$\bar{\Phi}(z) = -i\beta\left\{z - \sum_{n=0}^{\infty} \bar{B}_{2n}\zeta^{(2n)}(z)\right\}, \tag{34}$$

where the coefficients $\bar{B}_{2n}$ satisfy the equation

$$\bar{B}_{2n} - \sum_{m=0}^{\infty} \bar{P}_{nm}\bar{B}_{2m} = \bar{B}_{2n}^{(1)}. \tag{35}$$

Here,

$$\bar{B}_{2n}^{(1)} = \frac{1}{(2n)!}\bar{\Lambda}_{1,2n+1} \equiv \frac{1}{(2n)!}\Lambda_{1,2n+1}^{(y)}; \tag{36}$$

the matrix $\bar{P}_{nm}$ coincides in shape with P_{nm} from Eq. (28) and differs from the latter only by the substitution of $\Lambda_{2k+1,2n+1}$ by $\bar{\Lambda}_{2k+1,2n+1}$. Note that the substitution of Eqs. (29) and (34) into equality (12) written in the form $\tilde{\Phi}(z) = i\bar{\Phi}(z)$ gives the correlation between the coefficients

$$\bar{B}_{2n} = -\tilde{B}_{2n}. \tag{37}$$

For circular inclusions of radius R, the matrix of polarizabilities has a diagonal form,

$$\Lambda_{nm} = R^{2n}\frac{1-h}{1+h}\delta_{nm}; \quad h = \frac{\sigma_2}{\sigma_1}. \tag{38}$$

In this case, from Eq. (33) follows the equation

$$\begin{aligned} B_{2n} + \frac{1-h}{1+h}\sum_{m=0}^{\infty} B_{2m}\frac{(1n+2m)!}{(2n)!(2n+1)!}R^{4n+2}c_{n+m+1} \\ = R^2\frac{1-h}{1+h}\delta_{n0}, \end{aligned} \tag{39}$$

which coincides with equality (A.6) from [4].

4. EFFECTIVE CHARACTERISTICS

4.1. Conductivity

We use the complex potential $\Phi(z)$ from Eq. (29) to calculate the drop of voltage U_x on an elementary cell and the total current I_x through the latter in the direction of the x axis to derive, analogously with [4, 5],

$$\begin{aligned} U_x &= -2\beta a\left(1 + B_0\frac{\pi}{4a^2}\right), \\ I_x &= -2\sigma_1\beta a\left(1 - \beta_0\frac{\pi}{4a^2}\right). \end{aligned} \tag{40}$$

We find

$$\sigma_{xe} = \sigma_1\left(1 - B_0\frac{\pi}{4a^2}\right)\Big/\left(1 + B_0\frac{\pi}{4a^2}\right) \tag{41}$$

for the effective conductivity $\sigma_{xe} = I_x/U_x$ in the direction of the x axis (i.e., for the respective principal value σ_{xe} of the effective conductivity tensor $\hat{\sigma}_e$).

If the inclusion "radius" R is small compared with the lattice half-period a, then, according to Eqs. (19) and (9), we have

$$B_0 \approx B_0^{(1)} = \Lambda_{11}^{(x)} = -2s\alpha^{(x)},$$

so that, in a linear (with respect to the concentration $c = s/(2a)^2$) approximation, from Eq. (41) follows

$$\sigma_{xe} = \sigma_1(1 + 4\pi c\alpha^{(x)}). \tag{42}$$

Expression (42) coincides with the respective formula in [10] (with due regard for the fact that $\alpha^{(x)}$ in [10] is two times greater in magnitude).

Within the terms of $\sim R^{10}$ inclusive, we find for the coefficient B_0 from Eqs. (32), (28), and (19)

$$\begin{aligned} B_0 = {} & \Lambda_{11} - c_2\Lambda_{11}\left(\frac{1}{3}\Lambda_{31} + \Lambda_{13}\right) \\ & - c_4\left(\frac{1}{7}\Lambda_{11}\Lambda_{71} + 3\Lambda_{13}\Lambda_{51} + 5\Lambda_{31}\Lambda_{15} + \Lambda_{11}\Lambda_{17}\right) \\ & + c_2^2\Lambda_{11}\left[\frac{1}{9}(\Lambda_{31})^2 + \frac{1}{3}\Lambda_{13}\Lambda_{31} + \frac{1}{3}\Lambda_{11}\Lambda_{33} + (\Lambda_{13})^2\right] + \dots . \end{aligned} \tag{43}$$

For circular inclusions, it follows from Eq. (43) in view of (38) that

$$B_0 = \left\{1 + \frac{1}{3}c_2^2R^8\delta^2 + \dots\right\}R^2\delta, \qquad \delta = \frac{1-h}{1+h}, \tag{44}$$

which is in agreement with [4].

The terms written out in Eq. (43) are the first terms of the virial expansion for the coefficient B_0 and, thereby, for the conductivity σ_{xe}. Note that it is a quantity of the order of R/a, where R is the maximum "radius" of inclusion, that is the expansion parameter in Eq. (43) rather than the concentration (the fraction of the area being taken up) of inclusions. One can expect, however, that, as in the case of circular inclusions [4], the approximation given by Eq. (43) describes adequately the conductivity of the system for all values of $R/a \lesssim 1$. As R approaches a, an ever larger number of terms is to be taken into account in the expansion given by Eq. (43). One is inclined to think that the respective series will be convergent if the expansions given by Eqs. (8) and (10) converge (at $|z| > R$).

In the case when $\langle \mathbf{E} \rangle$ is directed along the y axis, we derive for the effective conductivity σ_{ye}, similarly to Eq. (41),

$$\sigma_{ye} = \frac{\sigma_1(1 - \bar{B}_0\pi/4a^2)}{1 + \bar{B}_0\pi/4a^2} \tag{45}$$

with the coefficient $\bar{B}_0$ from Eq. (34). For a reciprocal system ($h \longrightarrow 1/h$), it follows from Eq. (37), in particular, that $\tilde{\bar{B}}_0 = -\bar{B}_0$. Therefore, a comparison between σ_{ye} and $\tilde{\sigma}_{xe}$ gives

$$\tilde{\sigma}_{xe}\sigma_{ye} = \sigma_1\sigma_2, \tag{46}$$

which is the reciprocity relation [11].

4.2. Thermal Electromotive Force

In the case of a weak thermoelectric coupling, the following general expression [12] is valid for the principal value α_{xe} of the effective tensor of thermal emf $\hat{\alpha}_e$ of a two-component medium with structural anisotropy:

$$\alpha_{xe} = \alpha_1 - (\alpha_1 - \alpha_2)\Psi^{(x)}, \tag{47}$$

where

$$\Psi^{(x)} = \frac{\sigma_2}{\sigma_{xe}}\frac{\langle \mathbf{E}^{(x)} \cdot \mathbf{G}^{(x)} \rangle^2}{\langle E_x^{(x)} \rangle \langle G_x^{(x)} \rangle}. \tag{48}$$

Here, α_i is the thermoelectric coefficient of the ith component, and $\langle \dots \rangle^{(2)}$ is the integral over the area of inclusion, divided by the area of elementary cell. In Eq. (48), $\mathbf{E}^{(x)} = -\nabla\varphi$ is the electric field intensity and $\mathbf{G}^{(x)} = -\nabla T$ is the temperature field "intensity," which are calculated in the absence of thermoelectric coupling at $\langle \mathbf{E} \rangle \parallel x$ and $\langle \mathbf{G} \rangle \parallel x$.

The problems on electrical conductivity and thermal conductivity in the absence of thermoelectric effects change one into the other in the case of permutations $\sigma \longleftrightarrow \kappa$, $\mathbf{E} \longleftrightarrow \mathbf{G}$, and $\mathbf{j} \longleftrightarrow \mathbf{q}$, where κ is the thermal conductivity and $\mathbf{q} = \kappa\mathbf{G}$ is the heat flux density. Therefore, the results given in Sections 2 and 3 and in Subsection 4.1 are extended to the problem on thermal conductivity by means of substitutions $\sigma_i \longrightarrow \kappa_i$ and $\sigma_{\alpha e} \longrightarrow \kappa_{\alpha e}$, so that, for example,

$$\kappa_{xe} = \frac{\kappa_1(1 - \bar{\bar{B}}_0\pi/4a^2)}{1 + \bar{\bar{B}}_0\pi/4a^2}. \tag{49}$$

Here, two bars mark the coefficient obtained from B_0 at $\sigma_i \longrightarrow \kappa_i$.

In order to calculate $\Psi^{(x)}$, we will use formula (A.5) from Appendix A. In doing this, note that the electric potential

$$\varphi(\mathbf{r}) = \operatorname{Re}\Phi(z)$$

with $\Phi(z)$ from Eq. (29) assumes, in view of Eq. (23), the form

$$\begin{aligned} \varphi(\mathbf{r}) = \beta\Bigg\{ & r\cos\theta + \sum_{n=0}^{\infty}\left[B_2\frac{(2n)!}{r^{2n+1}}\right. \\ & \left. - \sum_{m=0}^{\infty} B_{2m}\frac{(2n+2m)!}{(2n+1)!}c_{n+m+1}r^{2n+1}\right]\cos(2n+1)\theta\Bigg\}. \end{aligned} \tag{50}$$

The substitution of Eq. (50) and a similar expansion for $T(\mathbf{r})$ (in which B_{2n} should be replaced by $\bar{\bar{B}}_{2n}$) into formula (A.5) gives

$$\int\limits_s (\mathbf{EG})d\mathbf{r} = 2\pi\beta\bar{\bar{\beta}}(\bar{\bar{B}}_0 - B_0)\left(\frac{\sigma_2}{\sigma_1} - \frac{\kappa_2}{\kappa_1}\right)^{-1}. \tag{51}$$

We substitute into Eq. (48) expression (51), $\langle E_x^{(x)}\rangle = U_x/(2a)$ with U_x from Eq. (40), and an analogous (with substitutions of B_0 by $\bar{\bar{B}}_0$ and β by $\bar{\bar{\beta}}$) expression for $\langle G_x^{(x)}\rangle$ to derive

$$\Psi^{(x)} = \frac{\sigma_2}{\sigma_{xe}}\frac{2(\bar{\bar{B}}_0 - B_0)}{\sigma_2/\sigma_1 - \kappa_2/\kappa_1}\frac{\pi}{4a^2} \times\left(1 + B_0\frac{\pi}{4a^2}\right)^{-1}\left(1 + \bar{\bar{B}}_0\frac{\pi}{4a^2}\right)^{-1}. \tag{52}$$

We use Eqs. (41) and (49) to eliminate the coefficients B_0 and $\bar{\bar{B}}_0$ from Eq. (52) and finally derive

$$\Psi^{(x)} = \left(\frac{\kappa_1}{\sigma_1} - \frac{\kappa_{xe}}{\sigma_{xe}}\right)\left(\frac{\kappa_1}{\sigma_1} - \frac{\kappa_2}{\sigma_2}\right)^{-1}. \tag{53}$$

Expression (53) is a generalization of the respective isotropic formula (see [13]) to the case of structurally (geometrically) anisotropic media.

4.3. Partial Quadratic Characteristics

Directly related to the effective conductivity tensor $\hat{\sigma}_e$ are the partial quadratic characteristics of the electric field intensity [14],

$$\psi_i^{(\alpha)} \equiv \langle(\mathbf{e}^{(\alpha)})^2\rangle^{(i)} = \frac{\partial\sigma_{\alpha e}}{\partial\sigma_i}, \qquad \mathbf{e}^{(\alpha)}(\mathbf{r}) = \frac{\mathbf{E}^{(\alpha)}(\mathbf{r})}{|\langle\mathbf{E}^{(\alpha)}\rangle|}. \tag{54}$$

Here, $\sigma_{\alpha e}$ (where $\alpha = x, y, z$) denotes the principal values of the tensor $\hat{\sigma}_e$, and $\mathbf{E}^{(\alpha)}(\mathbf{r})$ is the electric field intensity in the medium; the index α implies that $\langle\mathbf{E}^{(\alpha)}\rangle$ is directed along the axis α; and $\langle\ldots\rangle^{(i)}$ is the integral over the area (volume, in the three-dimensional case) of the ith component, divided by the area (volume) of sample.

We will use formula (51) in order to calculate $\psi_2^{(x)}$. We assume in Eq. (51) that $\kappa_1 = \sigma_1$ and $\bar{\bar{\beta}} = \beta$ and then perform the limiting transition $\kappa_2 \longrightarrow \sigma_2$ to derive

$$\int\limits_s \mathbf{E}^2 dr = -2\pi\sigma_1\beta^2\frac{\partial B_0}{\partial\sigma_2}. \tag{55}$$

From Eq. (55), in view of $= U_x/(2a)$ (with U_x from Eq. (40)), we find

$$\langle(\mathbf{e}^{(x)})^2\rangle^{(2)} = -2\sigma_1\frac{\partial B_0}{\partial\sigma_2}\frac{\pi}{4a^2}\left(1 + B_0\frac{\pi}{4a^2}\right)^{-2}. \tag{56}$$

One can readily see that the right-hand part of Eq. (56) is a derivative of σ_{xe} from Eq. (41),

$$\langle(\mathbf{e}^{(x)})^2\rangle^{(2)} = \frac{\partial\sigma_{xe}}{\partial\sigma_2}, \tag{57}$$

which coincides with Eq. (54) at $\alpha = x$ and $i = 2$.

4.4. Hall Coefficient

The effective Hall coefficient R_e in a low magnetic field $\mathbf{H}$ is expressed in terms of the Hall component σ_{ae} of the effective conductivity tensor $\hat{\sigma}_e$ as follows:

$$R_e = \frac{1}{H}\frac{\sigma_{ae}}{\sigma_{xe}\sigma_{ye}}. \tag{58}$$

Here, it is taken into account that the system being treated is, generally speaking, structurally anisotropic. In turn, σ_{ae} may be expressed in terms of the Hall components of individual components of σ_{ai} [14],

$$\sigma_{ae} = \sigma_{a2} + (\sigma_{a1} - \sigma_{a2})\varphi_a, \tag{59}$$

where

$$\varphi_a = 1 - \left\{\frac{\langle[\mathbf{E}^{(x)} \times \mathbf{E}^{(y)}]_z\rangle^{(2)}}{\langle E_x^{(x)}\rangle\langle E_y^{(y)}\rangle}\right\}. \tag{60}$$

Here, $\langle\ldots\rangle^{(2)}$ is the same as in Eq. (48).

We substitute into Eq. (B.12) $\varphi^{(x)} = \mathrm{Re}\,\Phi(z)$, $A^{(x)} = -\mathrm{Im}\,\Phi(z)$, and $\varphi^{(y)} = \mathrm{Re}\,\bar{\Phi}(z)$ with $\Phi(z)$ from Eq. (29) and $\bar{\Phi}(z)$ from Eq. (34) to derive

$$\int\limits_s [\mathbf{E}^{(x)} \times \mathbf{E}^{(y)}]_z d\mathbf{r} = 2\pi\beta\bar{\beta}\frac{B_0 + \bar{B}_0}{1 - h^2}. \tag{61}$$

We use relation (61) to derive from Eq. (60)

$$a = \frac{\sigma_{xe}\sigma_{ye} - \sigma_2^2}{\sigma_1^2 - \sigma_2^2}, \tag{62}$$

which is a generalization of the respective isotropic formula from [14] to the case of a system with structural anisotropy.

5. INCLUSION OF ELLIPTIC SHAPE

The problem on multipole polarizabilities of elliptic cylinder is solved in elliptic coordinates (μ, θ) [15],

$$x = z_0\cosh\mu\cos\theta, \quad y = z\sinh\mu\sin\theta. \tag{63}$$

The equation for the boundary of inclusion is preassigned by the equality $\mu = \mu_0$, so that $z_0 \cosh\mu_0 = a_0$ and $z_0 \sinh\mu_0 = b_0$ are the major and minor semiaxes of the ellipse, respectively. Hence,

$$z_0 = \sqrt{a_0^2 - b_0^2}, \quad \mu_0 = \frac{1}{2}\ln\frac{a_0 + b_0}{a_0 - b_0}. \qquad (64)$$

The ellipse matrix of polarizabilities is found in Appendix C. The following expression holds for its elements:

$$\Lambda^{(x)}_{2n+1,\,2m+1} = \frac{2}{2m+1}\left(\frac{z_0}{2}\right)^{2n+2m+2}$$
$$\times \sum_{k=n-m}^{n} (2n-2k+1) C^{k}_{2n+1} D^{2n-2k+1}_{2n+1} C^{m-n+k}_{2m+1}, \qquad (65)$$
$$m \le n.$$

The expression for $\Lambda^{(x)}_{2n+1,\,2m+1}$ at $m \ge n$ differs from Eq. (65) only in that the summation over k proceeds from zero to n. In Eq. (65), C^m_n is the binomial coefficient, and

$$D^{2n-2k+1}_{2m+1} = \frac{1}{2}\frac{(1-h)\sinh 2\xi}{\cosh\xi + h\sinh\xi}e^{\xi}, \qquad (66)$$
$$\xi = (2n-2k+1)\mu_0.$$

In particular, for $k = n$, we have, in view of Eq. (64),

$$D^1_{2n+1} = D^1_1 = \frac{a_0 b_0}{a_0 - b_0}\frac{1-h}{a_0 + hb_0}, \qquad (67)$$

so that

$$\Lambda^{(x)}_{11} = \frac{a_0 + b_0}{2}\frac{(1-h)a_0 b_0}{a_0 + hb_0}. \qquad (68)$$

The quantities $\Lambda^{(y)}_{2n+1,\,2m+1}$ may be found by the known values of $\Lambda^{(y)}_{2n+1,\,2m+1}$ using relation (14) written as

$$\Lambda^{(y)}_{2n+1,\,2m+1} = -\tilde{\Lambda}^{(x)}_{2n+1,\,2m+1}, \qquad (69)$$

where the tilde indicates the substitution $h \longrightarrow 1/h$.

APPENDIX A

In order to calculate the quantity entering the expression for the thermal emf,

$$\int_s \mathbf{E}\cdot\mathbf{G}d\mathbf{r},$$

where the integration is performed over the area s of the inclusion, we will treat the integral

$$J = \int_{r\le\rho} (\sigma_1\mathbf{qE} - \kappa_1\mathbf{jG})d\mathbf{r}, \qquad (A.1)$$

taken over the area of a circle of radius ρ ($R < \rho < a$). Note that the integrand in Eq. (A.1) is other than zero only inside the inclusion, so that

$$J = (\sigma_1\kappa_2 - \sigma_2\kappa_1)\int_s \mathbf{E}\cdot\mathbf{G}d\mathbf{r}. \qquad (A.2)$$

On the other hand, in view of the equations $\operatorname{div}\mathbf{j} = 0$ and $\operatorname{div}\mathbf{q} = 0$, we have

$$\sigma_1\mathbf{qE} - \kappa_1\mathbf{jG} = -\nabla(\sigma_1\varphi\mathbf{q} - \kappa_1 T\mathbf{j}), \qquad (A.3)$$

where it is taken that $\mathbf{E} = -\nabla\varphi$ and $\mathbf{G} = -\nabla T$. In view of Eq. (A.3), the integral in Eq. (A.1) may be transformed to a "surface" integral taken over the circumference of radius ρ that lies fully outside of the inclusion. As a result, from Eq. (A.1) follows

$$J = \sigma_1\kappa_1\int_0^{2\pi}\left(\varphi\frac{\partial T}{\partial r} - T\frac{\partial\varphi}{\partial r}\right)\bigg|_{r=\rho}\rho d\theta. \qquad (A.4)$$

We equate Eqs. (A.2) and (A.4) to find

$$\int_s \mathbf{E}\cdot\mathbf{G}d\mathbf{r} = \left(\frac{\kappa_2}{\kappa_1} - \frac{\sigma_2}{\sigma_1}\right)^{-1}$$
$$\times\int_0^{2\pi}\left(\varphi\frac{\partial T}{\partial r} - T\frac{\partial\varphi}{\partial r}\right)\bigg|_{r=\rho}\rho d\theta. \qquad (A.5)$$

For a solitary inclusion, we use the asymptotic forms given by Eq. (3) for the potential $\varphi(\mathbf{r})$ (and analogous asymptotic forms for $T(\mathbf{r})$) to derive from Eq. (A.5), in the limit $\rho \longrightarrow \infty$,

$$\int_s \mathbf{E}\cdot\mathbf{G}d\mathbf{r} = 4\pi\frac{\bar{\bar{\Lambda}}^{(x)} - \Lambda^{(x)}}{\bar{\bar{h}} - h}E_0^2, \qquad (A.6)$$

where $h = \sigma_2/\sigma_1$, $\bar{\bar{h}} = \kappa_2/\kappa_1$, and $\bar{\bar{\Lambda}}^{(x)} = \Lambda^{(x)}\bar{\bar{h}}$. In the limit $\bar{\bar{h}} \longrightarrow h$, from Eq. (A.6) follows

$$\int_s \mathbf{E}^2 d\mathbf{r} = 4\pi\frac{\partial\Lambda^{(x)}}{\partial h}E_0^2, \qquad (A.7)$$

where $\Lambda^{(x)}$ is the principal value of the dipolar polarizability tensor $\hat{\Lambda}$. In the case of arbitrary orientation of the inclusion, we have, instead of Eq. (A.7),

$$\int_s \mathbf{E}^2 d\mathbf{r} = 4\pi\left(\mathbf{E}_0\frac{\partial\hat{\Lambda}}{\partial h}\mathbf{E}_0\right). \qquad (A.8)$$

In the three-dimensional case, an analogous treatment using the asymptotic forms of the potential

$$\begin{gathered}\varphi(b r) = -\mathbf{E}_0\mathbf{r} + \frac{\mathbf{p}\mathbf{r}}{r^3} + \dots, \\ r \longrightarrow \infty, \quad \mathbf{p} = \hat{\Lambda}\mathbf{E}_0,\end{gathered} \tag{A.9}$$

also leads (for an integral over the volume of inclusion) to expressions (A.7) and (A.8).

APPENDIX B

In order to calculate the integral

$$\int_s [\mathbf{E}^{(x)} \times \mathbf{E}^{(y)}]_z d\mathbf{r},$$

entering the expression for the function φ_a in Eqs. (59) and (60), we will treat the quantity

$$I = \int_{r \le \rho} \{\sigma_1^2[\mathbf{E}^{(x)} \times \mathbf{E}^{(y)}]_z - [\mathbf{j}^{(x)} \times \mathbf{j}^{(y)}]_z\} d\mathbf{r}, \tag{B.1}$$

where the integration is performed over the circular area $r \le \rho$ ($R < \rho < a$). Because the integrand is other than zero only inside the inclusion, then

$$I = (\sigma_1^2 - \sigma_2^2)\int_s [\mathbf{E}^{(x)} \times \mathbf{E}^{(y)}]_z d\mathbf{r}. \tag{B.2}$$

Here, the integration is performed over the area s of the inclusion.

On the other hand, the integral in Eq. (B.1) may be transformed to a "surface" integral (contour integral, in the two-dimensional case being treated). In doing this, note that, by virtue of the equation $\operatorname{curl}\mathbf{E} = 0$, we have the equality

$$\mathbf{E}^{(x)} \times \mathbf{E}^{(y)} = -\operatorname{curl}\{\varphi^{(x)}\mathbf{E}^{(y)}\}, \tag{B.3}$$

where it is taken into account that $\mathbf{E}^{(x)} = -\nabla\varphi^{(x)}$. Therefore,

$$\begin{aligned}\frac{I_1}{\sigma_1^2} &\equiv \int_{r \le \rho} [\mathbf{E}^{(x)} \times \mathbf{E}^{(y)}]_z d\mathbf{r} \\ &= -\int_{r \le \rho} \operatorname{curl}\{\varphi^{(x)}\mathbf{E}^{(y)}\} d\mathbf{S},\end{aligned} \tag{B.4}$$

where the vector element $d\mathbf{S}$ of the area is directed along the z axis. The Stokes theorem is used to transform the integral over the surface of a circle of radius ρ into the integral over its contour,

$$\begin{aligned}\frac{I_1}{\sigma_1^2} &= -\oint \{\varphi^{(x)}\mathbf{E}^{(y)}\}\big|_{r = \rho} d\mathbf{l} \\ &= -\int_0^{2\pi} \{\varphi^{(x)}E_\tau^{(y)}\}\big|_{r = \rho}\rho d\theta.\end{aligned} \tag{B.5}$$

For the tangential component of intensity $\mathbf{E}^{(y)} = -\nabla\varphi^{(y)}$, we have

$$E_\tau^{(y)} = -r^{-1}(\partial\varphi^{(y)}/\partial\theta)$$

at $r = \rho$, so that from Eq. (B.5) follows

$$I_1 = \sigma_1^2 \int_0^{2\pi} \left\{\varphi^{(x)}\frac{1}{r}\frac{\partial\varphi^{(y)}}{\partial\theta}\right\}\Bigg|_{r = \rho} \rho d\theta. \tag{B.6}$$

We will further introduce the vector potential according to

$$\mathbf{j} = \sigma_1 \operatorname{curl}\mathbf{A}, \tag{B.7}$$

where $A_x = A_y = 0$ and $A_z = A$, so that

$$j_x = \sigma_1\frac{\partial A}{\partial y}, \quad j_y = -\sigma_1\frac{\partial A}{\partial x}. \tag{B.8}$$

Then, for the quantity

$$[\mathbf{j}^{(x)} \times \mathbf{j}^{(y)}]_z = \sigma_1\left\{\frac{\partial A^{(x)}}{\partial x}j_x^{(y)} + \frac{\partial A^{(x)}}{\partial y}j_y^{(y)}\right\} \tag{B.9}$$

in view of the equation $\operatorname{div}\mathbf{j} = 0$, we have

$$[\mathbf{j}^{(x)} \times \mathbf{j}^{(y)}]_z = \sigma_1\nabla\{A^{(x)}\mathbf{j}^{(y)}\}. \tag{B.10}$$

Therefore,

$$\begin{aligned}I_2 \equiv -\int_{r \le \rho} [\mathbf{j}^{(x)} \times \mathbf{j}^{(y)}]_z d\mathbf{r} &= -\sigma_1\int_0^{2\pi} \{A^{(x)}j_N^{(y)}\}\big|_{r = \rho}\rho d\theta \\ &= \sigma_1^2\int_0^{2\pi}\left\{A^{(x)}\frac{\partial\varphi^{(y)}}{\partial r}\right\}\Bigg|_{r = \rho}\rho d\theta.\end{aligned} \tag{B.11}$$

Here, j_N is the normal (to the $r = \rho$ contour) component of $\mathbf{j}$.

From Eqs. (B.6) and (B.11), we find the quantity $I = I_1 + I_2$, whose comparison with Eq. (B.2) ultimately produces

$$\begin{aligned}\int_s [\mathbf{E}^{(x)} \times \mathbf{E}^{(y)}]_z d\mathbf{r} &= \frac{1}{1 - h^2} \\ &\times \int_0^{2\pi}\left\{\varphi^{(x)}\frac{1}{r}\frac{\partial\varphi^{(y)}}{\partial\theta} + A^{(x)}\frac{\partial\varphi^{y}}{\partial r}\right\}\Bigg|_{r = \rho}\rho d\theta,\end{aligned} \tag{B.12}$$

where $h = \sigma_2/\sigma_1$.

APPENDIX C

In solving the problem on multipole polarizabilities of an elliptic cylinder, it is convenient to use a complex representation. In so doing, the coupling of coordinates given by Eq. (63) takes the form

$$z = x + iy = z_0 \cosh w, \quad w = \mu + i\theta \qquad (C.1)$$

with z_0 from Eq. (64). For the complex potential of the external field, we have

$$\Phi_0(z) = z^{2n+1} = z_0^{2n+1}(\cosh w)^{2n+1}. \qquad (C.2)$$

According to [8],

$$(\cosh w)^{2n+1} = \frac{1}{2^{2n}} \sum_{k=0}^{n} C_{2n+1}^{k} \cosh[(2n-2k+1)w]. \qquad (C.3)$$

Therefore, the complex potentials outside and inside the ellipse are sought in the form

$$\Phi^{(e)}(z) = 2\left(\frac{z_0}{2}\right)^{2n+1} \sum_{k=0}^{n} C_{2n+1}^{k} \{\cosh[2n-2k+1)w] + D_{2n+1}^{2n-2k+1} e^{-(2n-2k+1)w}\}, \qquad (C.4)$$

$$\Phi^{(i)}(z) = 2\left(\frac{z_0}{2}\right)^{2n+1} \times \sum_{k=0}^{n} C_{2n+1}^{k} B_{2n+1}^{2n-2k+1} \cosh[(2n-2k+1)w] \qquad (C.5)$$

with the real coefficients $D_{2n+1}^{2n-2k+1}$ and $B_{2n+1}^{2n-2k+1}$. We separate the real parts in Eqs. (C.4) and (C.5) and satisfy the boundary conditions

$$\varphi^{(e)} = \varphi^{(i)}, \quad \frac{\partial \varphi^{(e)}}{\partial \mu} = h \frac{\partial \varphi^{(i)}}{\partial \mu}, \quad \mu = \mu_0 \qquad (C.6)$$

to find the coefficients $D_{2n+1}^{2n-2k+1}$ (see formula (66)) and $B_{2n+1}^{2n-2k+1}$,

$$B_{2n+1}^{2n-2k+1} = \frac{e^{\xi}}{\cosh\xi + h\sinh\xi}, \quad \xi = (2n-2k+1)\mu_0. \qquad (C.7)$$

Then, we use the expansion

$$(1-\sqrt{1-x})^{n} = n\left(\frac{x}{2}\right)^{n} \sum_{k=0}^{\infty} \frac{1}{2k+n} C_{2k+n}^{k} \left(\frac{x}{4}\right)^{k} \qquad (C.8)$$

to find

$$e^{-(2n-2k+1)w} = \left(\frac{z-\sqrt{z^2-z_0^2}}{z_0}\right)^{2n-2k+1} = (2n-2k+1) \times \sum_{m=n-k}^{\infty} \frac{1}{2m+1} C_{2m+1}^{m-n+k} \left(\frac{z_0}{2z}\right)^{2m+1}. \qquad (C.9)$$

In view of Eqs. (C.2), (C.3), and (C.9), expression (C.4) takes the form

$$\Phi^{(e)}(z) = z^{2n+1} + 2\left(\frac{z_0}{2}\right)^{2n+1} \times \sum_{k=0}^{n} C_{2n+1}^{k} D_{2n+1}^{2n-2k+1}(2n-2k+1) \times \sum_{m=n-k}^{\infty} \frac{1}{2m+1} C_{2m+1}^{m-n+k} \left(\frac{z_0}{2z}\right)^{2m+1}. \qquad (C.10)$$

After the change of the order of summation in (C.10) and comparison with the general formula (10), we derive expressions (65) and (66) for $\Lambda_{2n+1,\,2m+1}^{(x)}$.

The even–even multipole polarizabilities $\Lambda_{2n,\,2m}^{(x)}$ are sought analogously. As a result,

$$\Lambda_{2n,\,2m}^{(x)} = \frac{2}{m}\left(\frac{z_0}{2}\right)^{2n+2m} \times \sum_{k=n-m}^{n-1} (n-k) C_{2n}^{k} D_{2n}^{2n-2k} C_{2m}^{m-n+k}, \quad 1 \le m \le n, \qquad (C.11)$$

where

$$D_{2n}^{2n-2k} = \frac{1}{2} \frac{(1-h)\sinh 2\xi}{\cosh\xi + h\sinh\xi} e^{\xi}, \quad \xi = 2(n-k)\mu_0. \qquad (C.12)$$

The expression for $\Lambda_{2n,\,2m}^{(x)}$ at $m \ge n$ differs from (C.11) in that the summation over k proceeds from zero to $n-1$. The polarizabilities $\Lambda_{2n,\,2m}^{(y)}$ are found from formulas of the type of (69).

REFERENCES

1. Yu. P. Emets, *Electrical Properties of Composites with Regular Structure* (Naukova Dumka, Kiev, 1986).
2. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **79**, 1561 (1980) [Sov. Phys. JETP **52**, 787 (1980)].
3. B. Ya. Balagurov, Zh. Tekh. Fiz. **53**, 428 (1983) [Sov. Phys. Tech. Phys. **28**, 269 (1983)].

4. B. Ya. Balagurov and V. A. Kashin, Zh. Éksp. Teor. Fiz. **117**, 978 (2000) [JETP **90**, 850 (2000)].
5. B. Ya. Balagurov and V. A. Kashin, Zh. Tekh. Fiz. **71**, 106 (2001) [Tech. Phys. **46**, 101 (2001)].
6. Lord Rayleigh, Philos. Mag. **34**, 481 (1892).
7. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
8. I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products* (Fizmatgiz, Moscow, 1962; Academic, New York, 1980).
9. A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 110 (1970) [Sov. Phys. JETP **32**, 63 (1970)].
10. B. Ya. Balagurov, Zh. Tekh. Fiz. **52**, 850 (1982) [Sov. Phys. Tech. Phys. **27**, 544 (1982)].
11. J. B. Keller, J. Math. Phys. **5**, 548 (1964).
12. B. Ya. Balagurov, Fiz. Tekh. Poluprovodn. (Leningrad) **21**, 1978 (1987) [Sov. Phys. Semicond. **21**, 1198 (1987)].
13. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **85**, 568 (1983) [Sov. Phys. JETP **58**, 331 (1983)].
14. B. Ya. Balagurov, Zh. Éksp. Teor. Fiz. **93**, 1888 (1987) [Sov. Phys. JETP **66**, 1079 (1987)].
15. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics* (McGraw-Hill, New York, 1953; Inostrannaya Literatura, Moscow, 1960), Vol. 2.

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 596–603.
Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 678–687.
Original Russian Text Copyright © 2001 by Gladkiĭ, Kirikov, Volk, Ivleva.

The Kinetic Characteristics of Polarization of Relaxor Ferroelectrics

V. V. Gladkiĭ[a], *, V. A. Kirikov[a], T. R. Volk[a], and L. I. Ivleva[b]

[a]*Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninskiĭ pr. 59, Moscow, 117333 Russia*
[b]*Scientific Center for Laser Materials and Technologies, Institute of General Physics, Russian Academy of Sciences, ul. Vavilova 38, Moscow, 117942 Russia*
**e-mail: glad@ns.crys.ras.ru*

Received February 8, 2001

Abstract—SBN crystals doped with rare-earth metal ions were studied to show that relaxor ferroelectrics had pronounced anomalies, which manifested themselves by the noncoincidence of the trajectories of the first several cycles of dielectric hysteresis loops, the absence of an unambiguous coercive field, and other special features of the kinetics of polarization. These anomalies were related to structural disorder of the crystals and a random internal electric field distribution and could only be observed in constant and quasi-static electric fields. A phenomenological analysis of the thermal activation stages of polarization relaxation was performed. The spectra of the energy distribution of potential barriers were reproduced in the approximation of their independence. Electric conductivity was shown to play an important role in the formation of giant barriers. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Relaxor ferroelectrics (relaxors) are solid oxide solutions with perovskite [1–3] or tungsten bronze [2–4] structures. The best studied representatives of the latter are $Sr_xBa_{1-x}Nb_2O_6$ (SBN) crystals with Sr concentrations in the range $0.75 \geq x \geq 0.25$ [4]. A characteristic feature of relaxors is substantial structural disordering caused by various reasons. For instance, Ba and Sr atoms in SBN only fill 5/6 of basis sites and are statistically distributed over two available structural channels [5]. As distinguished from usual uniform ferroelectrics, the phase transition to the polarized state and physical property anomalies in relaxors are spread over a wide temperature region (the Curie region). For instance, the permittivity ε function has a flat maximum and a noticeable dispersion at some T_m temperature [1–3].

Relaxors are characterized by high dielectric, piezoelectric, pyroelectric, electrooptical, and nonlinear optical characteristic values with substantial nonlinearity and weak temperature dependences because of phase transition smearing. For this reason, relaxors are promising materials for use in optics and piezoelectric devices. An attractive feature of these materials is also the possibility of modifying their properties by changing the chemical composition. For instance, increasing the concentration of Sr in SBN [2–4] and doping SBN with rare-earth metals [6–8] substantially decrease the T_m temperature, increase several parameters important for practical applications, and make characteristic relaxor properties more pronounced. SBN crystals have certain practical applications, for instance, in holography [9], because, when doped with Ce, Cr, and Co, they acquire a high sensitivity to recording and high coupling constants (“amplification factors”) of light waves in multiwave interactions. SBN crystals with certain rare-earth metal admixtures also offer promise for use in piezoelectric devices [6, 8]. One of the new possibilities of applying SBN is optical frequency conversion under quasi-phase matching conditions on regular domain structures [10–12].

The weak point common to all relaxors including SBN is irreproducibility of properties and their deterioration under external actions, in the first place, electric field actions. Detailed studies of polarization processes are therefore of great importance. The dielectric properties of relaxors were studied in alternating electric fields [1–3]. In this work, we report polarization measurements in constant and slowly varying (quasi-static) fields, which allows the contribution of long-lived metastable states characteristic of all nonuniform structures [13] to be taken into account. Our preliminary data on SBN of one composition showed [7, 14] that polarization processes in relaxors had very special properties and could not be described in terms of the traditional concepts of their occurrence in usual uniform ferroelectrics. The purpose of this work was to perform a comparative experimental analysis of the kinetics of polarization of crystals with different relaxor properties and a model uniform ferroelectric, triglycine sulfate (TGS).

1063-7761/01/9303-0596$21.00 © 2001 MAIK “Nauka/Interperiodica”

2. CRYSTALS AND EXPERIMENTAL PROCEDURE

The relation between the observed polarization and relaxor characteristics was studied for two SBN single crystals with 0.61 at. % Sr doped by various rare-earth metal admixtures, which allowed us to vary the degree of phase transition smearing and electric conductivity. The basis composition of SBN is congruent and possesses the best optical quality compared with other SBN crystals [4]. The crystals were grown using a modified Stepanov procedure at the Scientific Center for Laser Materials and Technologies, Institute of General Physics (Russian Academy of Sciences) [15]. One sample, SBN:(La + Ce), was doped with 1 wt % La_2O_3 and 0.1 wt % Ce_2O_3 in the molten state (the concentrations of the admixtures in the crystal were 0.44 at. % La and 0.023 at. % Ce according to measurements on a Comebax microanalyzer). The other sample, SBN:Nd, was doped with 0.5 wt % Nd_2O_3 (the concentration of the admixture in the crystals was roughly estimated at 0.7–0.8 at. % Nd).

The type of doping was selected from considerations of the possibility to control the phase transition temperature and the degree of its smearing upon the introduction of rare-earth admixtures [8]. The SBN:(La + Ce) crystal has the most smeared ε maximum at a comparatively low temperature T_m = (310–314) K [7, 8]; this crystal is a promising material for dynamic holography [16] because of its high electrooptical and light amplification coefficients. For this reason, expected anomalous polarization behavior near T_m was easier to study by the precision electrometric recording technique that we used thanks to a comparatively high electric resistance, which is known to increase as temperature decreases. The T_m temperature (340 K) is higher and the electric resistance lower for the SBN:Nd crystal, which makes it possible to estimate the influence of free charge carriers on polarization. Our interest in this composition stems from the observation that it exhibits the self-doubling effect of the generation frequency excited on Nd^{3+} ions in ferroelectric microdomains [12].

The samples were polished $2.5 \times 3 \times 0.7$ mm^3 plates of the polar z cut of crystals. The large faces were coated by a silver paste. The temperature of the samples was maintained constant to 0.03 K with the use of a cryostat.

Crystal polarization P was determined by precision compensation electrometric measurements. The key element of the scheme for measurements was an equal-arm bridge. The sample, reference capacitance C, and the low-resistance sources of constant voltage, V and v (applied to the sample and capacitor, respectively) were placed in the first, second, third, and fourth bridge arms. A V7-29 electrometer in the bridge diagonal was used as a null indicator. When the bridge is balanced, the voltage drop across the electrometer is zero, and the electric charge Q density on the electrodes of a sample of area S at time t is

$$\sigma(t) = \frac{Q(t)}{S} = \frac{Cv(t)}{S} = P(t) + \frac{E}{\rho}t, \qquad (1)$$

where $E = V/d$, d is the sample thickness, and ρ is the specific electric resistance. If ρ is large and the second term in (1) can be ignored, $\sigma(t)$ charge variations in time are only related to changes in $P(t)$. This condition is met in good dielectrics with high ρ values, and corrections for electric conductivity effects are virtually unnecessary.

The highest voltage and charge sensitivities of the bridge were 20 μV and 2×10^{-9} μC (at C = 10 pF), respectively. Voltage compensation in the bridge diagonal was performed with the use of a computer, which controlled peripheral devices. The $v(t)$ dependence of compensation voltage was displayed on a monitor in the real-time mode. A detailed description of the unit can be found in [13].

Polarization P was recorded using the following three measurement schemes: during continuously cooling and heating crystals in field $E = 0$ (pyroelectric effect), in quasi-static field E at T = const (dielectric hysteresis), and in the presence or absence of field E = const at T = const (polarization relaxation). In pyroelectric effect measurements, the rate of temperature variations was 0.3 deg/min. In dielectric hysteresis measurements, voltage V, which was varied from +300 to –300 V, was supplied from a B5-50 source controlled by a program block. Several repolarization cycles were recorded in voltage and time intervals multiple to 1 V and 1 s, respectively; the largest number of steps was 1200. Studies of the kinetics of polarization and depolarization were performed by measuring P changes in intervals of 0.25 to 1 min.

Electric resistance R was estimated at long times, when P relaxation virtually ceased, and the time dependence of charge Q was a straight line, $dQ/dt = V/R$. All measurements were performed for nonpolarized crystals; samples were heated above T_m and then cooled to the required temperature in the field $E = 0$.

3.RESULTS AND DISCUSSION

3.1. Pyroelectric effect and dielectric hysteresis in relaxor ferroelectrics. The temperature dependences of permittivity ε in an alternating field with a 1-kHz frequency are shown in Fig. 1. For SBN:(La + Ce), the ε maximum is lower and the phase transition in more smeared than for SBN:Nd. This observation and a noticeable frequency dispersion of temperature T_m corresponding to the e maximum in SBN:(La + Ce) [7, 8] are evidence that this composition has more pronounced relaxor properties. The temperature dependences of ΔP changes in nonpolarized SBN:(La + Ce) and SBN:Nd crystals are shown in Fig. 2. Numbers at curve points correspond to the sequence of temperature

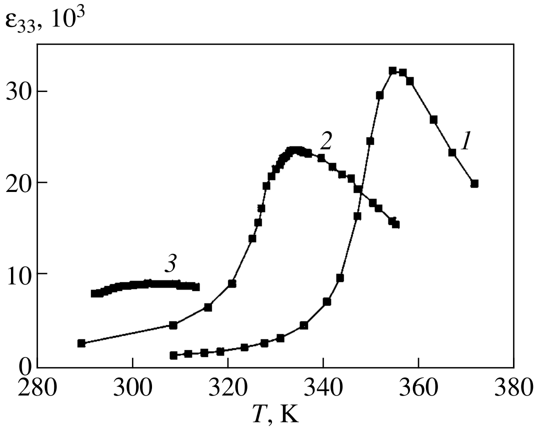


Fig. 1. Temperature dependences of permittivity ε_{33} of SBN crystals: (*1*) undoped, (*2*) doped by Nd, and (*3*) doped by La and Ce; measurements at 1 kHz.

and polarization changes with time. The $\Delta P(T)$ dependences in Fig. 2 correlate with the ε(T) dependences in Fig. 1. The first relaxor exhibits a smeared ε maximum and smooth ΔP variations at $T < T_m$, whereas a more pronounced ε maximum at T_m in the second relaxor corresponds with comparatively sharp ΔP changes. It is easy to verify that ΔP has a parabolic temperature dependence, $\Delta P = 0.18(\theta - T)^{1/2}$ (inset in Fig. 2b). The $\theta = 317$ K temperature is somewhat lower than T_m, which is likely to be related to the well-known low-frequency dispersion of ε [2].

The pyroelectric effect observed in nonpolarized crystals (Fig. 2) is, in the first place, evidence of the unipolar character of the samples, that is, of unequal volumes of regions with mutually opposite spontaneous polarization P_s directions. The stronger relaxor, SBN:(La + Ce), is much more unipolar (Fig. 2a). The $\Delta P(T)$ dependences recorded during cooling and heating do not coincide, and the temperature hysteresis magnitude depends on the rate of temperature variations. This is illustrated by relaxation of ΔP from point *4* to *5* at some fixed temperature in Fig. 2b (inset). Relaxation of ΔP is evidence of the formation of long-lived metastable states in the crystals after cooling. More details will be given below.

The quasi-static dielectric hysteresis loops of the relaxors are shown in Fig. 3. Solid circles correspond to the onset of repolarization. Numbers at the curves denote the sequence of P changes. As distinguished from the usual uniform ferroelectrics, the first loop cycles are open and noncoinciding curves with a decreasing P amplitude. After several cycles, the P amplitude virtually ceases to decrease, and all subsequent P trajectories coincide; that is, they become reproducible, and the loops acquire the familiar form. These low-frequency repolarization anomalies are more expressed in the first crystal, which, as mentioned, has other more pronounced relaxor properties.

The unusual shape of the loops is a direct substantiation of the earlier suggestions concerning the special features of the polar structure of relaxors [3] and can be phenomenologically explained as follows. Disordering of certain ions over various positions in the structure (for instance, disordering of Ba and Sr over two cation sites in SBN [4, 5]) should result in their concentration gradients, local internal electric fields E_i, and, as a consequence, a decrease in local symmetry. Local free energy should be an asymmetric two-minimum function of polarization in such a system,

$$F = -\alpha P^2 + \beta P^4 - (E_i + E)P, \qquad (2)$$

where E is the external field [2, 3]. Free energy F for $E_i + E < 0$ is schematically shown in the inset in Fig. 3. Field E_i, the depth of the minima, and the potential barriers between them are random values distributed over the crystal volume in a wide range of values. At $E = 0$, various crystal regions are in stable or metastable states corresponding to deep and shallow F minima, respectively, both with $P > 0$ and with $P < 0$. In an alternating field of a certain amplitude E, transitions of some crystal regions to the stable state may be observed, because fields lower barriers and accelerate relaxation. The reverse process is virtually impossible until $|E| \le |E_i|$. These regions therefore do not participate in further repolarization, and the P amplitude decreases. The difference of the P values at the beginning and at the end of the cycle of E variations can serve as a measure of the relative volume of the regions that do not participate in repolarization (are “quenched”). For instance, the relative volume of quenched regions in Fig. 3a approximately equals 40%. Qualitatively similar results were obtained for SBN in pulsed fields [8]. When temperature or field E decreases, all potential barriers become higher and the amplitude of P variations decreases [7]. Complete reproducible repolarization of the whole crystal volume is only possible in field E that exceeds the highest E_i field in the sample; it appears that the highest E_i value is fairly large.

3.2. Polarization relaxation in relaxor ferroelectrics. More complete information about the structure of the barriers can be obtained by analyzing polarization P relaxation in various constant fields E. Recall how such relaxation occurs in a usual uniform ferroelectric. In a uniform ferroelectric, $E_i = 0$ everywhere, F in (2) is a symmetrical function of P at $E = 0$, spontaneous polarization $P_s = (\alpha/2\beta)^{1/2}$, and the coercive field $E_c = (2\alpha/3)[(2\alpha/3)/\beta]^{1/2}$ coincides with the half-width of the loop and is well defined. At $E > E_c$, fast avalanche-like polarization occurs, and at arbitrary $E < E_c$, polarization is a slow thermally activated process without an initial P jump. Under both conditions, equilibrium polarization equals P_s, and the rate of relaxation increases as E grows [2]. By way of example, relaxation P of a TGS crystal in fields lower than E_c [13] is shown in Fig. 4a.

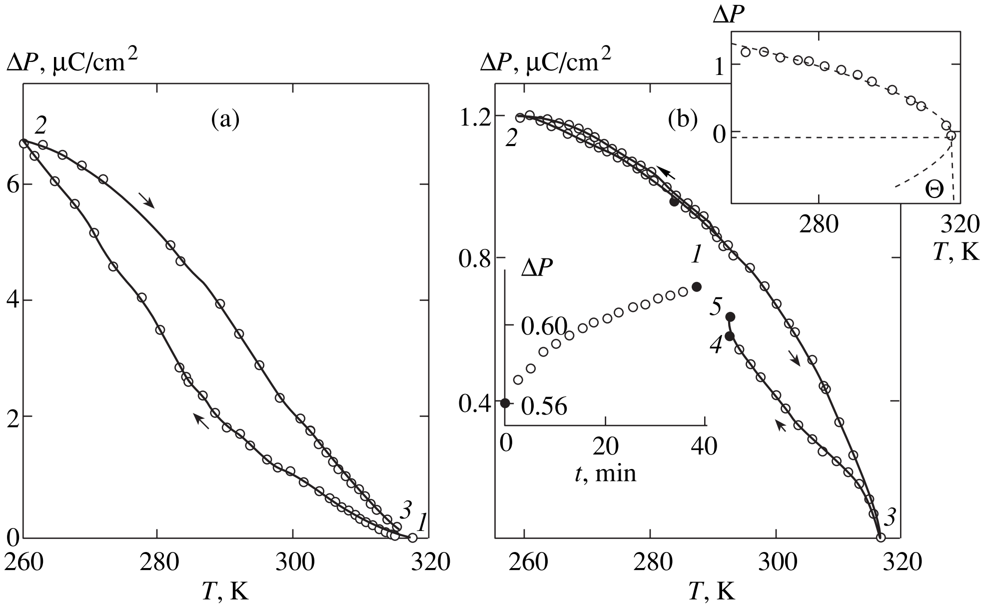


Fig. 2. Polarization ΔP measurements during cooling and heating (a) SBN : (La + Ce) and (b) SBN : Nd crystals in field $E = 0$. In the insets: relaxation of ΔP from point *4* to *5* and the $\Delta P(T)$ parabola.

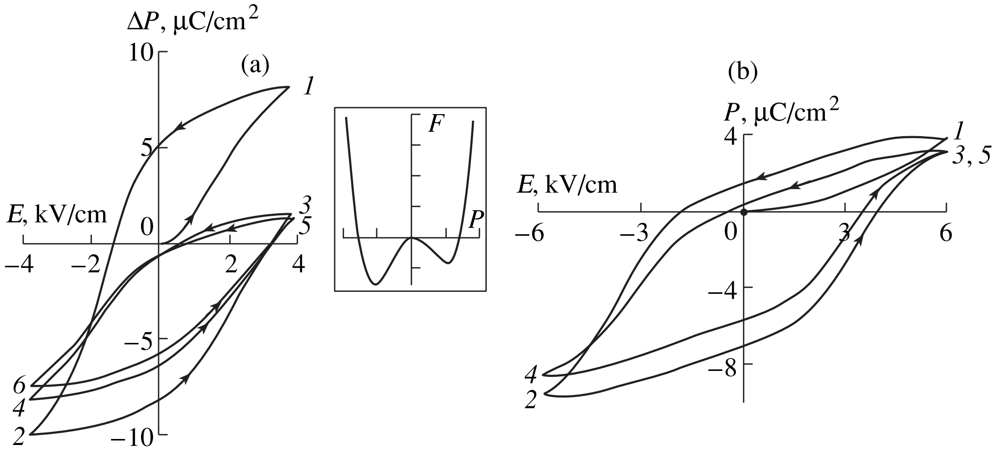


Fig. 3. Dielectric hysteresis loops of (a) SBN : (La + Ce) and (b) SBN : Nd; curve numbers correspond to the sequence of polarization P changes. In the inset: schematic drawing of local free energy function F versus P.

Polarization relaxation in the studied SBN crystals has basically other characteristics. Switching on field E, both larger and smaller than the loop half-width (formally, larger and smaller than E_c), causes P to change first jumpwise and then in a thermally activated manner (Fig. 4b). Jumps observed in a wide range of E values are evidence of the absence of an unambiguous coercive field E_c, which takes on different values at different crystal points. The P jump, naturally, increases at higher E fields because a larger crystal volume is polarized. A phenomenological analysis of the thermally activated relaxation stages shows that the concept of a certain equilibrium polarization P_e is also inapplicable to relaxors. Indeed, a certain limiting P_e value corresponds to each field E. This value increases as E grows, when crystal regions that remained "quenched" in lower fields become involved in the polarization process (Fig. 4b).

An analysis of the thermally activated relaxation stages was performed as in [17], in the approximation of independent relaxation centers whose contributions to polarization can be considered additive. Dimensionless polarization can then be written as

$$p(t) = \frac{P_e - P(t)}{P_e - P_0} = \int_0^{\infty} f(\tau)\exp\left(-\frac{t}{\tau}\right)d\tau. \quad (3)$$

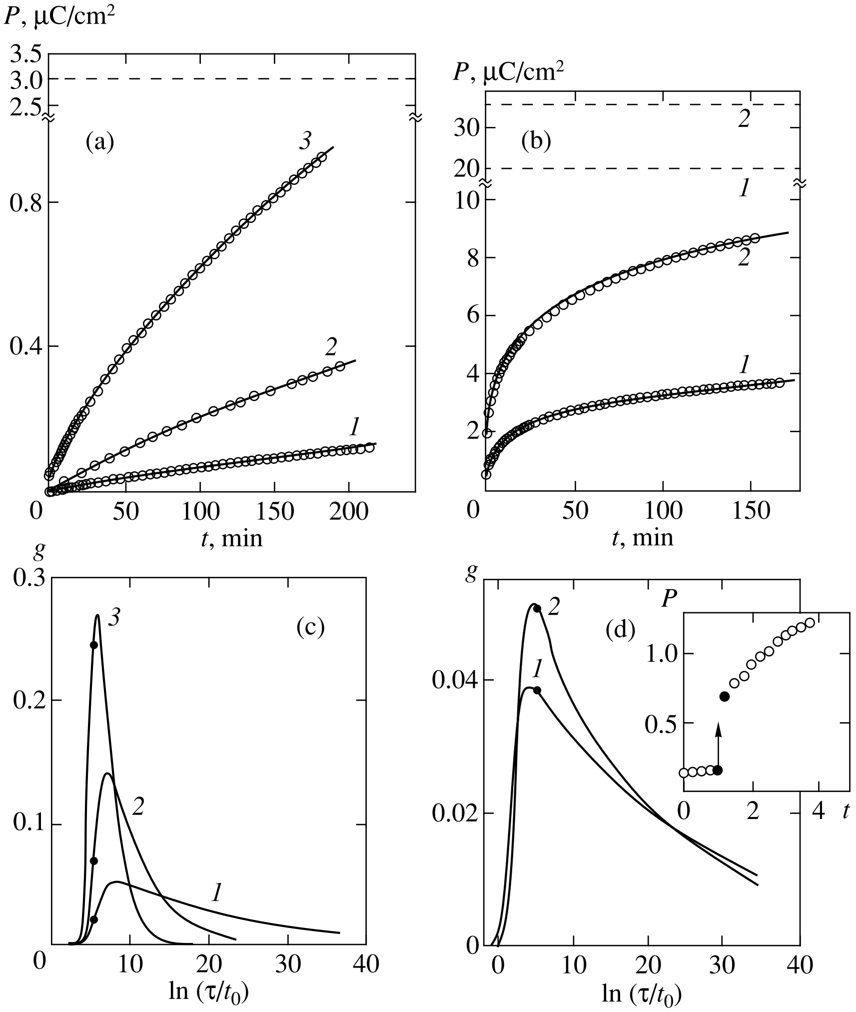


Fig. 4. (a, b) Polarization P relaxation and (c, d) $g(\ln\tau)$ spectra for (a, c) TGS and (b, d) SBN:(La + Ce): (a, c) field E = (*1*) 5.6, (*2*) 16, and (*3*) 25 V/cm, T = 293 K; (b, d) field E = (*1*) 300 and (*2*) 600 V/cm, T = 274 K. Solid curves are calculation results, experimental data are given by circles; dashed lines are equilibrium P_e values, t_0 = 1 min. In the inset: beginning of SBN relaxation in field E (P jump is marked by an arrow).

Here, P_0 is the initial polarization, P_e is the equilibrium polarization, $P(t)$ is the polarization at time t, and $f(\tau)$ is the normalized relaxation time τ distribution function. The $\tau^2 f(\tau)$ and $p(t)$ functions are related to each other by the Laplace integral transformation.

Our experimental data are described by the power time law

$$p(t) = 1/(1+t/a)^n. \qquad (4)$$

The $P(t)$ measurement results were approximated by (3) and (4) with three free parameters, P, a, and n, using the method of least squares and a standard program. The calculation results and experimental data are shown in Figs. 4a and 4b by solid curves and open circles, respectively. The deviation of the experimental values from the curves does not exceed 0.5%. Importantly, the larger the time interval of relaxation recording, the smaller the errors of determining the P_e, a, and n parameters [17]. The parameter values for TGS and SBN:(La + Ce) are listed in Table 1.

Empirical law (4) is likely to be universal for nonuniform systems of various types. Many of the nonexponential dependences observed earlier are its particular cases. For instance, $p(t) \sim 1/t^n$ at $t \gg a$, $p(t) \sim 1 - (n/a)t$ at $t \ll a$ [18], and $p(t) \sim 1 - n\ln(1 + t/a)$ at $n \ll 1$ [19]. Relaxation $p(t)$ in mixed $K_{1-x}Li_xTaO_3$ crystals in the vitreous state also obeys law (4) [20]. Compared with the long-known Kohlrausch law

$$p(t) \sim \exp\left(-\frac{t}{\tau}\right)^{\beta} \quad (\beta < 1),$$

which is also valid for our crystals at long times, law (4) has two advantages. First, it correctly describes relaxation even at short times. Secondly, a simple $f(\tau)$ function corresponds to this law no matter what the a and n parameter values.

Table 1. Relaxation and distribution spectrum $g(\tau)$ parameters for polarization of TGS and SBN:(La + Ce) crystals

Crystal	T, K	E, V/cm	P_e, μC/cm^2	a, min	n	τ_m, min	ΔU, eV	S
TGS	293	5.6	3.0	242.2	0.063	3800 ± 270	0.14	0.013
		16		269.4	0.220	1225 ± 85	0.09	0.027
		25		223.6	0.589	380 ± 30	0.06	0.168
SBN	274	300	24.5 ± 0.016	3.724 ± 0.012	0.045 ± 0.001	82.2 ± 0.3	0.15	0.14
		400	26.97 ± 0.019	4.35 ± 0.013	0.038 ± 0.001	114.5 ± 0.4	0.16	0.11
		600	38.41 ± 0.013	4.799 ± 0.07	0.058 ± 0.001	82.74 ± 0.15	0.14	0.16

Potential barrier U for a relaxation center is related to time τ by the Arrhenius equation $U = kT\ln(\tau/\tau_0)$, where τ_0 is the kinetic coefficient. For this reason, the $g(\ln\tau) = \tau f(\tau)$ dimensionless function can conveniently be used instead of $f(\tau)$; this function describes the distribution of $\ln\tau$ or the energy U distribution of barriers. For law (4), this function has the form [21]

$$g = \frac{1}{\Gamma(n)}\left(\frac{a}{\tau}\right)^n \exp\left(-\frac{a}{\tau}\right), \qquad (5)$$

where $\Gamma(n)$ is the gamma function.

The $g(\ln\tau)$ function has a maximum at $\tau_m = a/n$. If relaxation is slow, $n \ll 1$, the whole spectrum shifts to larger U and τ values, its width increases, and its maximum becomes lower. The width of the spectrum can be measured by the difference

$$\Delta(\ln\tau) = \ln\tau_2 - \ln\tau_1$$

or

$$\Delta U = kT\ln(\tau_2/\tau_1),$$

where τ_2 and τ_1 are the relaxation times τ corresponding to the inflection points of the $g(\ln\tau)$ curve, at which $d^2g/d(\ln\tau)^2 = 0$. Using (5) then gives

$$\Delta U = 2kT\ln\frac{(2n+1)+(4n+1)^{1/2}}{2n}$$

or

$$\Delta U = -2kT\ln n \quad \text{for} \quad n \ll 1.$$

The solid circles in Fig. 4 correspond to τ values equal to relaxation recording times $t_{\max}$. The spectra at $\tau > t_{\max}$ were obtained by extrapolating the experimental data to long times. For instance, for $\ln(\tau/t_0) = 10$ ($t_0 =$ 1 min), $\tau = 15$ days, and for $\ln(\tau/t_0) = 30$, $\tau = 10$ years. Clearly, the fraction of experimentally recorded relaxation processes, which equals area S bounded by the $g(\ln\tau)$ curve up to the $\ln(t_{\max}/t_0)$ point, is as a rule fairly small. The spectral parameters and areas S are listed in Table 1.

The $g(\ln\tau)$ spectrum of TGS is obviously different from that of SBN relaxors. The g spectrum of TGS monotonically shifts to small τ values and sharply narrows as field E increases within the range of values not exceeding the half-width of the hysteresis loop. Its behavior demonstrates a tendency toward simultaneous leveling and narrowing of energies U of all barriers as E approaches E_c (Fig. 4c). SBN does not exhibit such a clear-cut dependence (Fig. 4d), and its spectra for two different E values virtually coincide. This is a consequence of the participation of only some part of the crystal in polarization within the interval of E values used in our experiments. This part, certainly, increases as E grows, but the contributions of slow processes may fortuitously be almost equal.

3.3. Electric conductivity effects on polarization relaxation. If measurement processes are slow, the ΔQ contribution of electric conductivity to the observed charges should be noticeable. The ΔQ contribution and specific resistance ρ can be estimated by separating the linear and nonlinear constituents of charge Q relaxation. For SBN:(La + Ce) and SBN:Nd crystals, such estimates equal $\rho = 2 \times 10^{13}$ and 5×10^{11} Ω cm and $\Delta Q = Et/\rho = 0.03$ and 0.7 μC/cm^2, respectively, when field $E = 600$ V/cm is switched on for 15 min at $T = 273$ K. For SBN:(La + Ce), the electric conductivity ΔQ contribution is small compared with the recorded polarization P value, whereas, for SBN:Nd, this contribution is fairly large (see Fig. 3). The polarization data shown in Figs. 3 and 4 take into account corrections for conductivity. For SBN:Nd, polarization relaxation curves are neither given nor analyzed because of the large conductivity-induced error involved in separating the linear and nonlinear observed charge components.

The role played by electric conductivity in the formation of potential barriers for relaxation centers is quite obvious in depolarization processes. The external field is then zero, $E = 0$, and through conductance, which masks relaxation, is then absent. Relaxation can therefore very accurately be recorded and analyzed. Crystals were polarized in various electric fields E for 5 min. The field was then switched off, and depolarization was recorded (Fig. 5, inset). As during polarization, changes in P after the initial jump followed power time law (4). Circles in the figures are the experimental data, the calculation results are given by solid curves, and the equilibrium P_e values are given, by dashed lines. The higher field E, the larger the P_e value. Solid circles in the $g(\ln\tau)$ spectra signify the end of depolarization measurements.

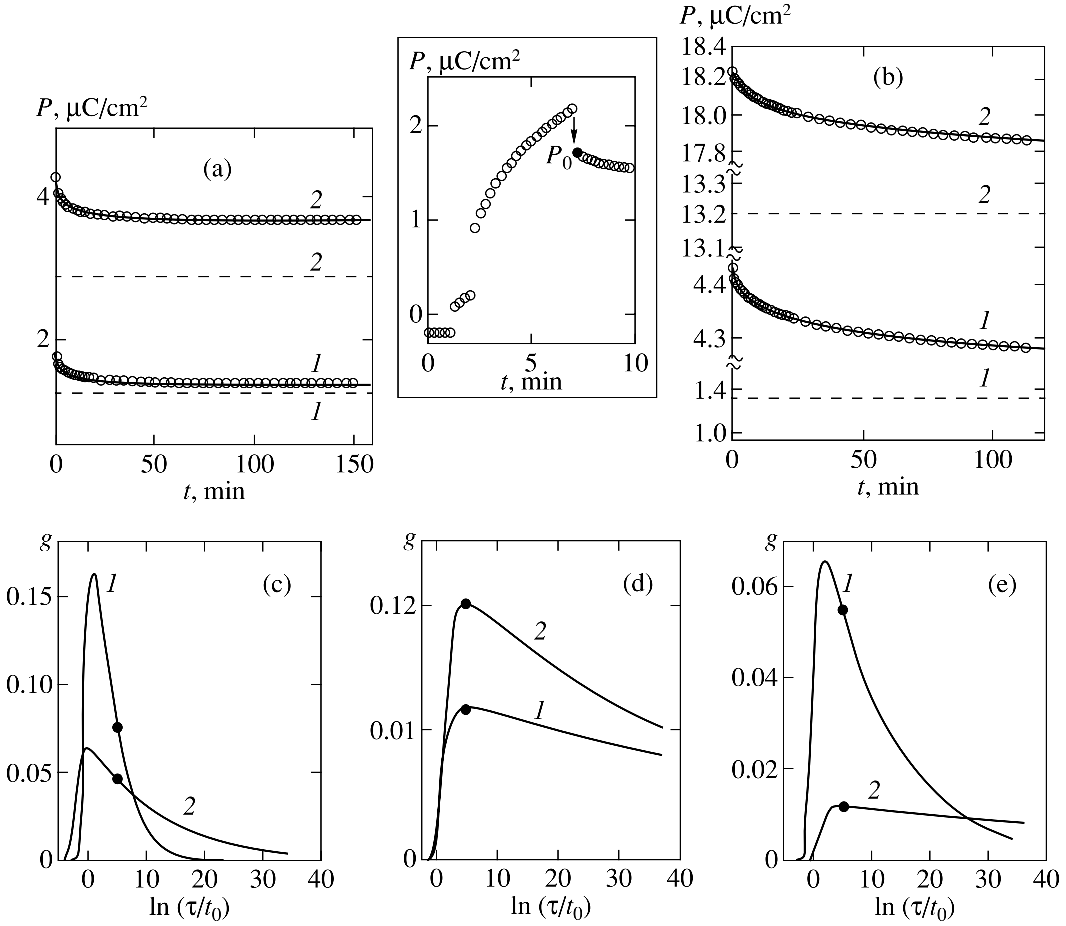


Fig. 5. (a, b) Depolarization and (c, d, e) $g(\ln\tau)$ spectra for (a, c) SBN:(La + Ce) and (b, d) SBN:Nd: (a, c) polarizing field E = (*1*) 500 and (*2*) 800 V/cm; (b, d) field E = (*1*) 1000 and (*2*) 2000 V/cm. Solid lines are calculation results, experimental data are given by circles; dashed lines are equilibrium P_e values, T = 274 K, t_0 = 1 min; (e) spectra of (*1*) SBN : (La+Ce) and (*2*) SBN:Nd reproduced for E = 1000 V/cm at 243 and 273 K, respectively. In the inset: beginning of SBN:(La + Ce) depolarization after switching off field E = 500 V/cm (P jump is marked by an arrow).

The spectrum of SBN:Nd is broader and includes giant relaxation times. The spectra of SBN:(La + Ce) and SBN:Nd crystals obtained after preliminary polarization in a 1-kV field at temperatures 60°C below those corresponding to ε maxima are compared in Fig. 5e. The relaxation and spectrum parameters are listed in Table 2.

The role played by free charge carriers in spectrum broadening and the appearance of giant barriers and relaxation times is obvious. Indeed, the screening times $\tau_s = \varepsilon\rho/4\pi$ are 1000 and 10 min for SBN:(La + Ce) and SBN:Nd, respectively. During measurements for t = 120 min after preliminary polarization, screening is only observed for the second crystal. This results in spatial charge redistribution and the appearance of an additional internal field, which increases asymmetry of local free energy F. It follows that screening in a crystal with a lower resistance decelerates depolarization and increases the equilibrium P_e value (crystal memory). In the SBN:(La + Ce) crystal with a larger resistance, the

Table 2. Relaxation and distribution spectrum $g(\tau)$ parameters for depolarization of SBN crystals

Admixture	T, K	E, V/cm	P_0, μC/cm^2	P_y, μC/cm^2	a, min	n	τ_m, min	ΔU, eV	S
La, Ce	274	500	1.75	1.217 ± 0.003	0.82 ± 0.06	0.273 ± 0.004	2.98 ± 0.26	0.061	0.65
		800	4.328	2.866 ± 0.006	0.094 ± 0.004	0.081 ± 0.001	1.16 ± 0.06	0.118	0.29
Nd	273	1000	4.432	1.312 ± 0.11	1.998 ± 0.3	0.012 ± 0.0004	166.5 ± 30	0.208	0.043
		2000	18.26	13.198 ± 0.057	2.92 ± 0.128	0.022 ± 0.0003	132.7 ± 7.4	0.18	0.076

role played by screening is insignificant. The data shown in Fig. 3 were used to estimate τ_s for the static permittivity $\varepsilon = 4\pi P/E$. Note in addition that screening is likely also to be responsible for the larger width of the hysteresis loop of SBN : Nd.

4. CONCLUSION

The observed anomalies of the kinetics of polarization and depolarization of relaxor ferroelectrics (noncoincidence of the first several cycles of dielectric hysteresis loops, the absence of an unambiguous coercive field, etc.) increase as relaxor properties become more pronounced and are obvious indications and a measure of structural disorder in these materials, which are characterized by the presence of long-lived metastable states. It is important that, for this reason, the anomalies only manifest themselves in constant or slowly varying (quasi-static) electric fields and have not been detected earlier, we believe, because of the use of overly fast measurement processes. The anomalies of SBN polarization can be used to qualitatively explain irreproducibility of properties and their deterioration under the action of fields known from the literature. The results of this work can be used in studies of transformations of arbitrary relaxor physical properties in electric fields and for the purpose of controlling their structural state. The structure of relaxors grows stabler, and the experimental data become reproducible after several slow repolarization cycles. Note also that similar disorder manifestations can, in a less pronounced form, be observed in ordinary uniform ferroelectric materials.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project nos. 99-02-17303 and 00-02-16624).

REFERENCES

1. G. A. Smolenskiĭ, V. A. Isupov, and A. I. Agranovskaya, Fiz. Tverd. Tela (Leningrad) **1**, 167 (1959).
2. M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Oxford Univ. Press, Oxford, 1977; Mir, Moscow, 1981).
3. L. E. Cross, Ferroelectrics **76**, 241 (1987).
4. Yu. S. Kuz'minov, *Ferroelectric Crystals for Laser Emission Control* (Nauka, Moscow, 1982).
5. P. B. Jamieson, S. C. Abrahams, and J. L. Bernstein, J. Chem. Phys. **48**, 5048 (1968).
6. R. R. Neurgaonkar, J. R. Oliver, W. K. Cory, *et al.*, Ferroelectrics **160**, 265 (1994).
7. V. V. Gladkiĭ, V. A. Kirikov, S. V. Nekhlyudov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 38 (2000) [JETP Lett. **71**, 24 (2000)].
8. T. R. Volk, V. Yu. Salobutin, L. I. Ivleva, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **42**, 2066 (2000) [Phys. Solid State **42**, 2129 (2000)].
9. G. L. Wood, W. W. Clark, M. J. Miller, *et al.*, IEEE J. Quantum Electron. **23**, 2126 (1987).
10. Y. Y. Zhu, J. S. Fu, R. F. Xiao, *et al.*, Appl. Phys. Lett. **70**, 1793 (1997).
11. S. Kawai, T. Ogawa, H. S. Lee, *et al.*, Appl. Phys. Lett. **73**, 768 (1998).
12. A. A. Kaminskiĭ, H. Garcia-Sole, S. N. Bagaev, *et al.*, Kvantovaya Élektron. (Moscow) **25**, 1059 (1998).
13. V. V. Gladkiĭ, V. A. Kirikov, S. V. Nekhlyudov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **39**, 2046 (1997) [Phys. Solid State **39**, 1829 (1997)].
14. V. V. Gladkiĭ, V. A. Kirikov, S. V. Nekhlyudov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **42**, 1296 (2000) [Phys. Solid State **42**, 1334 (2000)].
15. L. I. Ivleva, N. V. Bogodaev, N. M. Polozkov, *et al.*, Opt. Mater. **4**, 168 (1995).
16. T. Volk, Th. Woike, U. Doerfler, *et al.*, Ferroelectrics **203**, 457 (1997).
17. V. V. Gladkiĭ, V. A. Kirikov, E. S. Ivanova, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 499 (1999) [Phys. Solid State **41**, 447 (1999)].
18. A. K. Jonscher, *Dielectric Relaxation in Solids* (Chelsea Dielectrics Press, London, 1983).
19. V. V. Gladkiĭ, V. A. Kirikov, and E. S. Ivanova, Zh. Éksp. Teor. Fiz. **110**, 1 (1996) [JETP **83**, 161 (1996)].
20. F. Alberici, P. Doussineau, and A. Levelut, J. Phys. I **7**, 329 (1997).
21. V. I. Ditkin and A. P. Prudnikov, *Reference Book on Operational Calculus* (Vysshaya Shkola, Moscow, 1965).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 604–608.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 688–692.
Original Russian Text Copyright © 2001 by Bol'ginov, Genkin, Leviev, Ovchinnikova.

SOLIDS
Electronic Properties

A Nonlinear Response of the BiSrCaCuO Single Crystal in the Microwave Region

V. V. Bol'ginov, V. M. Genkin, G. I. Leviev*, and L. V. Ovchinnikova
Institute of Solid-State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia
**e-mail: leviev@issp.ac.ru*

Received January 25, 2001

Abstract—Third-harmonic microwave radiation of the BiSrCaCuO superconducting single crystal was studied. Two modes of microwave field–sample interactions were observed. In a weak field, a strong increase in the intensity of radiation after switching on a constant magnetic field, a hysteresis between opposite scan directions, and different harmonic amplitudes depending on the conditions of cooling (in the presence or absence of a magnetic field) were observed. These observations can be described by the generalized Ginzburg–Landau functional taking into account higher spatial derivatives of the order parameter. At a high intensity of incident waves, a magnetic field almost did not influence third-harmonic radiation, and, accordingly, hysteresis was absent. This is likely to be evidence that, at high powers, third-harmonic radiation arises as a result of generation of vortices under the action of a high-frequency magnetic field. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The microwave properties of the layered strongly anisotropic BiSrCaCuO superconductor have been studied from different points of view. A large number of studies have been concerned with interpreting microwave absorption in terms of the Josephson plasma resonance model [1–8]. The main idea of Josephson coupling between cuprate planes has also been experimentally verified by microwave measurements [9]. Owens [10] endeavored to determine the fractions of the liquid and solid phases in the vortex lattice of the BiSrCaCuO single crystals by measuring microwave absorption in a magnetic field. In [11], the authors were able to determine the temperature dependence of the λ_{ab} and λ_c penetration depths at a 10-GHz frequency from surface impedance measurements. Such measurements allow certain conclusions to be drawn on the character of interactions responsible for the formation of superconducting pairs, in particular, the quality of describing measurement results in terms of *s* or *d* models of order parameter symmetry. In all these investigations, a linear microwave response was studied. At the same time, it has been shown in several studies [12–14] that the greatest difference in the microscopic properties corresponding to different order parameter symmetry models manifests itself in the nonlinear response of the anisotropic high-T_c superconductor. For instance, the third-order response in the microwave region contains a singularity near the transition temperature in the *s*-pairing model, whereas an additional singularity of the $1/T$ type appears in the *d*-pairing model [14]. Studies of the nonlinear response of superconductors are also important for practical applications [15–17].

2. EXPERIMENTAL

In this work, we observed and studied radiation at the triple microwave frequency in the BiSrCaCuO superconducting crystal. The experiment was performed as follows. A $3 \times 3 \times 0.2$ mm^3 BiSrCaCuO single crystal with a $T_c = 90$ K transition temperature and a 1–1.5 K transition width (according to magnetic susceptibility measurements) was placed at the bottom of a bimodal cylindrical resonator with a 9-GHz fundamental frequency and a 27-GHz receiver frequency. The resonator was placed in a cryostat and, together with the cryostat, in an electromagnet. A system of Helmholtz coils and additional coils for compensating the earth field were used for measurements in a weak field (up to 180 Oe). A stronger field (up to 7 kOe) was generated by a Kapitsa magnet, into which the resonator could be introduced without remounting the sample. The system for stabilizing and varying temperature allowed measurements to be taken with a sufficient accuracy at a fixed temperature or the sample to be heated or cooled at the required rate. The microwave magnetic fields of both modes and the external magnetic field were collinear and lay in the plane normal to the *c* axis. Measurements were taken at temperatures from nitrogen (77 K) to temperatures substantially exceeding the transition point. The microwave part of the measuring circuit consisted of a high-power generator (a pulsed magnetron with a 0.8–10 μs pulse width), tuning and measuring devices (klystron generators, wavemeters, attenuators, and a power meter), and a superheterodyne receiver set to a frequency of 27 GHz. The harmonic radiation signal from the output of the superheterodyne receiver was fed to the input of an oscilloscope for visual observation and, simulta-

1063-7761/01/9303-0604$21.00 © 2001 MAIK "Nauka/Interperiodica"

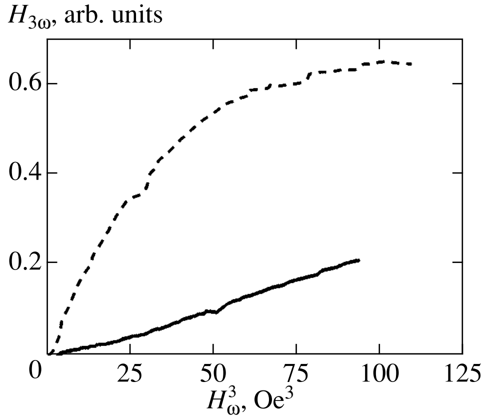


Fig. 1. Triple-frequency wave amplitude versus the cube of the incident wave amplitude in the absence (solid curve) and presence (dashed curve) of an external magnetic field (100 Oe) at 78 K.

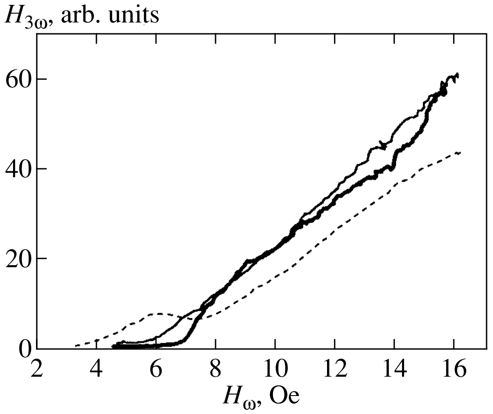


Fig. 2. Triple-frequency wave amplitude versus the incident wave amplitude in the absence (solid curve) and presence of weak (100 Oe, dashed curve) and strong (1200 Oe, thin solid curve) external magnetic fields at 78 K.

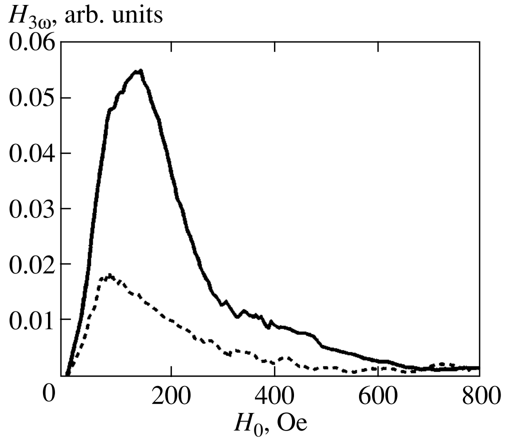


Fig. 3. Hysteresis character of the dependence of radiation intensity on a constant magnetic field. Solid and dashed curves were recorded while increasing and decreasing the field, respectively. Incident wave amplitude 1 Oe, $T = 78$ K.

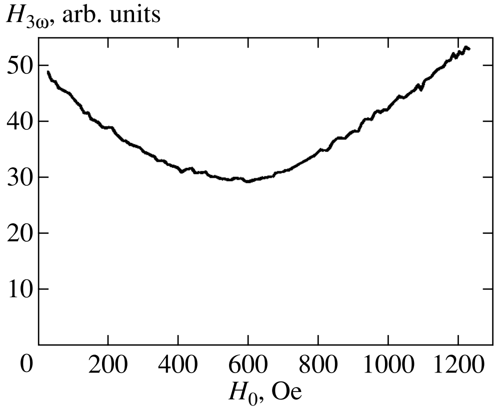


Fig. 4. External magnetic field dependence of triple-frequency wave amplitude. Incident wave amplitude 12 Oe, temperature 78 K.

neously, to the input of a stroboscopic integrator and, from the integrator, to a computer. The magnetron generator gave 50–100 pulses per second, which prevented heating of the single crystal by microwave currents. For weak signals at the integrator, averaging over many pulses could be performed to increase the signal-to-noise ratio. The unit allowed the rate of magnetic field sweep to be varied in a wide range both in the Helmholtz coils and in the Kapitsa magnet. The amplitude of the harmonic was recorded as a function of temperature, external magnetic field, and incident wave amplitude.

3. RESULTS

Odd harmonic radiation as distinguished from even harmonics is not forbidden by symmetry in a zero magnetic field. We observed radiation at the third-harmonic frequency. An unexpected finding was a strong dependence of the intensity of radiation at low incident wave amplitudes on constant magnetic field (Figs. 1 and 2). The curves shown in these figures were obtained with a single difference: when the curve shown in Fig. 2 was recorded, a 30 dB damping was introduced between the attenuator and the receiver input. Therefore, in the region of incident wave amplitudes shown in Fig. 1, the signal in Fig. 2 was either small or undetectable. Without a magnetic field, generation in Fig. 2 started at 7 Oe. Weak radiation was also observed at smaller amplitudes (Fig. 1); it is likely that, at a 7-Oe amplitude, a new and much stronger generation mechanism was switched on. At small pumping amplitudes, a well-defined hysteresis was observed (Fig. 3). The influence of the magnetic field on the intensity of harmonic radiation at large incident wave amplitudes was insignifi-

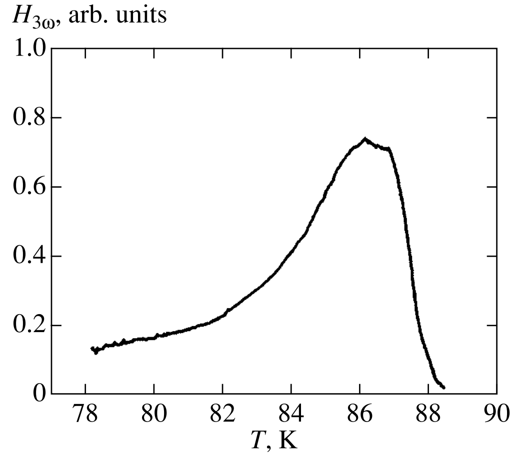


Fig. 5. Temperature dependence of triple-frequency wave amplitude in the absence of an external magnetic field. Incident wave amplitude 5 Oe.

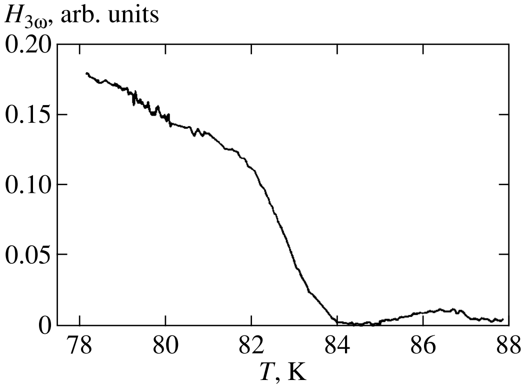


Fig. 6. Temperature dependence of triple-frequency wave amplitude in the presence of an external magnetic field (100 Oe). Incident wave amplitude 1 Oe.

cant, as follows from Fig. 4. This record was made while changing the magnetic field scan direction; it shows that hysteresis was absent. The temperature dependences of the intensity of frequency radiation in the absence and presence of a field are shown in Figs. 5 and 6, respectively. The record in Fig. 6 was made at a 1-Oe incident wave amplitude, and this was the reason for a strong influence of the external magnetic field on the shape of the curve. According to Fig. 6, the signal disappeared when the sample was heated to 85 K in the magnetic field and reappeared before the transition. The signal near the transition was independent of the field. Measurements were also performed using a different geometry, when all magnetic fields (microwave at the fundamental and harmonic frequencies and external) were oriented along the c axis of the crystal. In this geometry, we did not observe radiation stimulation by an external magnetic field.

4. DISCUSSION

First, consider experiments at small incident wave amplitudes, when the dependence of the intensity of triple-frequency radiation on the incident wave intensity can be considered cubic. The most interesting result is, we believe, a strong increase in the third-harmonic amplitude in the presence of a magnetic field. No such dependence was observed for YBaCuO single crystals [18]. In ceramic samples with a branched system of Josephson couplings between granules, even a weak magnetic field suppressed generation, and in fields of 15 Oe the signal was difficult to observe. Introducing a field into the sample at a temperature below the transition point has two consequences: a surface current related to the magnetization jump and vortices in the bulk arise. We believe that the intensity of radiation increases precisely because of the surface current. The following experiment lends support to this conclusion. A magnetic field was introduced at a low temperature, radiation arose, and the sample was heated to a temperature above the transition point to the normal state, in which radiation was absent. After this, the sample was cooled without varying the external magnetic field to the initial low temperature. The intensity of triple-frequency radiation was then several times lower than the initial intensity, although the number of vortices did not change during cooling. The distribution of vortices during cooling in a field is more uniform than when the field is introduced at a low temperature. This, however, does not refer to the thin layer near the surface, where the distribution of vortices is always strongly nonuniform [19, 20]. Precisely the smallness of this layer determines weak harmonic radiation during cooling in an external field.[1] Further, we suggest a phenomenological description of stimulation of harmonic radiation by the current based on the Ginzburg–Landau functional, which takes into account higher spatial derivatives of the order parameter than those used in considering a linear response. Let us write free energy f of a superconductor in the form

$$
\begin{aligned}
f = {} & \alpha|\psi^2| + \beta|\psi^4|/2 + \left|\left(-i\nabla_x - \frac{2e}{c}A_x\right)\psi\right|^2/2m_a \\
& + \left|\left(-i\nabla_y - \frac{2e}{c}A_y\right)\psi\right|^2/2m_a + \left|\left(-i\nabla_z - \frac{2e}{c}A_z\right)\psi\right|^2/2m_c \\
& + \delta\left(\left|\left(-i\nabla_x - \frac{2e}{c}A_x\right)^2\psi\right|^2 + \left|\left(-i\nabla_y - \frac{2e}{c}A_y\right)^2\psi\right|^2\right. \\
& + \left|\left(-i\nabla_x - \frac{2e}{c}A_x\right)\left(-i\nabla_y - \frac{2e}{c}A_y\right)\psi\right|^2 \\
& \left. + \left|\left(-i\nabla_y - \frac{2e}{c}A_y\right)\left(-i\nabla_x - \frac{2e}{c}A_x\right)\psi\right|^2\right)/4,
\end{aligned}
\qquad (1)
$$

---

[1] A similar situation arose in experiments with YBaCuO single crystals, in which, however, second-harmonic radiation was observed [18].

where ψ is the complex order parameter; A is the vector potential; α, β, and δ are the phenomenological expansion coefficients; e is the charge of the electron; and m_a and m_c are the effective masses along the crystal axes. The terms in the expansion of energy in powers of spatial derivatives that are proportional to the δ parameter are usually ignored in describing microwave responses. Let us assume that $A = 0$ and write the order parameter in the form

$$\psi = \phi e^{i\mathbf{kr}} \quad (2)$$

with a constant ϕ modulus. The $\mathbf{k}(k_x, k_y)$ and $\mathbf{r}(x, y)$ vectors are two-dimensional. At such an order parameter, the current is given by the equation

$$\mathbf{j} = 2e\phi^2\mathbf{k}(1 + m_a\delta k^2)/m_a. \quad (3)$$

Minimizing the free energy functional with respect to ϕ yields (4) (an algebraic relation between the order parameter and the α, β, and δ coefficients) and (5) (an expression for the current of superconducting electrons):

$$\alpha + \beta\phi^2 + \frac{k^2}{2m_a} + \frac{\delta k^4}{4} = 0, \quad (4)$$

$$\mathbf{j} = 2e\phi_0^2\mathbf{k}\frac{1 + m_a\delta k^2}{m_a} - \mathbf{k}k^2\frac{1 + 3m_a\delta k^2/2}{2m_a^2} - \frac{\mathbf{k}k^6\delta^2}{4\beta}, \quad (5)$$

where $\phi_0 = -\alpha/\beta$ is the equilibrium order parameter value corresponding to $\mathbf{k} = 0$. Set $\mathbf{k} = \mathbf{k}_0 + \mathbf{q}$, where $\mathbf{k}_0$ is determined by the constant constituent of the superconducting current, and $\mathbf{q} = -2e\mathbf{A}_\omega/c$ is determined by the $\boldsymbol{A}_\omega$ high-frequency vector potential. Contributions to the third-harmonic amplitude are made by the terms of (5) that are proportional to $\mathbf{A}_\omega^3$. These terms have the form

$$\delta q^3\phi_0^2, \quad q^3/m_a\beta, \quad \gamma q^3/m_a\beta, \quad \gamma^2 q^3/m_a\beta,$$

where $\gamma = \delta m_a k_0^2$. The first two terms give the harmonic amplitude independent of a constant current in the sample, and the last two terms describe radiation which only appears in the presence of a constant current. If $\gamma > 1$, the nonlinear current is determined by the last term in (5),

$$\mathbf{j}^{nl} = -\mathbf{k}k^6\delta^2/4\beta, \quad (6)$$

and we observe an increase in the intensity of radiation in the presence of a constant current. Equation (6) for a nonlinear source can be used to find the angular dependence of the current

$$(j^{nl})^2 \approx 9 + 288\cos^2\theta + 672\cos^4\theta + 256\cos^6\theta, \quad (7)$$

where θ is the angle made by $\mathbf{k}_0$ and $\mathbf{q}$. Figure 7 shows that dependence (7) satisfactorily correlates with the experimental angular dependence of the intensity of radiation. As has been mentioned above, stimulation of radiation by current is absent when the fields are oriented along the c axis. Within the framework of the phenomenological approach that we use, this corresponds to the smallness of spatial derivatives in this geometry because of the strong anisotropy of BiSrCaCuO.

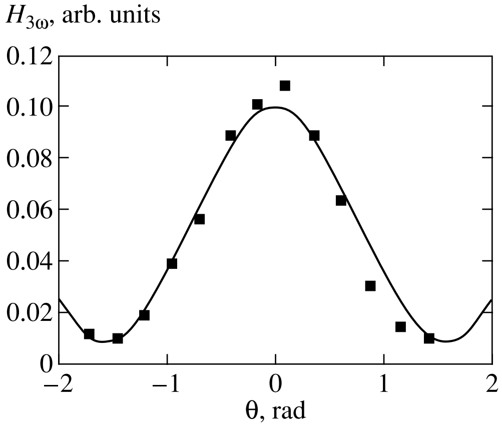


Fig. 7. Triple-frequency wave amplitude versus the angle between constant and microwave magnetic fields: (black squares) experiment; (solid curve) theory. Incident wave amplitude, 1 Oe; temperature, 78 K; external magnetic field, 100 Oe.

5. CONCLUSION

We showed that the intensity of the third harmonic emitted by a BiSrCaCuO superconducting single crystal in a microwave field could strongly increase in the presence of a constant magnetic field. An important role in radiation was played by the screening surface current. The experimental results can be understood using the Ginzburg–Landau theory with a generalized functional including higher spatial derivatives of the order parameter.

ACKNOWLEDGMENTS

The authors are indebted to G.A. Emel’chenko for providing the samples and D.V. Shovkun for the high-frequency susceptibility measurements and very useful discussions.

REFERENCES

1. E. B. Sonin, Phys. Rev. Lett. **79**, 3732 (1997).
2. Y. Matssuda, M. B. Gaifullin, K. Kumagai, *et al.*, Phys. Rev. B **55**, R8685 (1997).
3. E. B. Sonin, Phys. Rev. B **60**, 15430 (1999).
4. O. K. C. Tsui *et al.*, Phys. Rev. Lett. **73**, 724 (1994).
5. S. Sakamato *et al.*, Phys. Rev. B **53**, R14749 (1996).
6. M. D. Gaifulin, Y. Matssuda, and L. N. Bulaevskiĭ, Phys. Rev. Lett. **81**, 3551 (1998).
7. A. Yurgens *et al.*, Phys. Rev. B **59**, 7196 (1999).

8. N. Morozov, M. P. Maley, L. N. Bulaevskii, *et al.*, Phys. Rev. Lett. **82**, 1008 (1999).
9. R. Kleiner, F. Steinmeyer, G. Kunkel, and P. Muller, Phys. Rev. Lett. **68**, 2394 (1992).
10. F. J. Owens, J. Phys. Chem. Sol. **55**, 167 (1994).
11. T. Jacobs, S. Sridhar, Qaung Li, *et al.*, Phys. Rev. Lett. **75**, 4516 (1995).
12. D. Xu, S. K. Yip, and J. A. Sauls, Phys. Rev. B **51**, 16233 (1995).
13. B. P. Stojkovic and O. T. Valls, Phys. Rev. B **51**, 6049 (1995).
14. T. Dahm and D. J. Scalapino, J. Appl. Phys. **81**, 2002 (1997).
15. L. F. Cohen, A. Cowie, J. C. Gallop, *et al.*, J. Supercond. **10**, 85 (1997).
16. M. Hein *et al.*, IEEE Trans. Appl. Supercond. **7**, 1264 (1997).
17. T. Dahm, D. J. Scalapino, and B. A. Willemsen, J. Appl. Phys. **86** (7), 4055 (1999).
18. T. T. Bol’ginov, V. M. Genkin, G. I. Leviev, and L. V. Ovchinnikova, Zh. Éksp. Teor. Fiz. **115**, 2242 (1999) [JETP **88**, 1229 (1999)].
19. L. Burlachkov, Phys. Rev. B **47**, 8056 (1993).
20. E. Zeldov, A. I. Larkin, V. B. Geshkenbein, *et al.*, Phys. Rev. Lett. **73**, 1428 (1994).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 609–614.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 693–699.
Original Russian Text Copyright © 2001 by Lyubutin, Frolov.

SOLIDS
Electronic Properties

Spin Crossover in the Magnetic Phase Transition in $YBa_2(Cu_{1-x}Fe_x)_3O_{7\pm\delta}$ Superconductors

I. S. Lyubutin* and K. V. Frolov

Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninskii pr. 59, Moscow, 117333 Russia
**e-mail: lyubutin@ns.crys.ras.ru*

Received February 16, 2001

Abstract—All iron ions in the Cu1 and Cu2 local lattice sites of the $YBa_2(Cu_{0.9}{}^{57}Fe_{0.1})_3O_{7.01}$ superconductor with $T_c = 31$ K experienced magnetic ordering below $T_m = 22$ K. Therefore, at $T < T_m$, magnetic ordering coexisted with superconductivity. According to the Mössbauer spectroscopy data, iron ions in Cu2 (Fe2) sites were in the low-spin state at $T < T_m$ ($S = 3/2$ or $1/2$), whereas an equal number of iron ions in Cu1 (Fe1) sites were in the high-spin Fe^{3+} state ($S = 5/2$). The magnetic transition near T_m changed iron ion spin states—low-spin ions turned into high-spin ions, and vice versa. This preserved the spin balance between iron ions in the Cu1 and Cu2 layers. Control measurements on other samples of the $YBa_2(Cu_{1-x}Fe_x)_3O_{7\pm\delta}$ series substantiated these conclusions. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Experimentally studying correlations and competition between magnetism and superconductivity with theoretically analyzing magnetic pairing mechanisms is one of the main directions in studies of the nature of high-temperature superconductivity [1, 2]. The discovery of magnetic ordering of Fe ions in the superconducting compound $YBa_2(Cu_{1-x}Fe_x)_3O_7$ aroused much interest [3–10]. This effect was observed in the Mössbauer spectra of ^{57}Fe nuclei, whereas neutron diffraction did not detect any long-range magnetic order in the superconducting state [11–13]. Slow spin relaxation of Fe ions and spin glass-type ordering were considered as possible reasons for magnetic hyperfine interactions in these compounds [7, 8, 10, 12, 14]. In addition, the possibility of formation of iron ion clusters in Cu1 sites was analyzed [15–17]. It was, however, shown recently [13] that Fe ions in Cu1 regular copper sites with a square-planar oxygen coordination also experienced magnetic ordering, although these sites are usually not involved in clustering. It follows that the problem of coexistence, correlation, and competition between magnetism and superconductivity still remains unclarified and intriguing.

It was found with definiteness from the Mössbauer hyperfine interaction parameters (see [18] and the references therein) that, at room temperature, Fe ions in Cu2 sites of the $YBa_2Cu_3O_7$ (Y-1-2-3) superconductor doped with iron were in the Fe^{3+} high-spin state ($S = 5/2$), whereas Fe^{3+} ions in Cu1 sites were in a state with a lower spin ($S = 3/2$ and/or $1/2$). Nor could the presence of Fe^{4+} in Cu1 sites be ruled out [19, 20]. At low temperatures (in the superconducting state), iron ions were magnetically ordered in all Cu1 and Cu2 local sites. The Mössbauer lines of Fe in Cu1 sites were unambiguously identified in the Zeeman spectra, whereas the lines of Fe in Cu2 sites were difficult to observe because of complex superposition of several subspectra (see discussion in [21]).

In this work, we thoroughly studied the transformation of the Mössbauer spectrum magnetic components into paramagnetic doublets in $YBa_2(Cu_{1-x}{}^{57}Fe_x)_3O_{7\pm\delta}$ superconductors caused by the low-temperature magnetic transition near the T_m point. We observed that, at $T < T_m$, Fe ions in Cu2 sites were in a state with a lowered spin, whereas an equal number of Fe ions in Cu1 sites had the high spin. At $T > T_m$, spin crossover was observed: iron ions of both types changed their spin states and high-spin ions turned low-spin and vice versa.

2. EXPERIMENTAL

The $YBa_2(Cu_{0.9}{}^{57}Fe_{0.1})_3O_{7.01}$ sample enriched in the ^{57}Fe isotope by 96% was prepared by the standard ceramic procedure [22, 23] with final annealing in oxygen at 450°C for 24 h followed by cooling to room temperature in oxygen at a rate of 1 K/min. The content of oxygen was determined iodometrically and refined by processing the neutron diffraction data according to Rietweld [11, 18]. The sample was superconducting and had a $T_c = 31$ K superconducting transition temperature determined by measuring the electrical resistance by the inductive method [23]. The ^{57}Fe Mössbauer absorption spectra were recorded on a standard spectrometer with a constant acceleration in the transmis-

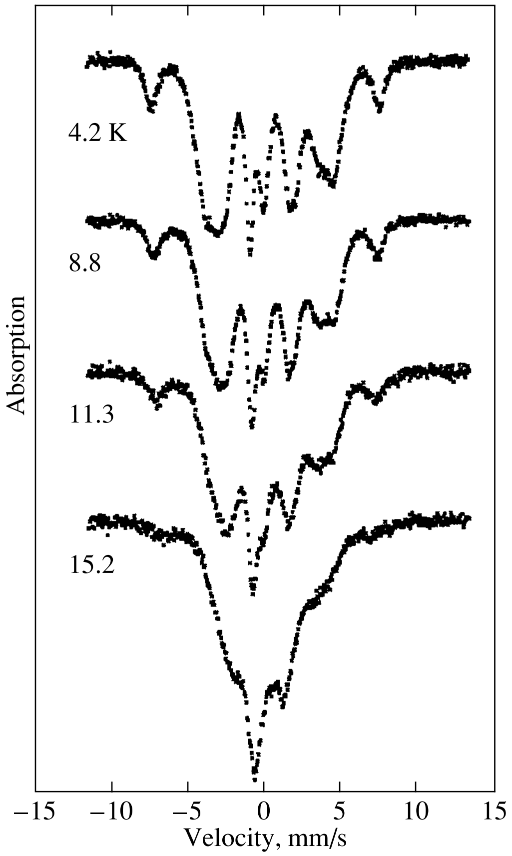


Fig. 1. Mössbauer spectra of ^{57}Fe nuclei in the $YBa_2(Cu_{0.9}{}^{57}Fe_{0.1})_3O_{7.01}$ superconductor at several temperatures below the T_m magnetic transition point.

sion geometry. A ^{57}Co(Rh) source of gamma quanta had room temperature. Measurements at 4.2–295 K were performed using a flow helium cryostat. The Mössbauer spectra were treated in terms of theoretical models by the method of least squares with the use of computer programs described in [18, 21, 22–26].

3. RESULTS AND DISCUSSION

A typical Mössbauer spectrum recorded at room temperature consisted of three quadrupole doublets ($D1$, $D2$, and $D3$). Their isomer shifts (ISs) and quadrupole splittings (QSs) are listed in Table 1. It was reliably established [18, 27] that the $D3$ doublet should be assigned to Fe^{3+} ions in Cu2 sites; its hyperfine interaction parameters were evidence of the high-spin state of these ions, $S = 5/2$. The $D1$ and $D2$ doublets corresponded to Cu1 local sites with Fe^{3+} ions in an intermediate ($S = 3/2$) and/or low-spin ($S = 1/2$) state [19, 23, 27]. The $D2$ doublet had broadened lines and in reality might contain several additional components (see [18]). Nor could the presence of Fe^{4+} ions in Cu1 sites be ruled out [19, 20].

The low-temperature Mössbauer spectra exhibited hyperfine magnetic splitting (Fig. 1). This was evidence that Fe ions were magnetically ordered in all local sites that they occupied in the 1-2-3 structure and that this magnetic order coexisted with superconductivity. The spectra could be described by three Zeeman subspectra $M1$, $M2$, and $M4$ (we use the notation introduced in [25, 26, 28]) with the mean hyperfine magnetic saturation fields $H_{hf}(4.2\ K) = 211$, 265, and 464 kOe, respectively. These subspectra corresponded to Fe ions in various Cu1 and Cu2 sublattice sites; their hyperfine coupling parameters are listed in Table 1. Their values closely agree with the data obtained by other authors [19, 29–31]. The resonance lines of the $M1$ and $M2$ magnetic components (with low H_{hf} field values) were broadened and overlapped each other. For this reason, we only estimated the mean values of their hyperfine coupling parameters with the use of special models (see [21, 24, 30, 32, 33]). The mechanisms of line broadening such as distribution of magnetic fields caused by chemical disordering, spin relaxation and spin glass-type behavior, and zero spin fluctuations in low-dimensional magnetic systems were discussed in [7, 9, 26, 29, 32, 34].

Table 1. Hyperfine interaction parameters for M_i magnetic components at $T = 4.2$ K and for D_i paramagnetic doublets at room temperature obtained from the Mössbauer spectra of $YBa_2(Cu_{0.9}{}^{57}Fe_{0.1})_3O_{7.01}$ superconductor

T, K	D_i, M_i	IS, mm/s	ε, mm/s	QS, mm/s	Γ, mm/s	A_{rel}, %	H_{hf}, kOe
4.2	$M1$	+0.10(4)	+0.59(5)	–	0.64(5)	29.5	211
	$M2$	+0.16(4)	+0.53(5)	–	1.22(5)	56.2	265
	$M4$	+0.56(3)	–0.54(5)	–	0.74(5)	14.3	464
298	$D1$	+0.057	–	1.993	0.354	33.5	–
	$D2$	–0.022	–	0.981	0.530	54.8	–
	$D3$	+0.369	–	0.651	0.337	11.7	–

Note: IS is the isomer shift with respect to iron metal, QS is the quadrupole splitting of paramagnetic spectra, ε is the quadrupole shift in magnetic spectra ($\varepsilon = (QS/4)(3\cos^2\theta - 1)$, where θ is the angle made by H_{hf} and the principal electric field gradient axis V_{zz}), Γ is the line half-width, A_{rel} is the relative area of the subspectrum, H_{hf} is the mean magnetic field on the iron nucleus. Parenthesized values are errors of measurements; if the latter are not given, they have typical values of ±0.005 for IS, QS, and Γ; ±1.5% for A_{rel}; and ±2 kOe for H_{hf}.

Field H_{hf} is usually related to the spin quantum number of iron (S) by the approximate equality $H_{hf} \approx 220\langle S_z\rangle$ kOe [13, 19, 26, 29, 32, 34], which as a rule holds in magnetic oxides, to describe the contribution to the field of ionic core spin polarization [35]. Considering that field H_{hf} can be lowered because of covalency or spin fluctuation effects [35], both H_{hf} field values for the $M1$ and $M2$ components (211 and 265 kOe) can correspond to the ($S = 3/2$) state of Fe^{3+}. The $H_{hf}(M1)$ value is also consistent with an intermediate spin state of Fe^{4+}. The low field values for the $M1$ and $M2$ components were explained on the assumption of the low-spin state of iron ions [8, 19, 21, 29, 31]. In this work, we found that not only fields H_{hf} had low values but also the isomer shifts of the $M1$ and $M2$ components were small (see Table 1, taking into account the temperature shift estimated below). A similar effect was recently observed by Prasanna *et al.* [31]. This is additional evidence that Fe ions in $M1$ and $M2$ local sites are in a state with a lowered spin (the $S = 3/2$ intermediate state is most probable [19]).

Consider the behavior of the less intense $M4$ component with the highest local magnetic field (the left and right outermost spectrum lines, Fig. 1). In the superconducting state, the $M4$ component has a large positive isomer shift, IS(4.2 K) ≈ +0.56 mm/s (with respect to iron metal at room temperature). Similar hyperfine parameters were obtained for this component in other studies [13, 19, 21, 24, 25, 29, 31, 34]. A further analysis can conveniently be performed after recalculating the IS(4.2 K) value to room temperature. It was shown in [19, 25] that, in 1-2-3-type compounds, the temperature dependence of the “center of gravity” of the Mössbauer spectrum was largely caused by the second-order Doppler shift, δ_R, and that $\partial\delta_R/\partial T$ approximately equaled -7×10^{-4} mm/(s K). It follows that the IS($M4$) value reduced to room temperature should approximately equal +0.35 mm/s. This value is typical of the high-spin state of Fe^{3+} ($S = 5/2$). The hyperfine field value for the $M4$ component, $H_{hf} = 464$ kOe, is also typical of the high-spin state of Fe^{3+} and is consistent with the IS value. This proves that the low-temperature Mössbauer spectra contain a component corresponding to Fe^{3+} ions in the high-spin state. It will, however, be shown that this component cannot be assigned to iron in Cu2 sites.

The intensity of the $M4$ component gradually decreases as temperature increases, whereas the $H_{hf}(M4)$ field value changes insignificantly. The temperature dependence of $H_{hf}(M4)$ was analyzed in terms of theoretical models in [13, 21, 28, 29, 31, 34] and explained on the assumption of low-dimensional magnetic ordering in these local sites. It was suggested in several studies [19, 21, 29] that the $M4$ component should be assigned to iron in Cu1 sites, which were in all likelihood coordinated by six oxygen ions [19, 21]. Recently, it was suggested that the $M4$ component corresponded to iron in Cu2 sites [31]. In the vicinity of $T_m \approx 20$ K (Figs. 1, 2), the Mössbauer spectra change sharply, which is evidence of the transition to the paramagnetic state. This transition is extended over some temperature interval and is accompanied by the appearance of paramagnetic doublets in the spectrum because of a decrease in the intensity of the magnetic components. Figure 2 clearly shows that the $M4$ magnetic component completely disappears at 19.7 K, whereas the $D3$ paramagnetic doublet (shown by arrows in Fig. 2) still does not appear at this temperature. The conclusion can be drawn that $M4$ does not transform into $D3$ as a result of the magnetic transition. It follows that the $M4$ and $D3$ components should be assigned to Fe ions in different local sites. The $D3$ doublet appears at a higher temperature (Fig. 2) from the magnetic component with a low $H_{hf} < 265$ kOe field; that is, it corresponds to iron ions with lower spins at $T < T_m$. The appearance of $D3$ from low H_{hf} field magnetic components was also

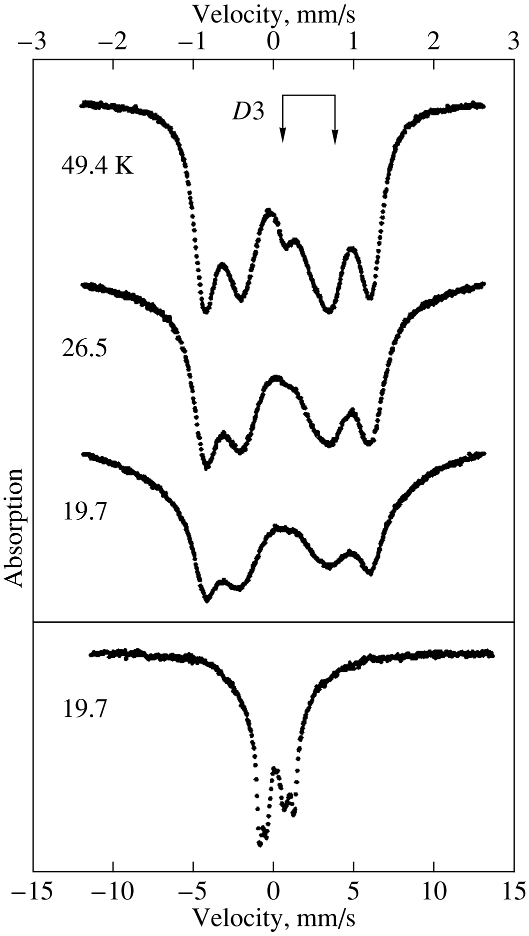


Fig. 2. Mössbauer spectra of ^{57}Fe nuclei in the $YBa_2(Cu_{0.9}{}^{57}Fe_{0.1})_3O_{7.01}$ superconductor at several temperatures near the T_m magnetic transition point. The positions of the $D3$ doublet lines are indicated by arrows. Velocity scales are different in the top and bottom parts of the figure.

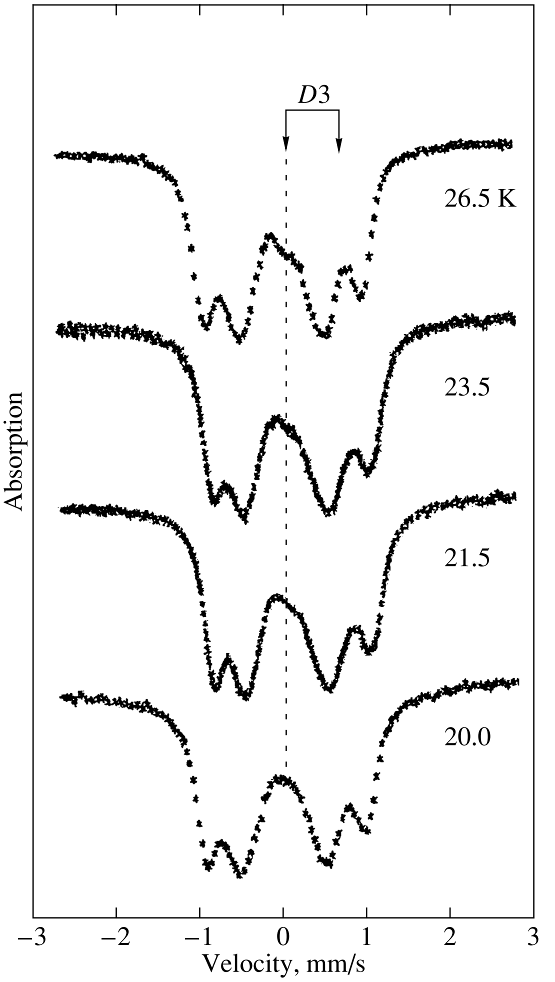


Fig. 3. Mössbauer spectra of ^{57}Fe nuclei in the $YBa_2(Cu_{0.9}{}^{57}Fe_{0.1})_3O_{7.07}$ superconductor at several temperatures near the T_m magnetic transition point. The positions of the $D3$ doublet lines are indicated by arrows.

observed in samples with a different iron concentration [13, 21].

A detailed analysis of the spectra recorded near T_m shows that the $M4$ component transforms into a $D2$-type component of iron in Cu1 sites, and its isomer shift at room temperature approximately equals –0.02 mm/s (see Table 1). It follows that, as a result of the magnetic transition at the T_m temperature, the IS value for Fe ions in ($M4$/$D2$) · Cu1 local sites decreases approximately by 0.37 mm/s. The $D3$ paramagnetic doublet of iron in Cu2 sites appears from magnetic components with low H_{hf} and IS values. Estimates taking into account the temperature shift (see Table 1) show that the IS value for Fe ions in $D3$-Cu2 sites increases as a result of the magnetic transition by a value equal to the decrease in IS for iron ions in $M4$-Cu1 sites. Several very important conclusions follow from this result.

(1) At $T < T_m$, the magnetic moments of iron ions in Cu2 sites are "frozen" even in the superconducting region, but these ions occur in a state with a lowered spin, most likely, with an intermediate Fe^{3+} spin value ($S = 3/2$). The suggestion of lowered spins of Fe^{3+} ions in Cu2 sites in the superconducting region was also made in [20].

(2) The $M4$ magnetic component with the highest H_{hf} field and a large isomer shift value corresponds to iron ions in Cu1 sites. Below T_m, iron ions of this type are in the high-spin Fe^{3+} state ($S = 5/2$).

(3) Fe^{3+} ions in Cu2 sites experience a transition from the magnetically ordered state with a lowered (intermediate) spin to the paramagnetic high-spin state (the $D3$ doublet) as temperature increases and the magnetic phase transition occurs at T_m. At the same time the high-spin Fe^{3+} state of ions in Cu1 sites (the $M4$ magnetic component) transforms into the low-spin paramagnetic state.

A thorough analysis of correlations between isomer shift values and the spin state of iron ions in various oxygen environments in the 1-2-3 phase was performed by Lines and Eibschutz [19], who used the data reported by Menil [36]. According to the diagrams given in [19], a change in the spin of Fe^{3+} ions from $S = 5/2$ to $S = 3/2$ without a change in the coordination number of iron causes an approximately 0.3 mm/s increase in the IS value. This value agrees closely with our experimental data and, therefore, substantiates the occurrence of the spin crossover suggested above.

It is likely that a certain number of Fe ions in Cu1 and Cu2 sites change their spins from low to high and vice versa as a result of the magnetic transition. This preserves the spin balance between the $M4$ and $D3$ local sites. We found that the areas below the $M4$ and $D3$ Mössbauer components were close to each other; that is, the numbers of iron ions in the corresponding sites were approximately equal (for this reason, the $M4$ component was erroneously assigned to iron in Cu2 sites in several studies). This leads us to conclude that each Fe ion in the Cu2 layer (Fe2 ion) has a paired Fe ion in the Cu1 layer (Fe1 ion); that is, Fe2–Fe1 iron ion pairs are formed. The total spin of such a pair is constant, but the spins of each pair member are different at temperatures above and below T_m. The spin state of the Fe1($S = 5/2$) + Fe2($S = 3/2$) pair existing at $T < T_m$ transforms into the Fe1($S = 3/2$) + Fe2($S = 5/2$) state at $T > T_m$. Curiously, two pair members occur in different copper layers, Cu1 and Cu2. This raises the question whether similar pairing of copper ions takes place and whether the observed effect has a bearing on pairing of superconducting electrons. In any event, iron ions in Cu2 planes have lower spins in the superconducting than in the normal (nonsuperconducting) state.

Note that the T_c temperature (T_c = 31 K) of the $YBa_2(Cu_{0.9}{}^{57}Fe_{0.1})_3O_{7.01}$ sample studied in this work was fairly close to $T_m \approx 22$ K. It is therefore not quite

Table 2. Hyperfine interaction parameters for the *M*4 magnetic component at 4.2 K and for the *D*3 paramagnetic component at room temperature for $YBa_2(Cu_{1-x}Fe_x)_3O_y$ superconductors saturated with oxygen

x	y	T_c	D_i, M_i	IS, mm/s	ε, mm/s	QS, mm/s	Γ, mm/s	A_{rel}, %	H_{hf}, kOe
0.07	6.93	44	*M*4	+0.59(5)	–0.48(5)	–	0.88(5)	10.9(1.5)	452(3)
			*D*3	+0.355(5)	–	0.638	0.305(5)	10.7(1.5)	–
0.10	7.07	43	*M*4	+0.58(5)	–0.45(5)	–	0.71(5)	16.8(1.5)	457(1)
			*D*3	+0.300(5)	–	0.592	0.341(5)	15.5(1.5)	–
0.15	7.15	9	*M*4	+0.53(5)	–0.48(5)	–	0.82(3)	20.8(1.5)	483(1)
			*D*3	+0.360(5)	–	0.632	0.355(5)	20.3(1.5)	–

Note: The notation is the same as in Table 1.

clear whether the superconducting or the magnetic transition is responsible for the observed spin crossover. To elucidate this point, we studied another sample, $YBa_2(Cu_{0.9}{}^{57}Fe_{0.1})_3O_{7.07}$, which was prepared similarly to the first one but annealed in oxygen for a longer time. Because of the higher oxygen content in the second sample, its T_c temperature increased to 43 K, whereas the T_m temperature decreased approximately to 16 K (some properties of the first and second samples were described in [13] and [28], respectively). The Mössbauer spectra of the second sample in the region of the low-temperature magnetic transition are shown in Fig. 3. These results allow us to draw two conclusions.

(1) The *D*3 doublet is observed in the second sample, as in the first one, at a higher temperature than that at which the *M*4 component disappears. The *D*3 doublet appears from the magnetically split background which remains after the major fraction of iron ions experiences the transition to the paramagnetic state. This lends support to the conclusion that the *D*3 and *M*4 components should be assigned to iron in different local sites.

(2) The spin crossover described above occurs in the magnetic rather than the superconducting transition.

The hyperfine coupling parameters for the *M*4 component in $YBa_2(Cu_{1-x}Fe_x)_3O_{7\pm\delta}$ are given in Table 2 for $x = 0.07$ and 0.15. These data were obtained in our preceding work [13], but have not been published yet. The data given in Table 2 substantiate the high IS value for the *M*4 component and closeness of the areas under the *M*4 and *D*3 components. This is evidence that the conclusions made above are valid for a wide range of iron concentrations in the 1-2-3 phase.

The magnetic behavior of 1-2-3 samples depleted of oxygen lent additional support to the idea of pairing Fe1 and Fe2 ions in *M*4–Cu1 and *D*3–Cu2 sites. The magnetic behavior of the *M*4 component under temperature variations was found [13, 21, 37] to follow the law characteristic of three-dimensional magnetic order, which was typical of the Cu2 sublattice. At the same time, iron ions in the other (Cu1) local sites did not obey this law, and their magnetic behavior including order parameter crossover [13, 21] was more complex.

4. CONCLUSION

In $YBa_2(Cu_{1-x}Fe_x)_3O_{7\pm\delta}$ superconductors, the *M*4 magnetic component with the highest $H_{hf} \approx 470$ kOe field value and the largest positive isomer shift IS should be assigned to high-spin Fe^{3+} ions in Cu1 sites. Iron ions in Cu2 sites have smaller H_{hf} and IS values, which is evidence of their low-spin state at $T < T_m$. The magnetic phase transition at the T_m point changes spins of all iron ions in Cu2 sites from low ($S = 3/2$) to high ($S = 5/2$). Simultaneously, an equal number of iron ions in Cu1 sites exchange their high spins ($S = 5/2$) for low spins ($S = 3/2$). It can be suggested that Fe^{3+} ions in Cu2 and Cu1 layers form pairs in superconducting compounds. Each pair member has different spin at temperatures above and below T_m, but the total spin of the pair remains constant as temperature varies.

ACKNOWLEDGMENTS

The authors are deeply indebted to A.Ya. Shapiro for the preparation of the samples and V.G. Terzieva for help with the experiments. This work was financially supported by the “Spin” project of the Ministry of Industry and Science of the Russian Federation (grant 99031) and the Russian Foundation for Basic Research (project no. 01-02-16769a).

REFERENCES

1. Yu. Izyumov, N. M. Plakida, and Yu. N. Skryabin, Usp. Fiz. Nauk **159**, 621 (1989) [Sov. Phys. Usp. **32**, 1060 (1989)].
2. V. J. Emery, J. Appl. Phys. **67**, 4666 (1990).
3. Z. Q. Qiu, Y. M. Du, H. Tang, *et al.*, J. Magn. Magn. Mater. **69**, L221 (1987).
4. X. Z. Zhou, M. Raudsepp, P. Pankhurst, *et al.*, Phys. Rev. B **36**, 7230 (1987).
5. I. S. Lyubutin, E. M. Smirnovskaya, V. G. Terziev, and A. Ya. Shapiro, Pis’ma Zh. Éksp. Teor. Fiz. **47**, 196 (1988) [JETP Lett. **47**, 238 (1988)].

6. T. Tamaki, T. Komai, A. Ito, *et al.*, Solid State Commun. **65**, 43 (1988).
7. I. Nowik, M. Kowitt, I. Felner, and E. R. Bauminger, Phys. Rev. B **38**, 6677 (1988).
8. I. S. Lyubutin and V. G. Terziev, in *Progress in High Temperature Superconductivity*, Ed. by V. L. Aksenov, N. M. Plakida, and N. N. Bogolyubov (World Sci. Publ., Singapore, 1989), Vol. 21, p. 281.
9. S. Nasu, Y. Oda, T. Kohara, *et al.*, in *Progress in High Temperature Superconductivity*, Ed. by Y. Murakami (World Sci. Publ., Singapore, 1989), Vol. 15, p. 214.
10. Z. Q. Qiu, Y. W. Du, H. Tang, and J. C. Walker, J. Magn. Magn. Mater. **78**, 359 (1989).
11. A. M. Balagurov, F. Bouree, I. S. Lyubutin, and I. Mirebeau, Physica C (Amsterdam) **228**, 299 (1994).
12. S. Katano, T. Matsumoto, A. Matsushita, *et al.*, Phys. Rev. B **41**, 2009 (1990).
13. I. S. Lyubutin, S. T. Lin, C. M. Lin, *et al.*, Physica C (Amsterdam) **248**, 235 (1995).
14. S. Suharan, J. Chadwick, D. B. Hannon, *et al.*, Solid State Commun. **70**, 817 (1989).
15. M. G. Smith, R. D. Taylor, and H. Oesterreicher, Phys. Rev. B **42**, 4202 (1990).
16. M. G. Smith, H. Oesterreicher, M. P. Maley, and R. D. Taylor, Physica C (Amsterdam) **204**, 130 (1992).
17. H. Oesterreicher, M. G. Smith, and R. D. Taylor, J. Magn. Magn. Mater. **104–107**, 497 (1992).
18. I. S. Lyubutin, S. T. Lin, C. M. Lin, *et al.*, Physica C (Amsterdam) **248**, 222 (1995).
19. M. E. Lines and M. Eibschutz, Physica C (Amsterdam) **166**, 235 (1990).
20. F. Hartmann-Boutron, C. Meyer, Y. Gros, *et al.*, Hyperfine Interact. **55**, 1293 (1990).
21. I. S. Lyubutin and T. V. Dmitrieva, Zh. Éksp. Teor. Fiz. **105**, 954 (1994) [JETP **78**, 511 (1994)].
22. A. M. Balagurov, G. M. Mironova, I. S. Lyubutin, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **3**, 615 (1990).
23. I. S. Lyubutin, V. G. Terziev, E. M. Smirnovskaya, and A. Ya. Shapiro, Physica C (Amsterdam) **169**, 361 (1990).
24. I. S. Lyubutin, V. G. Terziev, T. V. Dmitrieva, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **5**, 1423 (1992).
25. I. S. Lyubutin, V. G. Terziev, S. V. Luchko, *et al.*, Physica C (Amsterdam) **199**, 296 (1992).
26. I. S. Lyubutin, V. G. Terziev, T. V. Dmitrieva, *et al.*, Solid State Commun. **86**, 651 (1993).
27. P. Boolchand and D. McDaniel, Hyperfine Interact. **72**, 125 (1992).
28. I. S. Lyubutin, V. G. Terziev, T. V. Dmitrieva, *et al.*, Physica C (Amsterdam) **195**, 383 (1992).
29. M. Peng, C. W. Kimball, and B. D. Dunlap, Physica C (Amsterdam) **169**, 23 (1990).
30. J. Dengler and G. Ritter, Hyperfine Interact. **93**, 1653 (1994).
31. T. R. S. Prasanna, R. C. O'Handley, G. Kalongji, and G. C. Papaefthymiou, Phys. Rev. B **47**, 3374 (1993).
32. Q. A. Pankhurst, S. Suharan, and M. F. Thomas, J. Phys.: Condens. Matter **4**, 3551 (1992).
33. I. S. Lyubutin, V. G. Terziev, S. V. Luchko, *et al.*, Sverkhprovodimost: Fiz., Khim., Tekh. **5**, 1842 (1992).
34. C. W. Kimball, B. Dabrowski, Y. Liang, *et al.*, Hyperfine Interact. **72**, 153 (1992).
35. A. J. Freeman and R. E. Watson, in *Magnetism*, Ed. by G. T. Rado and H. Suhl (Academic, New York, 1965), Vol. 2B, p. 167.
36. F. Menil, J. Phys. Chem. Solids **46**, 763 (1985).
37. I. S. Lyubutin, Nucl. Instrum. Methods Phys. Res. B **76**, 276 (1993).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 615–624.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 700–711.
Original English Text Copyright © 2001 by Menushenkov, Klementev, Kuznetsov, Kagan.

SOLIDS
Electronic Properties

Superconductivity in $Ba_{1-x}K_xBiO_3$: Possible Scenario of Spatially Separated Fermi–Bose Mixture[¶]

A. P. Menushenkov[a], K. V. Klementev[a], A. V. Kuznetsov[a], and M. Yu. Kagan[b]

[a]*Moscow State Engineering Physics Institute, Moscow, 115409 Russia*
[b]*Kapitza Institute for Physical Problems, Moscow, 117334 Russia*

Received March 16, 2001

Abstract—We propose a new scenario for the metal–insulator phase transition and superconductivity in the perovskite-like bismuthates $Ba_{1-x}K_xBiO_3$ (BKBO) based on our EXAFS studies. We show that two types of charge carriers, the local pairs (real-space bosons) and the itinerant electrons, exist in the metallic compound $Ba_{1-x}K_xBiO_3$ ($x \geq 0.37$). The real-space bosons are responsible for the charge transport in semiconducting $BaBiO_3$ and for superconductivity in the metallic BKBO. The appearance of the Fermi liquid state as the percolation threshold is overcome ($x \geq 0.37$) explains the observed metal–insulator phase transition. Because bosons and fermions occupy different types of the octahedral BiO_6 complexes, they are separated in real space, and therefore, the spatially separated Fermi–Bose mixture of a new type is likely to be realized in the bismuthates. The nature of superconductivity is consistently explained in the framework of this scenario. A new superconducting oxide $Ba_{1-x}La_xPbO_3$ has been successfully synthesized to check our conclusions. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The concept of the extremely strong-coupling superconductivity with preexisted local pairs was first introduced by Shafroth [1] in the middle of the 1950s. His statement was that in the extremely type-II superconductors, where $\xi_0 k_F \lesssim 1$, the nature of the superconducting transition corresponds to the local pair formation (pairing in the real, rather than in the momentum space) at some relatively high temperature T^* and to the Bose–Einstein condensation (BEC) of pairs at a lower critical temperature $T_c < T^*$. Later on, Alexandrov and Ranninger [2] developed this concept for narrowband materials with an extremely strong electron–phonon coupling constant ($\lambda \gg 1$), where the standard Eliashberg theory [3] becomes inadequate. The key issue of their approach was the statement that in narrow bands, where the polaron formation is important, it is possible, in principle, to create the conditions for two polarons to effectively occupy the same potential well prepared within a self-consistent procedure.

At approximately the same time, Leggett and Nozieres [4, 5] developed a general theory that described a smooth interpolation between the BCS-type pairing in the momentum space for a weak electron–electron attraction and the pairing in the real space for a strong electron–electron attraction. There were two crucial points in their papers [4, 5]. (i) The results are valid independently of the actual nature of the short-range effective attraction between electrons; (ii) they have self-consistently investigated the standard equation for the superconducting gap and the equation for the conservation of the number of particles. The most important result of Nozieres and Leggett is that for $T^* > T_c$ (in other words, for the binding energy of a local pair $|E_b| > \varepsilon_F$), one has a normal bosonic metal that is responsible for the electron transport. Later on, Ranninger [6, 7] introduced the concept of the two-band Fermi–Bose mixture. His scenario involves the contributions of fermionic and bosonic quasiparticles to the electron transport for $T_c < T < T^*$.

Shortly after the discovery of high-T_c superconductors, Anderson [8] reintroduced the concept of local pairs. He also used two bands of the fermionic and bosonic quasiparticles. In his approach, the superconducting transition was related to the BEC in the bosonic band of charge excitations, the holons, while the presence of a large Fermi surface was guaranteed by the fermionic band of spin excitations, the spinons. Unfortunately, even this beautiful approach was not totally successful because, at least in one layer, the BEC of holons yields a superconducting pair with the charge *e* instead of 2*e* measured experimentally.

Later on, Geshkenbein, Ioffe, and Larkin [9] phenomenologically reconsidered the model of the Fermi–Bose mixture at the level of the Ginzburg–Landau expansion coefficients and showed that several important experiments in the underdoped high-T_c materials can be naturally explained within this form of the Ginzburg–Landau functional. At present, therefore, the question of whether a two-band Fermi–Bose mixture scenario is applicable to high-temperature superconducting (HTSC) materials remains open. The materials that can be best described

---

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

by this scenario are the bismuth family of high-T_c superconductors $Bi_2Sr_2CaCu_2O_{8+\delta}$, where $k_F\xi_0 \approx 2$, and the tunneling experiments of Rener *et al.* [10] signal the formation of a sufficiently large and stable pseudogap at temperatures well above T_c.

In this paper, we discuss the possibility of a two-band Fermi–Bose mixture scenario in an entirely different class of superconductors with a relatively high critical temperature $T_c \sim 30$ K, namely, in the barium bismuthates $Ba_{1-x}K_xBiO_3$. The key issue of our paper is the possibility of the existence of spatially separated subsystems of fermionic and bosonic quasiparticles in these materials. The cubic perovskites $BaPb_{1-x}Bi_xO_3$ (BPBO) and $Ba_{1-x}K_xBiO_3$ (BKBO) have been known and extensively studied since 1975 and 1988, respectively. A large number of the first-principle calculations have been carried out [11–17] to explain the high-temperature superconductivity in these compounds. However, most of the unusual properties of the $BaBiO_3$-family compounds mentioned in the early review by Uchida *et al.* [18] still remain unexplained. From our point of view, the reason is that these calculations were done in the mean-field approximation, and therefore, they adequately describe crystal structures with translational symmetry but cannot explain the peculiarities of the electron properties of BPBO and BKBO following from the local structure distortions. Such distortions, which were observed by the local sensitive methods such as the XPS [19], Raman scattering [20], and EXAFS [21–24], destroy the translation symmetry. Therefore, the mean-field approximation cannot describe the local electron structure.

Based on the low-temperature EXAFS study of the BKBO and BPBO compounds, the motion of the local electron pairs correlated with the lattice vibrations was established in our recent work [22]. We found that different electron fillings of the upper antibonding Bi6*s*-O2$p_{\sigma*}$ orbital of the BiO_6 octahedra lead to the formation of a double-well potential for the vibration of some part of the oxygen ions. The observed anomalies have been discussed in the framework of different theoretical models of superconductivity. We found that none of them can fully explain our experimental results and proposed a phenomenological model of the relationship between the local crystal and the local electron structures.

We present here a relatively new concept of superconductivity in bismuthates based on an interplay between spatially separated Fermi and Bose subsystems. In contrast to the previous theoretical models and calculations, we take the results obtained in a number of local sensitive measurements into account [19–24].

The paper is organized as follows. In Section 2, we present the basic experimental facts pertaining to the local electron and crystal structure peculiarities and consider their relation to the superconducting and the normal transport properties of BKBO. We then show how these basic facts can naturally be explained within the scenario of two spatially separated bands of the fermionic and bosonic quasiparticles. In Section 3, consequently, we discuss the nature of superconductivity in the bismuthates within this scenario. We conclude the paper by summarizing our model and discussing several additional experiments that would help give a definite answer as to whether our proposal is the only possibility for a superconducting pairing in bismuthates.

2. SPATIALLY SEPARATED FERMI–BOSE MIXTURE

$BaBiO_3$, which is a parent compound for the bismuthates $Ba_{1-x}K_xBiO_3$ and $BaPb_{1-x}Bi_xO_3$, represents a charge-density-wave (CDW) insulator having two gaps: an optical gap with $E_g = 1.9$ eV and an activation (transport) gap with $E_a = 0.24$ eV [18]. A partial replacement of Ba by K in BKBO causes the decrease if the gaps, and as a result, the insulator–metal transition and superconductivity are observed at the doping levels $x \gtrsim 0.37$. The superconductivity remains up to the doping level $x \approx 0.5$ corresponding to the solubility limit of K in BKBO, but the maximum critical temperature $T_c \approx 30$ K is achieved for $x \approx 0.4$ [25, 26].

A. Local Crystal Structure Peculiarities

A three-dimensional nature of the cubic perovskite-like structure of the bismuthates differs from the two-dimensional one in the HTSC cuprates. The building block in the bismuthates is the BiO_6 octahedral complex (the analogue of CuO_n ($n = 4, 5, 6$) in HTSC materials). The octahedral complexes are the most tightly bound items of the structure because of a strong covalence of the Bi6*s*–O2p_σ bonds. According to the crystallographic data [27], the crystal structure of a parent $BaBiO_3$ compound represents the alternating arrangement of the expanded and contracted BiO_6 octahedra (referred to as the “breathing” distortion) in the barium lattice. This alternation and the static rotation of the octahedra around the axis [110] produce a monoclinic distortion of the cubic lattice. As shown in [21, 22, 28], the larger soft octahedron corresponds to the BiO_6 complex with the completely filled Bi6*s*–O2*p* orbitals and the smaller rigid octahedron corresponds to the $Bi\underline{L}^2O_6$ complex. Here, $\underline{L}^2$ denotes the free level in the antibonding Bi6*s*–O2$p_{\sigma*}$ orbital of the smaller octahedral complex.

The K doping of $BaBiO_3$ is equivalent to the hole doping and leads to a partial replacement of the larger soft octahedra BiO_6 by the smaller rigid octahedra $Bi\underline{L}^2O_6$ [22]. This causes the decrease and the eventual disappearance of the static breathing and tilting distortions; the lattice must therefore contract despite the practically equal ionic radii of K^+ and Ba^{2+}. As a result, the average structure becomes a simple cubic one at the doping level $x = 0.37$ in accordance with the neutron diffraction data [29]. However, the local EXAFS probes [21, 22, 24] showed a significant difference of the local crystal structure from the average one. We found that the oxygen ions belonging to the different BiO_6 and $Bi\underline{L}^2O_6$ octahedra vibrate in a double-well potential, while those having an equivalent environment of the two equal $Bi\underline{L}^2O_6$ octahedra oscillate in a simple harmonic potential [21, 22]. This very unusual behavior is closely related to the local electron structure of BKBO.

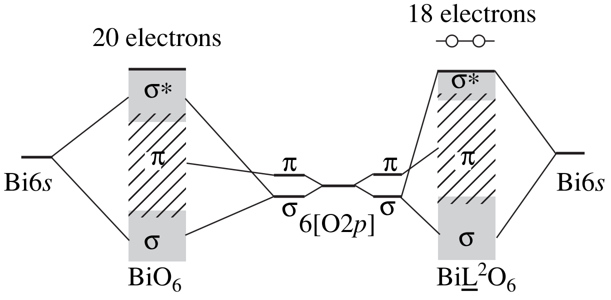


Fig. 1. The scheme of the electron structure formation for the different octahedral BiO_6 complexes.

B. Local Electron Structure

The coexistence of the two types of the octahedra in $BaBiO_3$ with different Bi–O bond lengths and strengths reflects the different electron structures of BiO_6 complexes. The valence band of $BaBiO_3$ is determined by the overlap of the Bi6*s* and O2*p* orbitals [11, 30], and because of a strong Bi6*s*–O2p_σ hybridization, the octahedra can be considered as quasimolecular complexes [31]. Each complex involves ten electron levels consisting of four bonding–antibonding Bi6*s*–O2p_σ orbitals and six nonbonding O2p_π orbitals. A monoclinic unit cell includes two octahedra and contains 38 valence electrons (10 from two bismuth ions, 4 from two barium ions, and 24 from six oxygen ions). All the Bi–O bond lengths must be equal and local magnetic moments must be present for the equal electron filling of the nearest octahedra ($Bi\underline{L}^1O_6 + Bi\underline{L}^1O_6$). However, the presence of two types of octahedral complexes and the absence of any local magnetic moment were observed experimentally [18, 32]. A scheme of the valence disproportionation $2Bi\underline{L}^1O_6 \longrightarrow Bi\underline{L}^2O_6 + BiO_6$ was then proposed [22] in which the numbers of occupied states are different in the neighboring octahedral complexes: the octahedron $Bi\underline{L}^2O_6$ contains 18 electrons and has one free level or a hole pair $\underline{L}^2$ in the upper antibonding Bi6*s*–O2$p_{\sigma*}$ orbital, while, in the octahedron BiO_6 with 20 electrons, the antibonding orbital is filled as shown in Fig. 1. It is quite natural that the $Bi\underline{L}^2O_6$ octahedra have stiff quasimolecular Bi–O bonds and a smaller radius, while the BiO_6 octahedra represent unstable molecules with the filled antibonding orbital and a larger radius.

Thus, in $BaBiO_3$, one has an alternating system of the two types of octahedral complexes filled with local pairs: the hole pairs in $Bi\underline{L}^2O_6$ complexes and the electron pairs in BiO_6 complexes. In other words, the parent compound is a system with the real-space [33] or hard-core [16] bosons (i.e., with at most one boson per site).

The local pair formation in $BaBiO_3$ was advocated previously, e.g., in [15, 16, 18, 34–38]. The binding mechanism for the pairs is probably of an electron nature [16, 35] in accordance with Varma's picture of the pairing due to the skipping of the valence "4+" by the Bi ion [35]. However, one cannot fully exclude the lattice mediated pairing [18, 34, 37] in accordance with de Jongh's statement [36] that the preference to retain the close-shell structures overcomes the Coulomb repulsion related to the intrasite bipolaron formation.

The local electron structure of $BaBiO_3$ combined with the real-space local crystal structure is presented in Fig. 2a. There are no free fermionic carriers in this system, and the conductivity is only due to the transfer of the carrier pairs [18, 28]. Experimentally, $BaBiO_3$ shows a semiconducting-type behavior with the energy gap $E_a = 0.24$ eV. This transport gap does not show up in photoconductivity, optical absorption, or photoacoustic measurements [39] and can be explained only as a two-particle transport with the activation energy $2E_a$ due to the delocalization of pairs. From our point of view, the transport gap value is defined by the combined effect of the intersite Coulomb repulsion and the local lattice deformation due to the static breathing distortion.

The optical gap, in similarity with the Varma [35] suggestion, costs the energy

$$\begin{aligned} E_g \;=\; E_b + 2E_a &= 2E(\mathrm{Bi\underline{L}^1O_6}) \\ &- [E(\mathrm{BiO_6}) + E(\mathrm{Bi\underline{L}^2O_6})] \end{aligned} \tag{1}$$

and is observed experimentally as the optical conductivity peak at the photon energy $h\nu = 1.9$ eV [18]. (Here, E_b is the pair binding energy related to the dissociation of pairs.)

In accordance with Eq. (1), the optical excitation must produce a local lattice deformation via the transformation of the two different octahedra into equivalent ones: $Bi\underline{L}^2O_6 + BiO_6 \xrightarrow{h\nu} 2Bi\underline{L}^1O_6$. Consequently, this dynamical local lattice deformation is manifested in the Raman spectra as an abnormally large amplitude of breathing-type vibrations of the oxygen octahedra if the resonance laser excitation with $h\nu = E_g$ is used [31, 37, 40]. The abrupt decrease of the mode amplitude was observed when lasers with different wavelengths were used [40].

It is important to emphasize that there are no free fermions in the system. Only the excited fermions can

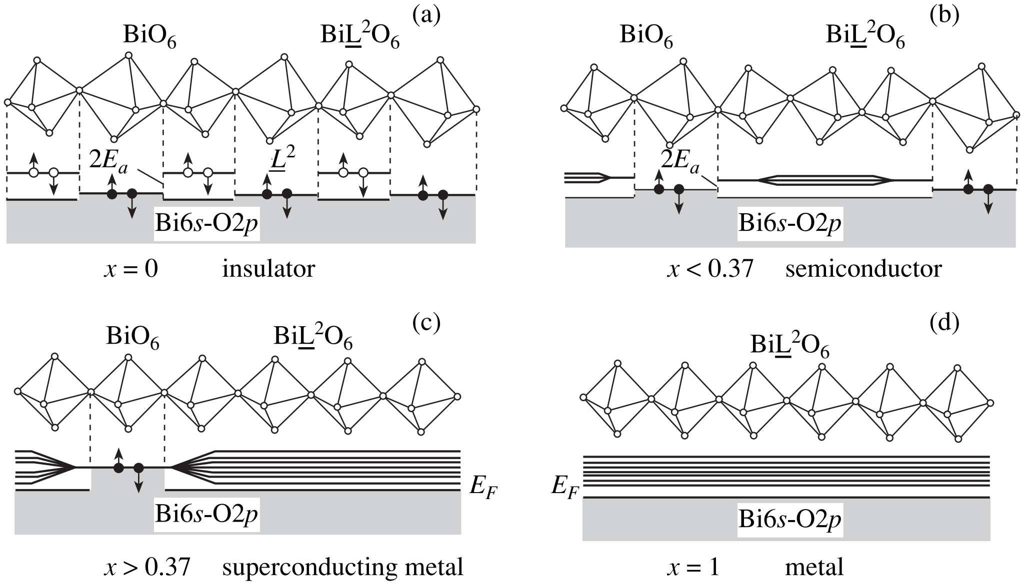


Fig. 2. The scheme of the insulator–metal phase transition for the K doping of $Ba_{1-x}K_xBiO_3$ in the framework of the relationship between the local crystal and the local electron structures. The local crystal structure of the octahedral complexes (at the top) and the local electron structure (at the bottom) are shown in pictures (a)–(d). The occupied states of the Bi6*s*–O2*p* valence band are marked by gray. $2E_a$ is the activation gap. Black and white circles with arrows denote, respectively, the electrons and the holes with the opposite spin orientations. (a) The monoclinic phase of the insulator $BaBiO_3$ compound. (b) The orthorhombic phase of a semiconducting BKBO at $0 < x < 0.37$. The splitting of the free level $\underline{L}^2$ at the spatial overlap of the $Bi\underline{L}^2O_6$ octahedra is sketched. (c) The undistorted cubic phase of a superconducting metal at $x > 0.37$. The formation of a Fermi iquid state is shown arising due to the overlap of the unoccupied fermionic band F with the occupied Bi6*s*–O2*p* valence band as the percolation threshold is reached. (d) The undistorted cubic phase of a nonsuperconducting metal at $x = 1$. A Fermi liquid state with the Fermi level E_F is shown.

be produced by the unpairing, and they do not give any contribution to the charge transport because of a high value of E_b. The bosonic and the fermionic subsystems are therefore separated both spatially and energetically, and hence, the Fermi–Bose mixture is absent in the parent compound.

This situation is illustrated in Fig. 3a, where we schematically present the one-particle density of states. For $x = 0$, the filled bosonic band is separated from the empty fermionic band (the excited band F') by the large optical gap E_g and from the empty bosonic band B by the smaller transport gap $2E_a$. The bosonic band plays the role of a conduction band for bosonic quasiparticles involved in the activation transport. In accordance with [41], the filled and empty bosonic bands have, respectively, the hole-like and electron-like dispersions in the representation of the one-particle density of states. Because bosons and fermions are always spatially separated (i.e., belong to different octahedra complexes), we show their densities of states in the different sides of the pictures.

C. Formation of the Fermi–Bose Mixture

The substitution of the two K^+ ions for two Ba^{2+} ones modifies the BiO_6 complex to the $Bi\underline{L}^2O_6$ one. As a result, the number of small stiff $Bi\underline{L}^2O_6$ octahedra increases as $n_0(1 + x)/2$ and the number of large soft BiO_6 octahedra decreases as $n_0(1 - x)/2$, where $n_0 = 1/a^3$ is the number of unit cells and a is the lattice parameter. Clusters of the $Bi\underline{L}^2O_6$ complexes are formed with doping, which contracts the lattice because of small radii and the rigid bonds of the $Bi\underline{L}^2O_6$ complexes.

The changes in the crystal structure are accompanied by essential changes in the local electron structure. A spatial overlap of the $\underline{L}^2$ levels leads to their splitting into an empty fermionic-like band F inside the $Bi\underline{L}^2O_6$– … –$Bi\underline{L}^2O_6$ Fermi cluster (see Fig. 2b). In the doping range $x < 0.37$, this sufficiently narrow band is still separated from the occupied Bi6*s*–O2*p* subband. The number of empty electron states in the F band increases with x as $\hat{n}_F = n_0(1 + x)$, while the number of local electron pairs decreases as $n_B = n_0(1 - x)/2$.

The free motion of the pairs is still prevented by the intersite Coulomb repulsion ($E_a \neq 0$), which is screened inside the clusters, however. When the Fermi clusters are formed, the conductivity occurs because of the motion of the pairs through the clusters of different lengths. The BKBO compounds demonstrate a semiconducting-type conductivity changing from a simple activation type to Mott's law with variable-range hopping [42].

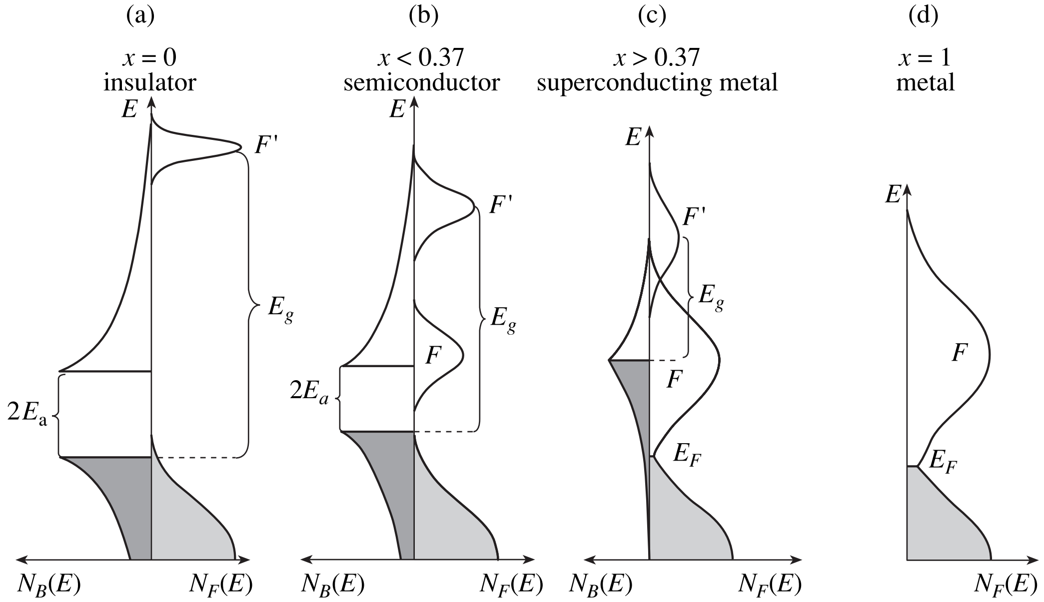


Fig. 3. A sketch of the one-particle density of states for $Ba_{1-x}K_xBiO_3$. The contributions of the bosons $N_B(E)$ and the fermions $N_F(E)$ are depicted separately because bosonic and fermionic states are spatially separated. The filled (dark gray) and the unoccupied (transparent) bosonic bands correspond, respectively, to the contributions of the electron and the hole pairs. The bands are separated by the activation gap $2E_a$, which is lowered with the doping level x. An empty fermionic band F' corresponding to the destruction of pairs is separated from the occupied bosonic band by the optical gap E_g. An empty fermionic band F is formed due to the splitting of the free level $\underline{L}^2$, which arises from the spatial overlap of the $Bi\underline{L}^2O_6$ octahedra. The filled fermionic band (gray) represents the Bi6s–O2p valence band. The band F' and bosonic bands decrease as the doping increases, because the number of the electron pairs decreases, while the band F grows due to the increase in the number of free levels. A Fermi liquid state is formed (c, d) as a result of the overlap between the F band and the Bi6s–O2p valence band.

At the doping level $x \approx 0.37$ (see Fig. 2c and Fig. 3c), the following cardinal changes occur.

(i) The breathing and the rotational static lattice distortions transform to the dynamic ones. At the Bose and the Fermi cluster borders, where all the oxygen ions belong to both the BiO_6 and the $Bi\underline{L}^2O_6$ octahedra, the local breathing dynamic distortion is observed as the oxygen ion vibration in a double-well potential [21, 22].

(ii) The infinite three-dimensional percolating Fermi cluster is formed from the $Bi\underline{L}^2O_6$ octahedra along the [100]-type directions. The empty fermionic band overlaps the filled one, and F therefore becomes a conduction band. Overcoming the percolation threshold provides the insulator-metal phase transition and the formation of the Fermi liquid state for $x \geq 0.37$. The valence electrons previously localized in the $Bi\underline{L}^2O_6$ complexes become itinerant inside the infinite Fermi cluster.

(iii) The pair localization energy disappears, $E_a \approx 0$, and therefore, local electron pairs originating from the BiO_6 complexes can move freely providing a bosonic contribution to the conductivity. In the metallic phase, two types of carriers are present: the itinerant electrons from the $Bi\underline{L}^2O_6$ complexes (fermions) and the delocalized electron pairs from the BiO_6 complexes (bosons). Although the normal state conductivity is mainly due to the fermionic subsystem, the contribution of the bosonic subsystem was also observed by Hellman and Hartford [43] as the two-particle normal state tunneling.

As a result, at doping levels $x \geq 0.37$, we have a new type of a spatially separated mixture of the bosonic and fermionic subsystems describing both metallic and superconducting properties of BKBO. We stress that because fermions and bosons belong to the complexes with different electron structures, the Fermi and Bose subsystems are spatially separated at any doping level. These subsystems are related by

$$2n_B + \hat{n}_F = 2n_0 \tag{2}$$

and

$$\frac{2n_B}{\hat{n}_F} = \frac{1-x}{1+x}. \tag{3}$$

The pair destruction is prevented by a sufficiently high value of the binding energy, which becomes apparent as the pseudogap $E_b = E_g \approx 0.5$ eV [44] in superconducting compositions. The unpairing is possible only under the optical excitation to the band F' (see Fig. 3c), which does not play any role in the charge transport.

At $x = 1$, all the BiO_6 octahedra are transformed into the $\mathrm{Bi}\underline{L}^2\mathrm{O}_6$ ones. The Bose system disappears ($n_B = 0$) together with the excited fermionic band F'. Therefore, KBiO_3 must be a nonsuperconducting Fermi liquid metal (see Fig. 2d and Fig. 3d).

We note that a metallic KBiO_3 compound exists only hypothetically because the potassium solubility limit $x \approx 0.5$ is exceeded in BKBO. However, BaPbO_3, which can be viewed as an electron analogue of KBiO_3, demonstrates metallic but not superconducting properties. Recent attempts to synthesize KBiO_3 at a high pressure have shown that only $\mathrm{K}_{1-y}\mathrm{Bi}_y\mathrm{BiO}_3$ with a partial replacement of the K^+ ions by the Bi^{3+} ones is formed [45]. This replacement must lead to the appearance of the BiO_6 octahedra with the local electron pairs, and the compound $\mathrm{K}_{1-y}\mathrm{Bi}_y\mathrm{BiO}_3$ must therefore be superconducting in accordance with the above discussion. Indeed, superconductivity with $T_c = 10.2$ K was experimentally observed in this compound [45].

Our analysis implies that BaPbO_3 must be superconducting at a partial substitution of the Ba^{2+} ions for the trivalent ions because this substitution produces local electron pairs as in the case of $\mathrm{K}_{1-y}\mathrm{Bi}_y\mathrm{BiO}_3$. Thus, using the La^{3+} doping, it is possible to obtain the spatially separated Fermi–Bose mixture in BaPbO_3. Recently, we have successfully produced $\mathrm{Ba}_{1-x}\mathrm{La}_x\mathrm{PbO}_3$ using the high-pressure synthesis technique. Superconductivity at $T_c = 11$ K observed in this new compound [46] is a direct evidence in favor of the above scenario.

At the end of this section, we note that our understanding of the insulating state in the parent BaBiO_3 is very similar to the theoretical model by Taraphder *et al.* [15, 16]. We agree with the authors of [15, 16] on the following principal positions: (i) the presence of the electron-mediated (Varma's type) pairing mechanism; (ii) the existence of the charge $\pm 2e$ bosonic bound states that dominate transport properties of BaBiO_3; (iii) the explanation of the nature of both the transport and the optical gaps.

However, our description of the K-doped systems strongly differs from their model. Going from insulating BaBiO_3 to superconducting BKBO ($x \geq 0.37$), Taraphder *et al.* were forced to change the nature of the pairing mechanism from the real-space pairing to the k-space one. Thus their description of the superconducting state does not differ from the traditional BCS description that has been discussed for BKBO, e.g., in [12, 13, 47].

Using our EXAFS results [21, 22], we consistently explain the insulating and the superconducting states in BKBO within a single approach. In contrast to [15, 16], we showed that the real-space bosons do not disappear in the metalic region of BKBO and that they are responsible for superconductivity. At the same time, the Fermi liquid appears in the BKBO system because of the overlapping of the occupied valence band levels and the free ones when the percolation threshold $x = 0.37$ is overcome. An interplay of these Bose and Fermi subsystems explains the main properties of BKBO.

3. DISCUSSION

Taking the existence of the double-well potential in $\mathrm{Ba}_{1-x}\mathrm{K}_x\mathrm{BiO}_3$ into account, one can consider superconductivity in this compound in the framework of the anharmonic models for HTSC [48, 49]. As shown in these models, if the oxygen ions move in a double-well potential, an order-of-magnitude enhancement of the electron-lattice coupling constant follows automatically from a consistent treatment of this motion.

However, as shown above, the double-well potential arises in the bismuthates from different electron fillings of the nearest octahedra and the tunneling of local pairs between them. The existence of the double-well potential in the metallic phase of BKBO ($x \geq 0.37$) therefore indicates that the real-space bosons do not decay with doping. There are at least two additional experimental confirmations of this fact: (i) the observation of the optical pseudogap in superconducting composition [44]; (ii) the existence of two types of charge carriers with heavy and light masses [43, 50]. These experimental facts allow us to consider superconductivity in the bismuthates as the motion of local electron pairs. This motion is correlated with the oxygen ion vibrations in the double-well potential and leads to the transformation of the Bose octahedral complexes to the Fermi ones and vice versa in the dynamical exchange process $\mathrm{BiO}_6 \longleftrightarrow \mathrm{Bi}\underline{L}^2\mathrm{O}_6$. The interplay between the Bose and the Fermi subsystems is closely related to the superconductivity analyzed below in more detail.

A. The Fermi–Bose Mixture Hamiltonian

We first consider the Hamiltonian of the Fermi–Bose mixture (see, e.g., [51])

$$H = H_B + H_F + H_{FB} \tag{4}$$

in the spatially separated case,

$$\begin{aligned} H_B &= \int d^3r\Phi^*(\mathbf{r})\left[-\frac{\hbar^2}{2m_B}\nabla^2\right]\Phi(\mathbf{r}) \\ &+ \frac{1}{2}U_{BB}\int d^3r d^3r'\Phi(\mathbf{r})\Phi^*(\mathbf{r}')\delta(\mathbf{r}-\mathbf{r}')\Phi(\mathbf{r}')\Phi(\mathbf{r}), \\ H_F &= \int d^3R\Psi^*_\alpha(\mathbf{R})\left[-\frac{\hbar^2}{2m_F}\nabla^2\right]\Psi_\alpha(\mathbf{R}) \\ &+ \frac{1}{2}U_{FF}\int d^3Rd^3R'\Psi^*_\alpha(\mathbf{R})\Psi^*_\beta(\mathbf{R}') \\ &\times\delta(\mathbf{R}-\mathbf{R}')\Psi_\beta(\mathbf{R}')\Psi_\alpha(\mathbf{R}), \\ H_{FB} &= \int d^3Rd^3r\Phi^*(\mathbf{r})\Psi^*_\alpha\mathbf{R} \\ &\times U_{FB}(\mathbf{r}-\mathbf{R})\Phi(\mathbf{r})\Psi_\alpha(\mathbf{R}), \end{aligned} \tag{5}$$

where Φ and Ψ_α are bosonic and fermionic operators, and m_B and m_F are the effective masses of bosons and fermions. The most important property of Eq. (5) is that fermions and bosons are spatially separated, and moreover, the potential U_{FB} has a double-well shape at the boundary between the bosonic and fermionic subsystems. This fact is crucial for matching our model with experimental data on the local structure and for the explanation of superconductivity. Obviously, realistic Hamiltonians are more complicated, and the quantities U_{BB}, U_{FF}, and U_{FB} include not only the direct interactions between bosons and fermions, but also indirect interactions via the lattice. The last contribution could be dynamical in principle, and the picture could therefore change in time, thereby leading to a dynamical version of the real-space separation. We note that the bosonic Hamiltonian H_B in Eq. (4) is generated from the electron band with a strong attraction between electrons (see [15, 16, 41]). We thus have two generic spatially separated electron bands: one with a strong attraction between electrons and the other with a repulsion.

We must solve Hamiltonian (4) together with the equation for the particle number conservation obtained from Eqs. (2),

$$2n = \frac{2N}{\Omega} = \frac{1}{\Omega}\int 2n_B(\mathbf{r})d^3r + \frac{1}{\Omega}\int \hat{n}_F(\mathbf{R})d^3R, \qquad (6)$$

where

$$n_B(\mathbf{r}) = \langle \Phi^*(\mathbf{r})\Phi(\mathbf{r})\rangle,$$
$$\hat{n}_F(\mathbf{R}) = \sum_\alpha \langle \Psi^*_\alpha(\mathbf{R})\Psi_\alpha(\mathbf{R})\rangle, \qquad (7)$$

and Ω is the volume of the system. In accordance with Eqs. (3), we also have $2n_B/\hat{n}_F = (1-x)/(1+x)$.

We consider the system of equations (4)–(7) in the Hartree–Fock approximation. By analogy with the Fermi–Bose mixture ^{7}Li–^{6}Li in a magnetic trap [51], it is convenient to introduce the effective external potentials for fermions $U_{\mathrm{ext}}(\mathbf{R})$ and bosons $U_{\mathrm{ext}}(\mathbf{r})$ as

$$U_{\mathrm{ext}}(\mathbf{r}) = \frac{1}{\Omega}\int U_{FB}(\mathbf{r}-\mathbf{R})\hat{n}_F(\mathbf{R})d^3R, \qquad (8)$$

$$U_{\mathrm{ext}}(\mathbf{R}) = \frac{1}{\Omega}\int U_{FB}(\mathbf{r}-\mathbf{R})n_B(\mathbf{r})d^3r. \qquad (9)$$

Following [52], we then obtain

$$H_{HF} = \tilde{H}_B + \tilde{H}_F,$$
$$\tilde{H}_B = \int d^3r\Phi^*(\mathbf{r})\left[-\frac{\hbar^2}{2m_B}\frac{\partial^2}{\partial\mathbf{r}^2} + U_{BB}(\mathbf{r})n_B(\mathbf{r}) + U_{\mathrm{ext}}(\mathbf{r})\right]\Phi(\mathbf{r}), \qquad (10)$$

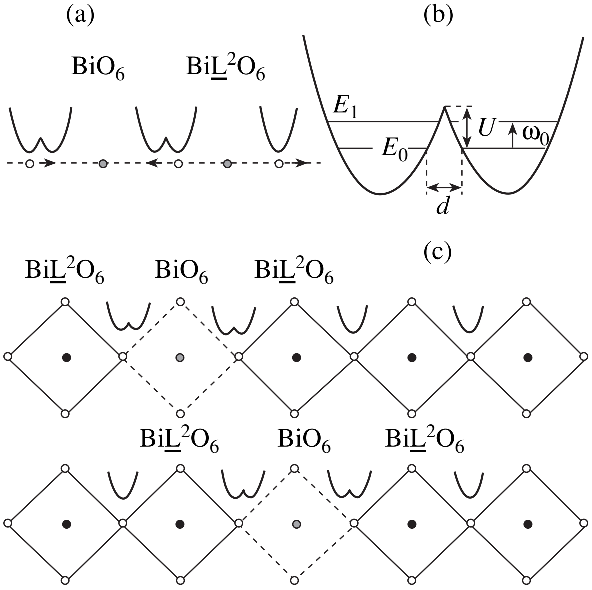


Fig. 4. A sketch of the dynamic exchange $BiO_6 \longleftrightarrow Bi\underline{L}^2O_6$ is shown in the BiO_2 plane of the octahedra. (a) The breathing mode of the vibrations along the [100]-type direction of two neighboring octahedra with different electron structures. The BiO_6 octahedron transforms to the $Bi\underline{L}^2O_6$ one and vice versa due to the electron pair tunneling between the octahedra. An oxygen ion belonging to these octahedra oscillates in the double-well potential. An oxygen ion belonging to the equivalent neighboring $Bi\underline{L}^2O_6$ octahedra oscillates in a simple parabolic potential. (b) The double-well potential with the energy levels for the vibration of the oxygen ion. The following parameters describe the tunneling barrier between the wells in $Ba_{0.6}K_{0.4}BiO_3$ at low temperatures [22]: the tunneling frequency $\omega_0 \approx 200$ K, the barrier height $U \approx 500$ K, the barrier width $d \approx 0.07$ Å. (c) The motion of a local electron pair centered on the BiO_6 octahedron through the $Bi\underline{L}^2O_6 \ldots Bi\underline{L}^2O_6$ Fermi cluster. For detailed explanations, see the text.

$$\tilde{H}_F = \int d^3R\Psi^*_\alpha(\mathbf{R})\left[-\frac{\hbar^2}{2m_F}\frac{\partial^2}{\partial\mathbf{R}^2} + U_{FF}(\mathbf{R})\hat{n}_F(\mathbf{R}) + U_{\mathrm{ext}}(\mathbf{R})\right]\Psi_\alpha(\mathbf{R}). \qquad (11)$$

Equations (10) and (11) self-consistently describe the interaction between the Fermi and Bose subsystems. However, in contrast to the Fermi–Bose mixture models discussed previously (see Section 1), the relation between the concentration of the bosonic and fermionic quasiparticles is independent of the temperature and depends only on the doping level x in accordance with Eq. (3).

The theoretical analysis of the above equations is beyond the scope of this paper and will be published elsewhere. We only note that the potential $U_{\mathrm{ext}}(\mathbf{r})$ in (10), which is the contraction of the Fermi–Bose interaction potential $U_{FB}(\mathbf{r}-\mathbf{R})$ with the fermionic density

$\hat{n}_F(\mathbf{R})$, also has a double-well shape, and the real-space bosons therefore live in the effective double-well external potential. The maximum of this external potential corresponds to the point $\mathbf{r} = \mathbf{R}$ at the boundary between the Fermi and the Bose clusters. Our estimate for the superconductivity critical temperature is heavily based on this fact.

B. Superconductivity in $Ba_{1-x}K_xBiO_3$

Taking the existence of the double-well potential in $Ba_{1-x}K_xBiO_3$ into account, one can consider superconductivity in this compound as a long-range order that is established via the local pair tunneling from one Bose cluster to a nearest one over the Fermi cluster along [100]-type directions.

The pair transfer process correlated with the oxygen vibration (in other words, the dynamic exchange) is illustrated in Fig. 4. The oxygen belonging to the two neighboring octahedra BiO_6 and $Bi\underline{L}^2O_6$ vibrates in the double-well potential, and hence, the electron pair tunneling between the neighboring octahedra occurs when the ion tunnels through the potential barriers between the wells. Because of this interconnection between the pair and the oxygen tunneling processes, we can estimate the matrix element of the pair tunneling as $t_B \propto \omega_0 e^{-D}$, where ω_0 is the tunneling frequency,

$$D = \frac{1}{\hbar}\int_{x_0}^{x_1} |p| dx \approx \frac{d}{\hbar}\sqrt{2MU}$$

is the semiclassical transparency of the barrier in the double-well potential, U and d are the barrier height and width, and M is the oxygen ion mass. We note that the relatively small tunneling frequency $\omega_0 = 200$ K (see Fig. 4) already incorporates all the polaronic effects.

A local pair is transferred from one Bose cluster to the nearest one over a Fermi cluster, which, depending on the doping level, consists of several octahedra. The pairs overcome the Fermi cluster step by step. A single step corresponding to the pair transfer into a neighboring octahedron occurs simultaneously with the oxygen ion tunneling in the double-well potential. The tunneling frequency ω_0 is therefore the same for each step. Assuming that the steps are independent events, the probability of overcoming the Fermi cluster can be obtained as the product of the probabilities of each step. The matrix element of the pair tunneling through the Fermi cluster can then be estimated as $\hat{t}_B \propto \omega_0 e^{-\langle N\rangle D}$, where the average number of steps (which is proportional to the Fermi cluster linear size) can be obtained from the ratio of the concentrations of the $Bi\underline{L}^2O_6$ and BiO_6 octahedra. This gives the number of steps

$$\langle N\rangle \approx \left(\frac{1+x}{1-x}\right)^{1/3}.$$

A natural assumption is that the critical temperature of the onset of superconductivity is of the order of the Bose–Einstein condensation temperature $T_c \propto \hat{t}_B a^2 n_B^{2/3}$ in the bosonic system with the large effective mass $m_B \propto 1/\hat{t}_B a^2$. We recall that $a^3 n_B = (1-x)/2$ in our case. For the parameters of the double-well potential obtained in [22] (see also Fig. 4), we estimated $T_c \sim 50$ K in $Ba_{0.6}K_{0.4}BiO_3$, which is larger than the measured $T_c \approx 30$ K.

However, this estimate does not account for the phase coherence arising due to the relation between the vibrations of oxygen ions and the transfer of pairs. When a pair is transferred from one octahedron to another, the lattice has sufficient time to relax due to the longitudinal stretching phonons, each time forming a new configuration before the next tunneling event occurs. Taking the breathing like character of the oxygen ion vibrations in the double-well potential into account (see Fig. 4), it is natural to suppose that the breathing mode of each octahedron is coordinated with its neighbors to guarantee a resonant tunneling along [100]-type axes in the system.

From the dispersion of longitudinal phonon modes studied by the inelastic neutron scattering [53], it follows that the breathing-type vibrations with the wave vector $q_b = (\pi/a, 0)$ are energetically favorable in the superconducting compositions of BKBO. Hence, a long-range correlation of vibrations must occur at low temperatures when only the low-energy states are occupied. The bandwidth of the longitudinal stretching mode is of the order 100 K, and the temperature $T \sim T_c$ is sufficiently high to excite the nonbreathing-type longitudinal stretching phonons with the wave vectors shorter than q_b. The thermal excitation of these phonons leads to the destruction of the long-range correlation between the breathing-type vibrations and hence destructively affects the long-range phase coherency of the local pair transfer.

We note that the anomalous dispersion of the longitudinal stretching phonons observed in [53] reflects the lattice softening with the decrease in temperature due to the existence of a double-well potential in the superconducting compositions of BKBO. A similar dispersion was also observed in the high-T_c cuprates $La_{1.85}Sr_{0.15}CuO_4$ and $YBa_2Cu_3O_7$ [54]. The problem of the T_c limitation due to the phase coherence destruction is now extensively discussed (see [55] for a review). In view of the recent experimental evidence by Müller *et al.* [56] for the coexistence of small bosonic and fermionic charge carriers in $La_{2-x}Sr_xCuO_4$, we also envisage applying our scenario to HTSC cuprates. Because the underdoped HTSC materials are similar to the phase separation on antiferromagnetic (AFM) and paramagnetic (PM) clusters [57–59], we suppose that the lattice here can play an assistant role by providing a pair tunneling between the superconducting PM metallic clusters via the insulating AFM barrier. It can also serve as a limitation for the estimate of the effective critical temperature for the superconducting transition in cuprates.

4. CONCLUSIONS

We briefly summarize the key positions of our conception.

1. The parent compound $BaBiO_3$ represents a system with the initially preformed local electron and hole pairs. Each pair is spatially and energetically localized inside the octahedron volume. The localization energy of a pair determines the transport activation gap E_a. The binding energy of a pair is given by $E_b = E_g - 2E_a$, where $E_g$is the optical gap.

2. The spatially separated Fermi–Bose mixture of a new type is possibly realized in the superconducting compositions $Ba_{1-x}K_xBiO_3$ for $x \geq 0.37$. The bosonic bands are responsible for the two-particle normal state conductivity. The overlap of the empty fermionic band F with the occupied valence band Bi6*s*–O2*p* provides the insulator–metal phase transition and produces the Fermi liquid state. This state strongly shunts the normal state conductivity arising from the two-particle Bose transport.

3. The fermionic band F' connected with the pair destruction does not play any role in the transport. The excitation energy is sufficiently high to guarantee against the destruction of bosons (the pair binding energy for superconducting compositions is $E_b \approx 0.5$ eV).

4. The pair localization energy is absent for $x \geq 0.37$ ($E_a = 0$), and therefore, the bosonic and the fermionic subsystems are separated only spatially. The interplay between them is due to the dynamic exchange $Bi\underline{L}^2O_6 \longleftrightarrow BiO_6$, which causes the free motion of local pairs in the real space.

5. The pairing mechanism in the bismuthates is more likely of the Varma type (because of skipping the "4+" valence by the Bi ion) rather than of phonon-mediated origin. The existence of local pairs and their tunneling between the neighboring octahedra is closely related to the presence of a double-well potential that describes the vibration of the oxygen ions. The lattice is more likely involved in the superconductivity by providing the phase coherence for the motion of local pairs in the real space.

We finally emphasize that the scenario of the Fermi–Bose mixture allows us to qualitatively describe the insulator–metal phase transition and the superconducting state in BKBO in the framework of a single approach. To some extent, this scenario explains the contradictions between the result of the local sensitive and integral experimental methods [18–24, 31, 40, 53]. In addition, we successfully synthesized a new superconducting oxide $Ba_{1-x}La_xPbO_3$ that can be considered as direct evidence in favor of our model.

Nevertheless, additional experiments are required to make a definite conclusion about the nature of superconductivity in these systems.

We propose two direct experiments to test our model. (i) To provide the Raman scattering experiment of the superconducting $Ba_{0.6}K_{0.4}BiO_3$ compound using a resonance optical excitation in the range of the optical pseudogap $E_g \approx 0.5$ eV. In this case, a sharp increase in the amplitudes of some Raman modes due to local dynamic distortions must be observed at the pair destruction energy in accordance with our model. (ii) To provide measurements of the inelastic neutron scattering in the $Ba_{0.5}K_{0.5}BiO_3$ and $BaPbO_3$ samples. We expect that the dispersion of the longitudinal stretching mode should decrease with a change in the K doping from $x = 0.4$ to $x = 0.5$ and should be absent in the metallic $BaPbO_3$ compound.

Moreover, it is important to carry out more precise measurements of the specific heat in the bismuthates for $T \sim T_c$. In the three-dimensional Bose gas, the specific heat behaves as $C_B \sim (T/T_c)^{3/2}$ for the temperatures $T < T_c$ and $C_B = \text{const}$ for $T \gg T_c$. As a result, there is a λ-point behavior of the specific heat for $T \sim T_c$. However, the Fermi–Bose mixture gives an additional contribution from the Fermi gas, $C_F \sim \gamma T$. This contribution could in principle destroy the λ-point behavior of the specific heat in the Fermi–Bose mixture. We note that the currently available experimental results in the bismuthates [60] signal a smooth behavior of the specific heat near T_c, because, in all the experiments, the contributions of the Fermi and Bose gases are masked by a larger lattice contribution.

ACKNOWLEDGMENTS

We acknowledge fruitful discussions with N.M. Plakida, Yu. Kagan, P. Fulde, P. Woelfe, and A.N. Mitin. This work was supported by the Russian Foundation for Basic Research (project no. 99-02-17343) and the Federal Program "Superconductivity" (project no. 99010). M.Yu.K. is grateful for the grant no. 96-15-96942 from the President of Russia.

REFERENCES

1. M. R. Shafroth, Phys. Rev. B **100**, 463 (1955); Solid State Phys. **10**, 422 (1960).
2. A. S. Alexandrov and J. Ranninger, Phys. Rev. B **23**, 1796 (1981).
3. G. M. Eliashberg, Zh. Éksp. Teor. Fiz. **43**, 1105 (1963) [Sov. Phys. JETP **16**, 780 (1963)].
4. A. J. Leggett, in *Modern Trends in Theory of Condensed Matter: Lecture Notes of the 1979 Karpatz Winter*

School, Ed. by A. Pekalski and J. Przystowa (Springer-Verlag, Berlin, 1980), p. 14.

5. P. Nozieres and S. Schmitt-Rink, J. Low-Temp. Phys. **59**, 195 (1985).
6. J. Ranninger and S. Robaszkiewicz, Physica B (Amsterdam) **135**, 468 (1985).
7. B. K. Chakraverty, J. Ranninger, and D. Feinberg, Phys. Rev. Lett. **81**, 433 (1998).
8. P. W. Anderson, Science **235**, 1196 (1987).
9. V. B. Geshkenbein, L. B. Ioffe, and A. I. Larkin, Phys. Rev. B **55**, 3173 (1997).
10. Ch. Rener, B. Revaz, J.-Y. Genoud, *et al.*, Phys. Rev. Lett. **80**, 149 (1998).
11. L. F. Mattheiss and D. R. Hamann, Phys. Rev. B **28**, 4227 (1983).
12. M. Shirai, N. Suzuki, and K. Motizuki, J. Phys.: Condens. Matter **2**, 3553 (1990).
13. G. Vielsack and W. Weber, Phys. Rev. B **54**, 6614 (1996).
14. A. I. Liechtenstein, I. I. Mazin, C. O. Rodriguez, *et al.*, Phys. Rev. B **44**, 5388 (1991).
15. A. Taraphder, H. R. Krishamurthy, R. Pandit, *et al.*, Europhys. Lett. **21**, 79 (1993).
16. A. Taraphder, H. R. Krishamurthy, R. Pandit, *et al.*, Phys. Rev. B **52**, 1368 (1995).
17. V. Meregalli and S. Y. Savrasov, Phys. Rev. B **57**, 14453 (1998).
18. S. Uchida, K. Kitazawa, and S. Tanaka, Phase Transit. **8**, 95 (1987).
19. M. Qvarford, V. G. Nazin, A. A. Zakharov, *et al.*, Phys. Rev. B **54**, 6700 (1996).
20. N. V. Anshukova, A. I. Golovashkin, V. S. Gorelik, *et al.*, J. Mol. Struct. **219**, 147 (1990).
21. A. P. Menushenkov, K. V. Klementev, P. V. Konarev, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **67**, 977 (1998) [JETP Lett. **67**, 1034 (1998)].
22. A. P. Menushenkov and K. V. Klementev, J. Phys.: Condens. Matter **12**, 3767 (2000).
23. S. Salem-Sugui, Jr., E. E. Alp, S. M. Mini, *et al.*, Phys. Rev. B **43**, 5511 (1991).
24. Y. Yacoby, S. M. Heald, and E. A. Stern, Solid State Commun. **101**, 801 (1997).
25. R. J. Cava, B. Batlogg, J. J. Krajewski, *et al.*, Nature **332**, 814 (1988).
26. Shiyou Pei, J. D. Jorgensen, B. Dabrowski, *et al.*, Phys. Rev. B **41**, 4126 (1990).
27. D. E. Cox and A. W. Sleight, Acta Crystallogr., Sect. B: Struct. Cryst. Chem. **B35**, 1 (1989).
28. A. P. Menushenkov, Nucl. Instrum. Methods Phys. Res. A **405**, 365 (1998).
29. Shiyou Pei, J. D. Jorgensen, B. Dabrowski, *et al.*, Phys. Rev. B **41**, 4126 (1990).
30. A. W. Sleight, J. L. Gillson, and P. E. Bierstedt, Solid State Commun. **17**, 27 (1975).
31. Shunji Sugai, Jpn. J. Appl. Phys., Suppl. **26**, 1123 (1987).
32. Y. J. Uemura, B. J. Sternlieb, D. E. Cox, *et al.*, Nature **335**, 151 (1988).
33. N. Mott, Supercond. Sci. Technol. **4**, S59 (1991).
34. T. M. Rice and L. Sneddon, Phys. Rev. Lett. **47**, 689 (1981).
35. C. M. Varma, Phys. Rev. Lett. **61**, 2713 (1988).
36. L. J. De Jongh, Physica C (Amsterdam) **152**, 171 (1998).
37. S. Sugai, Solid State Commun. **72**, 1187 (1989).
38. J. Yu, X. Y. Chen, and W. P. Su, Phys. Rev. B **41**, 344 (1990).
39. S. Tajima, S. Uchida, A. Masaki, *et al.*, Phys. Rev. B **32**, 6302 (1985).
40. S. Tajima, M. Yoshida, N. Koshizuka, *et al.*, Phys. Rev. B **46**, 1232 (1992).
41. M. Yu. Kagan, R. Frésard, M. Capezzali, and H. Beck, Phys. Rev. B **57**, 5995 (1998).
42. E. S. Hellman, B. Miller, J. M. Rosamilia, *et al.*, Phys. Rev. B **44**, 9719 (1991).
43. E. S. Hellman and E. H. Hartford, Jr., Phys. Rev. B **52**, 6822 (1995).
44. S. H. Blanton, R. T. Collins, K. H. Kelleher, *et al.*, Phys. Rev. B **47**, 996 (1993).
45. N. R. Khasanova, A. Yamamoto, S. Tajima, *et al.*, Physica C (Amsterdam) **305**, 275 (1998).
46. A. P. Menushenkov, A. V. Tsvyashchenko, D. V. Eremenko, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **43**, 591 (2001) [Phys. Solid State **43**, 613 (2001)].
47. Wei Jin, M. H. Degani, R. K. Kalia, *et al.*, Phys. Rev. B **45**, 5535 (1992).
48. N. M. Plakida, V. L. Aksenov, and S. L. Drechsler, Europhys. Lett. **4**, 1309 (1987).
49. J. R. Hardy and J. W. Flocken, Phys. Rev. Lett. **60**, 2191 (1988).
50. J. H. Lee, K. Char, Y. W. Park, *et al.*, Phys. Rev. B **61**, 14815 (2000).
51. T. Mijakawa, T. Suzuki, and H. Yabu, Phys. Rev. A **62**, 063613 (2000).
52. M. J. Bijlsma, B. A. Heringa, and H. T. C. Stoof, Phys. Rev. A **61**, 053601 (2000).
53. M. Braden, W. Reichardt, W. Schmidbauer, *et al.*, J. Supercond. **8**, 595 (1995).
54. L. Pintschovius and W. Reichardt, in *Physical Properties of High Temperature Superconductors*, Ed. by D. M. Ginsberg (World Scientific, Singapore, 1994), Vol. IV.
55. J. Orenstein and A. J. Mills, Science **288**, 468 (2000).
56. K. A. Müller, Guo-Meng Zhao, K. Conder, *et al.*, J. Phys.: Condens. Matter **10**, L291 (1998).
57. D. Jorgensen, B. Dabrowski, S. Pei, *et al.*, Phys. Rev. B **38**, 11337 (1988).
58. V. Yu. Pomyakushin, A. A. Zakharov, A. M. Balagurov, *et al.*, Phys. Rev. B **58**, 12350 (1998).
59. V. J. Emery and S. A. Kivel’son, Physica C (Amsterdam) **209**, 597 (1993).
60. S. E. Stupp, M. E. Reeves, D. M. Ginsberg, *et al.*, Phys. Rev. B **40**, 10878 (1989).

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 625–629.
Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 712–717.
Original Russian Text Copyright © 2001 by Meilikhov.

SOLIDS
Electronic Properties

High-Temperature Conduction of Granular Metals

E. Z. Meilikhov
Institute of Molecular Physics, Kurchatov Institute Russian Research Centre,
pl. Kurchatova 1, Moscow, 123182 Russia
e-mail: meilikhov@imp.kiae

Received March 26, 2001

Abstract—In theoretical and experimental studies of conductivity associated with intergranular tunneling of electrons in nanocomposites (granular metals), only the range of relatively low temperatures was traditionally investigated, in which only electron transitions involving singly-charged metal granules are significant. In this mode, the temperature dependence of conductivity is exponential. However, experiments show that the type of conduction of nanocomposites at high temperatures changes significantly. In the model proposed in this article, the features of conduction of granular metals at high temperatures are attributed to multiply charged granules with a large spread of the size. The conclusions of the model are in good agreement with the experimental results. *© 2001 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Granular metals (metal–insulator nanocomposites), which have the form of small metallic inclusions (with a radius $r = 1$–100 nm) in an insulator matrix, possess a number of unique properties determined by the volume concentration x of the conducting phase [1–3]. There exists a critical value of x_c such that the material possesses metallic properties for $x > x_c$ and is an insulator with thermally activated conduction for $x < x_c$. It was found that the conduction is associated with tunnel transitions of charge carriers between granules. In this respect, its mechanism is close to the well-known hopping conduction over impurities in doped semiconductors. However, in contrast to “pointlike” impurity centers in a semiconductor, granules in a nanocomposite have a finite size. For this reason, only the tunnel transitions between nearest (or next to nearest) granules are “geometrically” allowed. This distinguishes the mechanism of tunnel conduction of nanocomposites from the hopping conduction of impurity semiconductors with a variable range, in which jumps of any length are admissible in principle.

In addition, we must take into account the fact that only an insignificant part of granules in a nanocomposite effectively participate in the low-temperature conduction in the case of a large spread of grain size. As a matter of fact, the equilibrium concentration of singly charged granules (and the more so multiply charged granules) is small at low temperatures. Consequently, the conductivity of a nanocomposite is determined by tunnel transitions of electrons only between singly charged granules and the most abundant neutral granules. The probability of such transitions depends on their activation energy Δ and the distance l between granules. Usually, these factors produce opposite effects: upon a decrease in the granule radius r, the energy $\Delta \sim e^2/\varepsilon r$ increases and the distance l decreases. For this reason, the current mainly flows over granules of the so-called optimal size r_{opt} which decreases upon heating [4]. Hence the low-temperature conductivity G is described by the well-known “1/2 law”

$$G \propto \exp[-(T_0/T)^{1/2}],$$

where $T_0 = T_0(x)$ is a certain characteristic temperature.

Experiments show, however, that the “1/2 law” is violated at relatively high temperatures [5]: the conductivity starts increasing much more rapidly with temperature. In all probability, this is due to the fact that the number of multiply charged granules increases significantly upon an increase in temperature and their contribution to the conductivity cannot be neglected any longer. The present work is devoted to an analysis of precisely this mode of conduction of nanocomposites.

2. EQUILIBRIUM IONIZATION OF NANOCOMPOSITE GRANULES

In the following analysis, we will assume that the temperature is sufficiently high so that the interaction of charges at granules, which leads to charge correlations and to the emergence of the Coulomb gap in the density of charged states, can be disregarded [6]. In actual practice, this means that $T \gtrsim 1$ K.

We begin the analysis of the equilibrium ionization of nanocomposite granules with a system consisting of granules of the same radius r. The concentration N_m of granules with charge me (e is the absolute value of the electron charge and $m = 0, \pm1, \pm2, \ldots$) and with the

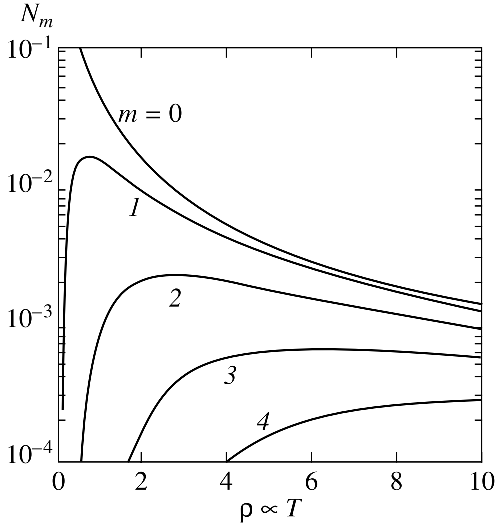


Fig. 1. Temperature dependences of concentration N_m of granules with charge *me*. Singularities in the temperature dependence of conductance should be expected for the value of $\rho \approx 0.5$ corresponding to the maximum concentration of singly charged granules.

Coulomb energy $(me)^2/\varepsilon r$ (ε is the permittivity of the medium) is obviously given by

$$\frac{N_m}{N} = \frac{\exp(-m^2/\rho)}{\sum\limits_{m=-\infty}^{\infty} \exp(-m^2/\rho)}, \qquad (1)$$

where N is the total concentration of granules and $\rho = r(\varepsilon kT/e^2)$ is the reduced radius of granules.

Relation (1) shows that the charge distribution of granules in our case is Gaussian. The width of this distribution is determined by parameter ρ (i.e., the relation between the thermal and Coulomb energies) and increases with temperature. At low temperatures ($\rho \ll 1$), we have predominantly singly ionized granules ($m = \pm 1$), while, at high temperatures ($\rho \gg 1$), the charge spectrum of granules may be very broad. For example, the fraction of doubly ionized granules for $\rho = 4$ is only half the fraction of singly ionized granules (Fig. 1).

Another important consequence of relation (1) is that the concentration N_m of granules with a certain charge *me* is a nonmonotonic function of temperature. It can be seen from Fig. 1 that the initial increase in this concentration with temperature is replaced by its subsequent decrease. Clearly, all these circumstances must be taken into account while calculating the high-temperature conductivity of a nanocomposite.

3. CONDUCTIVITY OF A NANOCOMPOSITE WITH GRANULES OF THE SAME SIZE

If we introduce the symbol Z_m denoting a granule with a charge *me*, all intergranular electron transitions occurring in the system can be presented in the form of a set of "reactions"

$$Z_m + Z_n \longleftrightarrow Z_{m+1} + Z_{n-1}, \qquad (2)$$

each of which makes a "feasible" contribution G_{mn} to the total conductivity G. The change in energy as a result of such a transition is given by

$$\frac{\Delta_{mn}}{kT} = \frac{m-n+1}{\rho}. \qquad (3)$$

In the case when a "reaction" is endothermic (i.e., $\Delta_{mn} > 0$), the lacking energy is supplied by phonons. Consequently, the conductance of the system can be written in the form

$$G = \sum_{m,n} G_{mn},$$

$$G_{mn} \propto N_m N_n P_{mn} \exp\left(-\frac{l_{mn}}{\lambda}\right),$$

$$P_{mn} = \begin{cases} \dfrac{\Delta_{mn}/kT}{\exp(\Delta_{mn}/kT) - 1}, & \Delta_{mn} \ge 0 \\ 1, & \Delta_{mn} < 0, \end{cases} \qquad (4)$$

where the factor P_{mn} is connected with the Planck distribution of phonons. Here, $l_{mn} \propto N^{-1/3}$ is the mean distance between neighboring granules, which determines the typical width of the potential barrier overcome by tunneling electrons and $\lambda \sim \hbar/(mW)^{1/2}$ is the electron wavelength in the insulator (W is the height of the tunnel barrier, which virtually coincides with the half-width of the band gap in the insulator).

It can be seen from relations (4) that the largest contribution to the total conductivity comes from (m, n) transitions between granules whose charges are not very large. However, the range of "admissible" charges expands with increasing temperature. As a result, the set of dominant transitions becomes richer and richer upon heating. This is illustrated in Fig. 2, showing schematically the sets of the most significant transitions for three values of the parameter $\rho \propto T$. The height of each vertical segment is proportional to the contribution G_{mn} of the corresponding transition, and its position x on the abscissa axis on all the diagrams is chosen so that each value of x corresponds to a transition (m, n), where $m = [x]$ and $n = 10(x - [x])$, $[x]$ being the integer closest to x.

The temperature dependence of the total conductance for the given case is presented in Fig. 3. In accordance with Fig. 2, only transitions (–1, 0) and (0, 1) for which $\Delta_{mn} = 0$ are significant at low temperatures. For this reason, we have

$$G \propto (N_{-1}N_0 + N_0N_1) \propto \exp(e^2/\varepsilon rkT).$$

At high temperatures, this simple exponential dependence is violated. On the one hand, we can expect a decrease in conductance due to the above-mentioned decrease in the concentration of granules with a fixed charge upon heating (see Fig. 1). On the other hand, it must increase due to the contribution of multiply ionized granules to conductance. As a result, the temperature dependence of the total conductivity of a nanocomposite deviates considerably (in the downward direction) from the exponential dependence in the range of high temperatures. This deviation starts at $\rho \approx 0.5$, i.e., at the temperature for which transitions with a nonzero energy balance start playing a significant role in accordance with Fig. 2. At a still higher temperature ($\rho \gg 1$), the total conductivity approaches saturation.

4. CONDUCTIVITY OF A NANOCOMPOSITE WITH GRANULES OF VARIOUS SIZES

Let us now calculate the conductivity of a nanocomposite in a more realistic system formed by granules of various sizes whose spread is characterized by the distribution function $f(\rho)$. By virtue of the detailed balance principle, the concentration of the m-fold ionized granules with radii ranging from ρ to $\rho + d\rho$ is defined, as before, by a formula of type (1):

$$dN_m^{(\rho)} = Nf(\rho)d\rho\frac{\exp(-m^2/\rho)}{\sum\limits_{m=-\infty}^{\infty}\exp(-m^2/\rho)}, \tag{5}$$

where $Nf(\rho)d\rho$ is the total concentration of such granules.

Let us consider the reaction

$$Z_m^{(1)} + Z_n^{(2)} \longleftrightarrow Z_{m+1}^{(1)} + Z_{n-1}^{(2)}, \tag{6}$$

in which an electron passes from a granule of radius ρ_1 to a granule of radius ρ_2. The energy effect $\Delta_{mn}^{(12)}$ of such a reaction is given by

$$\frac{\Delta_{mn}^{(12)}}{kT} = \frac{m+1/2}{\rho_1} - \frac{n-1/2}{\rho_2}. \tag{7}$$

The contribution $dG_{mn}^{(12)}$ of such transitions to the total conductivity is expressed by the formula

$$dG_{mn}^{(12)} \propto dN_m^{(\rho_1)}dN_n^{(\rho_2)}P_{mn}^{(12)},$$

$$P_{mn}^{(12)} = \begin{cases} \dfrac{\Delta_{mn}^{(12)}/kT}{\exp(\Delta_{mn}^{(12)}/kT) - 1}, & \Delta_{mn}^{(12)} \ge 0 \\ 1, & \Delta_{mn}^{(12)} < 0, \end{cases} \tag{8}$$

where $dN_m^{(\rho_1)}$ and $dN_n^{(\rho_2)}$ are defined by formula (5).

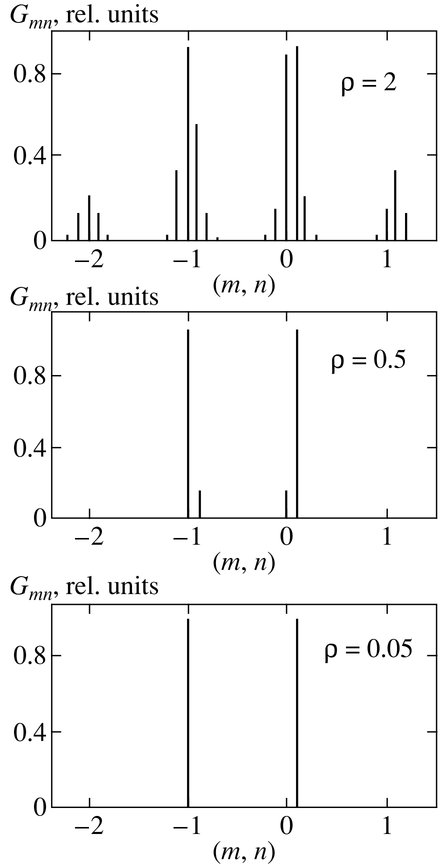


Fig. 2. Contributions G_{mn} of significant intergranular transitions at various temperatures (the temperature increases from the lower to the upper diagram).

Let us first analyze the case of low temperatures (ρ_1, $\rho_2 \gg 1$), when the conductance is associated with the electron transitions (–1, 0) and (0, 1) making identical contributions to it. In this case, we have

$$\frac{\Delta_{(-1,0)}^{(12)}}{kT} = -\frac{\Delta_{(0,1)}^{(12)}}{kT} = \frac{\rho_1 - \rho_2}{2\rho_1\rho_2},$$

and, hence, transitions (–1, 0) occur only for $\rho_1 < \rho_2$ (electron transition from a small granule with a single negative charge to a larger neutral granule), while transitions (0, 1) take place only for $\rho_1 > \rho_2$ (electron transition from a large neutral granule to a small granule with a single positive charge). In both cases, we can assume that $P_{mn}^{(12)} = 1$. Let us estimate, for example, the contribution $G_{(0,1)}^{(12)} = \int dG_{(0,1)}^{(12)}$ from the second of these transitions, assuming, for the sake of definiteness, that the size distribution function for granules has the form $f(\rho) = (1/\rho_0)\exp(-\rho/\rho_0)$ [7]. Using relations (8), we obtain

$$G_{(0,1)}^{(12)} \propto \iint\limits_{\rho_1 > \rho_2} \exp\left(-\frac{\rho_1}{\rho_0}\right)\exp\left(-\frac{\rho_2}{\rho_0}\right) \times \exp\left(-\frac{1}{\rho_2}\right)d\rho_1 d\rho_2 \tag{9}$$

$$= \rho_0\int\limits_0^\infty \exp\left[-\left(\frac{2\rho_2}{\rho_0} + \frac{1}{\rho_2}\right)\right]d\rho_2.$$

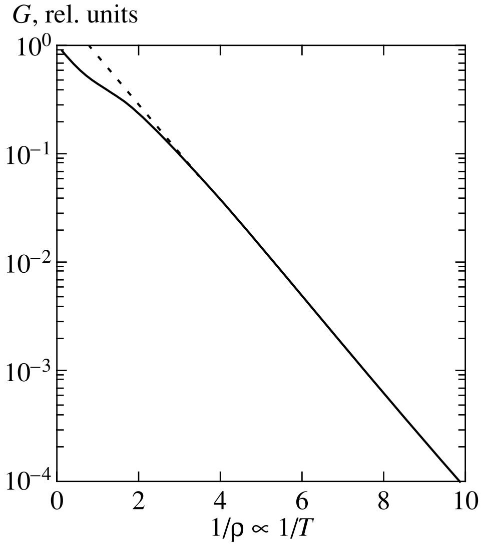


Fig. 3. Temperature dependence (solid curve) of the conductance of a nanocomposite with granules of the same size. The dashed line corresponds to the exponential dependence $G \propto \exp(-\mathrm{const}/T)$.

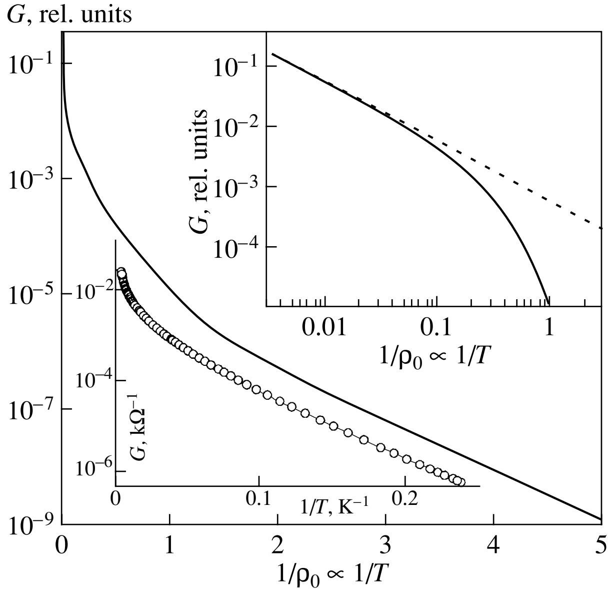


Fig. 4. Temperature dependence (solid curves) of the conductance of a nanocomposite with granules of different sizes. The dashed line in the upper inset corresponds to the linear dependence $G \propto \rho_0 \propto T$. The lower inset shows the experimental curve from [5].

The integrand function attains its maximum value for $\rho_2 = \sqrt{\rho_0/2}$ (this is the so-called optimal size of granules) and the estimate obtained by the steepest descent method is given by

$$G^{(12)}_{(0,\,1)} \propto \exp[-\sqrt{8/\rho_0}].$$

The same contribution to the conductance comes from transitions (–1, 0) so that the low-temperature dependence of the total conductance has the form

$$G \propto \exp\left[-\sqrt{\frac{8e^2}{\varepsilon r_0 kT}}\right], \qquad (10)$$

where r_0 is the mean radius of granules, which determines their size distribution function:

$$f(r) = \frac{1}{r_0}\exp\left(-\frac{r}{r_0}\right).$$

Thus, the conductance at low temperatures is described by the “1/2 law”:[1]

$$\frac{\partial \ln\ln G}{\partial T} = -\frac{1}{2}.$$

In the model under investigation, this law is observed due to the fact that the size of singly charged (positive or negative) granules through which the conduction predominantly takes place increases upon cooling. Naturally, this occurs as long as the optimal granule radius

$$r_{\mathrm{opt}} = r_0\sqrt{e^2/\varepsilon r_0 kT},$$

is below their maximum size $r_{\max}$ (for a given system), i.e., for $kT > e^2 r_0/\varepsilon r^2_{\max}$. At lower temperatures, we have

$$G \propto \exp[-(e^2/\varepsilon r_{\max} kT)].$$

Let us now consider the case of high temperatures ($\rho_1, \rho_2 \gg 1$) for which the electron transitions involving multiply charged granules become significant. In this case, for the former distribution function

$$f(\rho) = (1/\rho_0)\exp(-\rho/\rho_0),$$

we obtain from (4) and (7)

$$G \propto \sum_{m,\,n}\int_{\rho_1}\int_{\rho_2}\frac{\exp[-(\rho_1/\rho_0 + m^2/\rho_1)]}{\rho_0\psi(\rho_1)} \times \frac{\exp[-(\rho_2/\rho_0 + n^2/\rho_2)]}{\rho_0\psi(\rho_2)}P^{(12)}_{mn}\,d\rho_1 d\rho_2, \qquad (11)$$

where

$$\psi(u) = \sum_m \exp\left(-\frac{m^2}{u}\right)$$

(for $u \gg 1$, we can use the asymptotic relation $\psi(u) = 2.37u^{3/2}$).

The temperature dependence of conductance determined with the help of relation (11) is presented in Fig. 4. The strong variation in the conduction in the high-temperature range ($\rho_0^{-1} \ll 1$) is associated with the gradual inclusion of new conduction channels through multiply charged granules. The same dependence presented in

---

[1] It is well known that this law is also preserved for other types of size distribution functions for granules under the condition $\partial f(r)\partial r < 0$ [4].

different coordinates (solid curve in the upper inset in Fig. 4) shows that, at high temperatures, the conductance is a power (and not exponential) function of temperature. For $\rho_0 \sim 10$, the temperature dependence of the conductance attains the asymptotic form (dashed line in the upper inset) corresponding to the linear function $G \propto \rho_0 \propto T$.

For the sake of comparison, the lower inset in Fig. 4 shows the experimental temperature dependence of the conductivity of the nanocomposite $Fe_x(SiO_2)_{1-x}$ ($x \approx 0.4$) from [5]. While comparing this dependence with the theoretical curve, it should be borne in mind that the granule size in the nanocomposite under investigation lies in the interval 10 Å $< r <$ 1000 Å in accordance with electron-microscopic data, the number of small granules being much larger than the number of large granules. This means that $r_0 \sim 100$ Å. In addition, we must take into account the fact that the effective permittivity ε of the nanocomposite is much higher[2] than the permittivity $\varepsilon_d \approx 3.5$ of its dielectric component SiO_2 due to the closeness to the percolation transition. Setting $\varepsilon = 50$, we find that $\rho_0 \approx 0.25$ (or $1/\rho_0 \approx 4$) corresponds to $T = 4$ K. It can be seen from Fig. 4 that the matching between the theoretical and experimental dependences is satisfactory.

ACKNOWLEDGMENTS

This work was supported financially by the Russian Foundation for Basic Research (project nos. 99-02-16955 and 98-02-17191) and by the PICS–RFBR program (grant 98-02-22037).

REFERENCES

1. Philos. Mag. B **65**, Special Issue (1992).
2. C. J. Adkins, in *Metal–Insulator Transitions Revisited*, Ed. by P. P. Edwards and C. N. R. Rao (Taylor & Francis, London, 1995); J. Phys.: Condens. Matter **1**, 1253 (1989).
3. P. Sheng, Philos. Mag. B **65**, 357 (1992).
4. E. Z. Meilikhov, Zh. Éksp. Teor. Fiz. **115**, 624 (1999) [JETP **88**, 819 (1999)].
5. B. A. Aronzon, A. E. Varfolomeev, D. Yu. Kovalev, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 944 (1999) [Phys. Solid State **41**, 857 (1999)].
6. E. Cuevas, M. Ortuño, and J. Ruitz, Phys. Rev. Lett. **71**, 1871 (1993).
7. A. B. Pakhomov, X. Yan, N. Wang, *et al.*, Physica A (Amsterdam) **241**, 344 (1997).

Translated by N. Wadhwa

---

[2] According to [4], $\varepsilon \approx \varepsilon_d/[1-(x/x_c)^{1/3}] \approx 15\varepsilon_d$ for $x/x_c = 0.8$.

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 630–641.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 718–730.

Anomalous Penetration of an Electromagnetic Signal in a Thin Metallic Plate under Strong Magnetodynamic Nonlinearity Conditions

S. A. Derevyanko*, G. B. Tkachev, and V. A. Yampol'skiĭ
Usikov Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine, Kharkov, 61085 Ukraine
**e-mail: stanislv@ire.kharkov.ua*
Received April 10, 2001

Abstract—A new phenomenon of anomalous transparency of a metallic plate upon the passage of a strong transport current is predicted. It is shown that the electromagnetic field of an incident wave may be carried from the skin layer to the opposite face by electrons trapped by the alternating intrinsic magnetic field of the transport current even under the conditions of the extremely anomalous skin effect for which the skin depth is the smallest parameter having the dimensions of length. The mechanism of the rf field transport effectively operates at low temperature, when the mean free path of charge carriers is large. The wave field distribution in the plate is analyzed. The cyclotron resonance at trapped electrons, which emerges due to their periodic return to the skin layer, is predicted and qualitatively analyzed. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is well known that normal metals under standard conditions are opaque to electromagnetic radiation: waves are localized in the skin layer due to the skin effect. If, however, a strong constant magnetic field is applied parallel to the sample surface, the electromagnetic field may be transferred from the skin layer to the bulk of the metal. In plates of a finite thickness, the transport of the field of a wave from the surface of incidence to the opposite face makes the sample transparent. Similar effects, which are known as the effects of anomalous penetration, are caused by individual movement of charged particles in a magnetic field. An electromagnetic field penetrates into the bulk of the metal owing to electrons "carrying" the rf field from the skin layer and then "reproducing" it in the bulk of the metal. Azbel [1] was the first to pay attention to such effects. He considered the field transport along a chain of electron trajectories under cyclotron resonance conditions. The theory of anomalous penetration was developed further by Kaner and coauthors. They proved that the transport along a trajectory in a magnetic field parallel to the metal surface is observed in a wide frequency range of electromagnetic field and not only under cyclotron resonance conditions [2]. Other mechanisms of anomalous penetration other than that predicted in [1] were also considered. The variety of the trajectory-type anomalous penetration effects was analyzed in review [3].

The dynamics of charge carriers in a metal can be affected not only by an external constant magnetic field, but also by the magnetic component of the field of a wave or the intrinsic magnetic field of the transport current. The magnetic field of the wave, which is much stronger than the electric field in the bulk of the metal, is responsible for the nonlinearity in the Maxwell equations owing to the dependence of conductivity on the magnetic field. The mechanism of this type of nonlinearity is effective at low temperatures, for which the mean free path of charge carriers is quite large. This mechanism is not associated with a strong deviation of the electron subsystem from equilibrium as, for example, in semiconductors and is known as the magnetodynamic mechanism. Various manifestations of magnetodynamic nonlinearity are considered in reviews [4, 5]. A nontrivial example of a magnetodynamic nonlinearity is the generation of so-called current states discovered experimentally for some metals [6]. A considerable deviation of static current–voltage characteristics (IVC) for thin metallic plates from Ohm's law towards a decrease in resistance associated with the magnetodynamic nonlinearity mechanism was theoretically predicted in [7] and experimentally observed in [8]. The nonlinear interaction of an electromagnetic wave incident on a metallic plate with the transport current passing through it was analyzed in [9]. It was proved that under the conditions of weak spatial dispersion, when the electromagnetic field penetration depth δ is much larger than the sample thickness d, the electrodynamic response of the metal exhibits a number of peculiarities. For example, the time dependence of the electric field at the plate surface is characterized by a series of kinks and peaks.

The variety of the above-mentioned effects indicates that the influence of the intrinsic magnetic field of a wave or the transport current on the dynamics of charge carriers may be significant. For this reason, it would be interesting to find out whether the magnetodynamic nonlinearity may lead to the penetration of a radio wave to large distances from the irradiated surface as in an external constant magnetic field.

In the present work, a new effect of anomalous penetration of an electromagnetic field into a metallic plate upon the passage of a strong transport current is predicted. We consider thin samples and the extremely anomalous skin effect, for which the following inequalities are satisfied:

$$\delta \ll d \ll l. \tag{1}$$

Here, δ is the electromagnetic field penetration depth, d is the plate thickness, and l is the mean free path of charge carriers. One of the sample faces is exposed to a plane monochromatic wave of a preset frequency and amplitude. In the absence of transport current, the field of the wave is mainly localized in the skin layer of thickness δ. It was proved long ago by Pippard [10] that the main contribution of the rf conductivity in the anomalous skin effect comes only from a small fraction of electrons ("effective" electrons) moving at small angles $\sim\delta/l \ll 1$ to the sample surface. In zero magnetic field, effective electrons are obviously unable to carry the field of the wave from the skin layer since they remain in this layer during the entire mean free time. The transport of the rf field to the bulk of the sample is carried out only by noneffective electrons moving at large angles to the faces of the plate (as compared to d/l). However, such electrons receive from the field a much smaller amount of energy than "grazing" particles, and the electromagnetic field created by them at the opposite face of the sample is much weaker than on the irradiated surface.

It will be shown below that the situation changes radically if a strong transport current I passes through the sample (the geometry of the problem is depicted in Fig. 1). The electron trajectories start experiencing a considerable effect of the intrinsic magnetic field of the current, which is distributed antisymmetrically over the sample thickness. It is equal to zero at the middle of the plate and assumes equal and opposite values (H and $-H$) at the opposite faces:

$$H = \frac{2\pi I}{cD}. \tag{2}$$

Here, D is the plate width (the size in the direction of the magnetic field of the current) and c is the velocity of light. The alternating field of the transport current leads to the emergence of a new group of charge carriers (in addition to conventional "flying" electrons), which are trapped in an effective potential well. The trajectories of flying and trapped electrons are presented schematically in Fig. 1, where these trajectories are marked as *1* and *2*, respectively. Trapped carriers move in almost planar periodic curves "oscillating" about the plane $x = 0$ at which the field of the current changes its sign. It can be seen from Fig. 1 that a fraction of trapped electrons (namely, those which get into the skin layer) are able to carry the rf field to the opposite surface of the sample. In this way, the field of the wave penetrates anomalously owing to its transport by the trapped electrons. The physical pattern of this phenomenon is similar to the transport of an rf field along a chain of electron orbits in a uniform magnetic field [3].

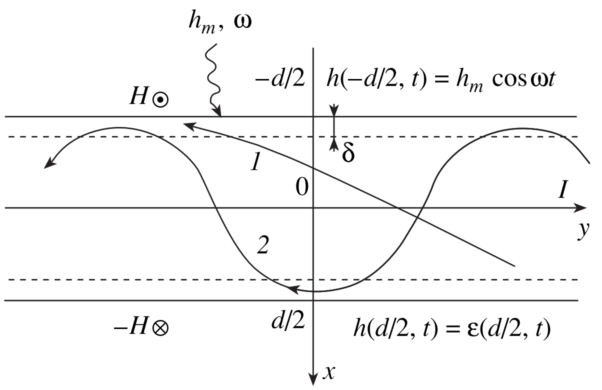


Fig. 1. Geometry of the problem. Schematic diagram of the trajectories of flying (*1*) and trapped (*2*) particles.

In the next section, we will formulate the problem and write the Maxwell equations with the boundary conditions for the total electric and magnetic fields of the current and the wave. In Section 3, the conductivity operators for flying and trapped electrons are analyzed and the Maxwell equations are linearized in the relatively weak field of a radio wave. The specific nature of the conductivity operator for trapped electrons, as well as the requirements imposed by the boundary conditions, leads to a complex distribution of the electromagnetic field in the plate: it turns out that the symmetric and antisymmetric electric field components (over the sample thickness) differ significantly in magnitude and are characterized by different penetration depths. Under the strong nonlinearity conditions, the electric field of the wave is practically antisymmetric, which means that it is carried by trapped electrons from the skin layer to the opposite surface of the metal without attenuation. The asymptotic forms of the conductivity operator are obtained in Section 4 for a strong spatial dispersion ($\delta \ll d$), and a linear integrodifferential equation describing the electric field distribution in the plate is derived. In Section 5, a consistent qualitative analysis of the problem is carried out on the basis of the noneffectiveness concept; correct order-of-magnitude estimates for the signal penetration depth and for the surface impedance of the plate are obtained without solving the complex integrodifferential equation. The pattern of electric field distribution in the sample is numerically analyzed in Section 6. In Section 7, the possibility of observing the cyclotron resonance at trapped electrons is indicated and the necessary conditions for observing such an effect are formulated. Qualitative formulas describing the resonance oscillations of the surface impedance are also given in this section.

The range of applicability of the obtained results is outlined in Conclusions, where some considerations concerning the possibility of experimental observation of the predicted effects are put forth.

2. FORMULATION OF THE PROBLEM

Let us consider a thin metallic plate of thickness d carrying a direct transport current I. We choose the x axis along the normal to the metal surface so that the plane $x = 0$ coincides with the middle of the sample. The boundary $x = -d/2$ is exposed to a plane electromagnetic wave whose magnetic vector is collinear to the magnetic field vector of the current. We direct the y axis along the current and the z axis parallel to the magnetic fields of the current and of the wave (see Fig. 1). The magnetic field in the metal is the sum of the intrinsic field $H(x)$ of the transport current and the field $h(x, t)$ of the wave:

$$\mathbf{H}(x, t) = \{0, 0, \mathcal{H}(x, t)\} = \{0, 0, H(x) + h(x, t)\}. \quad (3)$$

The electric field $\mathbf{E}(x, t)$ in the sample has only the y component which can also be naturally presented as the sum of the uniform potential field $E_0(t)$, which is responsible for the passage of the transport current, and the nonuniform vortex component $\mathcal{E}(x, t)$ of the field of the electromagnetic wave:

$$\mathbf{E}(x, t) = \{0, E(x, t), 0\} = \{0, E_0(t) + \mathcal{E}(x, t), 0\}. \quad (4)$$

We assume that the length L and width D of the plate (i.e., its dimensions along the y and z axes, respectively) are much larger than thickness d. In addition, we consider the diffuse scattering of electrons at the sample surface.

In the geometry adopted by us, the Maxwell equations have the form

$$\begin{aligned} -\frac{\partial \mathcal{H}(x, t)}{\partial x} &= \frac{4\pi}{c} j(x, t), \\ \frac{\partial E(x, t)}{\partial x} &= -\frac{1}{c}\frac{\partial \mathcal{H}(x, t)}{\partial t}, \end{aligned} \quad (5)$$

where $j(x, t)$ is the y component of current density. The boundary condition for Eqs. (5) at the irradiated boundary $x = -d/2$ has the form

$$\mathcal{H}(-d/2, t) = h_m \cos\omega t + H, \quad (6)$$

where $H = 2\pi I/cD$ is the magnetic field of the current at the plate surface and h_m is the amplitude of the wave. At the opposite face of the sample, the magnetic and electric components of the wave transmitted through the sample must coincide; i.e., the following equality must hold:

$$h(d/2, t) = \mathcal{E}(d/2, t). \quad (7)$$

The signal incident on the plate is assumed to be weak; i.e., the magnetic and electric components h and $\mathcal{E}$ of the wave field are much smaller than the intrinsic field $H(x)$ of the current and the potential electric field $E_0(t)$, respectively. Consequently, Maxwell's equations can be linearized in the fields $h(x, t)$ and $\mathcal{E}(x, t)$. In addition, we will disregard time dispersion; i.e., we assume that the frequency ω of the alternating field is much smaller than the relaxation frequency ν of charge carriers:

$$\omega \ll \nu. \quad (8)$$

3. DYNAMICS OF CHARGE CARRIERS. ASYMPTOTIC FORMS OF CONDUCTIVITY OPERATORS

Owing to the relative weakness of the radiowave field, we can study the electron dynamics disregarding the effect of this field on electron trajectories and, hence, simplify the problem as compared to that considered in [7] and [9]. Following the general algorithm proposed in [9], we chose the vector potential gauge of the field induced by the transport current in the form

$$A(x) = \int_0^x H(x')\,dx'. \quad (9)$$

It was mentioned above that the magnetic field of the current is distributed antisymmetrically and vanishes at the middle of the plate:

$$H(0) = 0.$$

The vector potential $A(x)$, however, is a symmetric function of coordinate x. It is negative almost everywhere and attains its maximum value equal to zero only at point $x = 0$. The integrals of motion of an electron in the antisymmetric field of the current are the total energy equal to the Fermi energy and the generalized momenta

$$p_z = m v_z, \quad p_y = m v_y - eA(x)/c$$

(m is the electron mass). The generalized momenta can be used to express the electron velocity components $v_x(x)$ and $v_y(x)$ in a plane perpendicular to the magnetic field. In the case of a spherical Fermi surface of radius $p_F = mv$, we have

$$\begin{gathered} v_y(x) = \frac{p_y + eA(x)/c}{m}, \\ |v_x(x)| = [v_\perp^2 - v_y^2(x)]^{1/2}, \quad v_\perp = [v^2 - v_z^2]^{1/2}. \end{gathered} \quad (10)$$

It is well known [7, 9] that electrons can be divided into the following three groups depending on the magnitude and sign of the integral of motion p_y: trapped, flying, and surface carriers. The last group is insignificant in the case of diffuse scattering of charge carriers at the sample boundaries considered here (these electrons play a significant role only in the case of specular reflection).

The conductivity of trapped electrons depends to a considerable extent on the transport current. In the static case, when the external varying field is equal to zero, the strong nonlinearity mode corresponds to the values of current for which the contribution from the trapped electrons to conductivity is much larger than the contribution of flying electrons. According to [7], the nonlinearity parameter r determining the relative contribution of trapped electrons coincides with the ratio of the mean free path l to the characteristic length of the arc of the trajectory of a trapped particle:

$$r = \frac{l}{(Rd)^{1/2}}, \quad R = \frac{cp_F}{eH}. \tag{11}$$

Here, R is the characteristic radius of curvature of an electron trajectory in the magnetic field of the current, and e is the elementary charge. Thus, the nonlinearity parameter r is equal to the number of oscillations performed by a trapped electron relative to the plane $x = 0$ before it experiences a scattering act. Henceforth, we will consider only the case of strong nonlinearity, i.e.,

$$r \gg 1. \tag{12}$$

In order to find the conductivity operators for flying and trapped electrons, we must carry out a rather complicated and cumbersome procedure of solving the Boltzmann kinetic equation. We will not present the corresponding calculations here since they are described in detail in [7]. It should only be noted that the current density in [7] was obtained for a uniform electric field. In the case of strong spatial dispersion considered by us here, the expression for current density is different since the electric field cannot be taken out of the symbol of the conductivity operator. The corresponding formulas for the current density j_{fl} of flying and j_{tr} of trapped electrons have the form

$$\begin{aligned} j_{fl} = \hat{\sigma}_{fl}\mathcal{E}(x,t) = \frac{3\sigma_0}{2\pi l}\int_0^v \frac{dv_z}{v} \\ \times \int_{p_y^*}^{mv_\perp} \frac{dp_y}{p_F}\frac{v_y(x)}{|v_x(x)|}\int_{-d/2}^{d/2} dx'\frac{v_y(x')}{|v_x(x')|} \\ \times \exp(-\nu|\tau(x,x')|)\mathcal{E}(x'), \end{aligned} \tag{13}$$

$$\begin{aligned} j_{tr}(x) = \hat{\sigma}_{tr}(x,t) = \frac{3\sigma_0}{\pi l}\int_0^v \frac{dv_z}{v} \\ \times \int_{-mv_\perp-(e/c)A(x)}^{p_y^*} \frac{dp_y}{p_F}\frac{v_y(x)}{|v_x(x)|}\left\{\frac{\cosh\nu\tau(x_1,x)}{\sinh\nu T}\right. \\ \times \int_{x_1}^{x_2} dx'\frac{v_y(x')}{|v_x(x')|}\cosh\nu\tau(x',x_2)\mathcal{E}(x') \\ \left. + \int_{x_1}^{x} dx'\frac{v_y(x')}{|v_x(x')|}\sinh\nu\tau(x,x')\mathcal{E}(x')\right\}, \end{aligned} \tag{14}$$

where

$$\tau(x,x') = \int_x^{x'} \frac{dx''}{|v_x(x'')|}, \tag{15}$$

$x_1 = -x_2 \equiv -x^*$ are the turning points for electron trajectories (which are the roots of the equation $|v_x| = 0$)

$$p_y^* = -mv_\perp - (e/c)A(d/2) \tag{16}$$

is the upper boundary (in momenta p_y) of the region of existence of the trapped electrons (the lower boundary of this region is $p_y = -mv_\perp$), and σ_0 is the conductivity of a bulk sample. The quantity $2T \equiv 2\tau(x_1, x_2)$ is equal to the period of motion of a trapped electron.

In the zeroth approximation, the Maxwell equations (5) give a nonlinear current–voltage characteristic of the plate considered in [7]. We are interested in the expansion terms linear in the field of the wave, which are responsible for the electrodynamic response. In the linear approximation, we seek the solutions in the form

$$\mathcal{E}(x,t) = \mathcal{E}(x)\exp(-i\omega t), \quad h(x,t) = h(x)\exp(-i\omega t).$$

In this case, Maxwell's equations assume the form

$$\begin{aligned} \mathcal{E}'(x) &= \frac{i\omega}{c}h(x), \\ -h'(x) &= \frac{4\pi}{c}[\hat{\sigma}_{tr}\mathcal{E}(x) + \hat{\sigma}_{fl}\mathcal{E}(x)]. \end{aligned} \tag{17}$$

The boundary conditions (6) and (7) can be written in the form

$$h(-d/2) = h_m, \quad h(d/2) = \mathcal{E}(d/2). \tag{18}$$

The strong nonlinearity mode presumes large values of currents. However, we will assume that the current is not very large, so that the condition

$$d \ll R, \tag{19}$$

which allows us to disregard the curvature of the trajectories of the flying electrons, holds along with inequality (12). Inequality (19) allows us to take all the quantities in zero magnetic field in expression (13) for the conductivity operator for flying electrons. In this case, formulas (10), (15), and (16) are transformed to

$$p_y^* = -mv_\perp, \quad v_y = \frac{p_y}{m}, \quad v_x = \sqrt{v_\perp^2 - p_y^2/m^2}. \tag{20}$$

Consequently, the conductivity operator (13) can be written in the form

$$j_{fl} = \frac{3\sigma_0}{2\pi l}\int_0^{v}\frac{dv_z}{v}\int_{p_y^*}^{mv_\perp}\frac{dp_y}{p_F}\frac{p_y^2/m^2}{v_\perp^2 - p_y^2/m^2} \times \int_{-d/2}^{d/2}\exp\left(-\frac{\nu|x-x'|}{\sqrt{v_\perp^2 - p_2^2/m^2}}\right)\mathcal{E}(x')dx'. \tag{21}$$

In this integral, we carry out the substitution $t = p_y/mv_\perp$. This gives

$$j_{fl} = \frac{3\sigma_0}{2\pi l}\int_0^{v}\frac{dv_z}{v}\frac{2mv_\perp}{p_F}\int_0^1 dt\frac{t^2}{1-t^2} \times \int_{-d/2}^{d/2}\exp\left(-\frac{\nu|x-x'|}{v_\perp\sqrt{1-t^2}}\right)\mathcal{E}(x')dx'. \tag{22}$$

We change the order of integrations with respect to t and x. In this case, the internal integral assumes the form

$$I(\alpha) = \int_0^1 dt\frac{t^2}{1-t^2}\exp\left[-\frac{\alpha}{(1-t^2)^{1/2}}\right]. \tag{23}$$

This integral is equal to the Macdonald function $K_0(\alpha)$ to within exponentially small terms. The quantity α is defined as

$$\alpha = \frac{\nu|x-x'|}{v_\perp} \lesssim \frac{\nu d}{v_\perp}.$$

Everywhere except close neighborhoods of the reference points on the Fermi surface, where the value of v is close to zero, we have

$$\alpha \sim d/l \ll 1.$$

A simple analysis shows that the contribution to the conductivity of flying electrons from a narrow region in the vicinity of the reference points on the Fermi sphere, where $\alpha \gtrsim 1$, is small as compared to the contribution of the remaining part of the Fermi surface in parameter $r^{-3/2}$. Generally speaking, only the neighborhood of the extremal cross section of the Fermi surface, where $v_\perp \approx v$, is significant in the outer integral in Eq. (21). Consequently, the conductivity operator of flying electrons can be written in the form

$$j_{fl}(x) = \frac{3\sigma_0}{\pi l}\int_{\pi l}^{v}\frac{dv_z}{v}\frac{mv_\perp}{p_F}\int_{-d/2}^{d/2}K_0\left(\frac{|x-x'|}{l}\right)\mathcal{E}(x')dx'. \tag{24}$$

Since the argument of the Macdonald function in this equation is a small quantity, we can use the following asymptotic form:

$$K_0(t) = -\left\{\ln\frac{t}{2} + \gamma\right\}, \quad t \longrightarrow 0,$$

where $\gamma \approx 0.577216$ is the Euler constant. In addition, it should be noted that the double integral in Eq. (24) is reduced to the iterated integral. The outer integral with respect to v_z can be evaluated and is equal to $\pi/4$. Consequently, we ultimately have

$$j_{fl}(x) = -\frac{3\sigma_0}{4l}\int_{-d/2}^{d/2}\left[\ln\frac{|x-x'|}{2l} + \gamma\right]\mathcal{E}(x')dx'. \tag{25}$$

Let us now write the asymptotic form of the conductivity operator for trapped electrons. In the main approximation in $r^{-1} \sim \nu\tau \ll 1$, formula (14) gives

$$j_{tr}(x) = \frac{3\sigma_0}{\pi l}\int_0^{v}\frac{dv_z}{v} \times \int_{-mv_\perp-(e/c)A(x)}^{p_y^*}\frac{dp_y}{p_F}\frac{v_y(x)}{|v_x(x)|}\frac{1}{\nu T}\int_{-x^*}^{x^*}dx'\frac{v_y(x')}{|v_x(x')|}\mathcal{E}(x'). \tag{26}$$

In contrast to the conductivity operator for flying electrons, the conductivity operator for trapped electrons is a complex function of the distribution of the magnetic field induced by the transport current.

The simple estimates presented in Section 5 show that the conductivity of trapped particles is higher than the conductivity of flying electrons by a factor of $r^{2/3} \gg 1$. Nevertheless, as will be shown in the next section, the current of flying electrons must be taken into account while solving the Maxwell equations in order to satisfy the boundary conditions (18).

4. TWO-COMPONENT ELECTROMAGNETIC FIELD DISTRIBUTION IN A PLATE

In spite of the complex form of the conductivity operators (25) and (26), a number of conclusions concerning the field distributions in a metal can be drawn simply from a qualitative analysis of the Maxwell equations with the current density $j_{tr} + j_{fl}$, without actually solving these equations. To begin with, we note that formulas (9) and (10) imply that the vector potential $A(x)$ and, hence, the velocities $v_y(x)$ and $v_x(x)$ are even functions of coordinate x. This means that the current density $j_{tr}(x)$ is also an even function of x in the main approximation in parameter r^{-1}. Consequently, if we neglect the contribution of flying electrons, the total current density $j(x) = j_{tr}(x)$ in the Maxwell equations (17) is an even function. The second equation from (17) then implies that the varying magnetic field $h(x)$ must be distributed antisymmetrically. However, this contradicts the boundary conditions. Indeed, in accordance with Eqs. (18), the magnetic field at the boundary $x = -d/2$ must be equal to the preset value h_m, while at the opposite boundary it must practically vanish. The latter statement is the main consequence of the fact that the electric field in a metal is much smaller than the magnetic field in parameter $\omega\delta/c$, which in turn follows from the first of Eqs. (17) (see estimate (57) in Section 5). Therefore, in order to satisfy the boundary conditions (18), the magnetic field $h(x)$ must contain commensurate even and odd components.

This means that, in order to solve the formulated boundary value problem correctly, we must take into account the contribution of flying electrons, which allows us to satisfy both the equations and the boundary conditions.

Let us present the electric field $\mathscr{E}(x)$ of the wave as the sum of symmetric $\mathscr{E}_+(x)$ and antisymmetric $\mathscr{E}_-(x)$ components. It should be observed that the conductivity operator (26) for trapped electrons nullifies the antisymmetric component of the field. On the other hand, the conductivity operator (25) for flying electrons does not nullify this component, and the presence of odd terms in the rf current density is due only to flying electrons.[1] It turns out that the boundary conditions (18) and Eqs. (17) can be satisfied simultaneously only if the antisymmetric component of the electric field is much larger than the symmetric component:

$$\mathscr{E} = \mathscr{E}_+ + \mathscr{E}_-, \quad \mathscr{E}_- \gg \mathscr{E}_+. \tag{27}$$

In order to verify this, we write the Maxwell equations (17) separating even and odd components in all the quantities. In doing so, we take into account the fact that, in view of the specific nature of the conductivity operator (25) for flying electrons (whose kernel depends only on the modulus of the difference $x - x'$), the antisymmetric component of the current density for flying particles is determined only by the antisymmetric component of the electric field, while the symmetric current density component for flying electrons is determined only by the symmetric component of the electric field and is much smaller than the current density (26) for trapped electrons (see Section 5 for details). As a result, we have

$$\begin{aligned} -h'_+(x) &= (4\pi/c)j_-(x) = (4\pi/c)\hat{\sigma}_{fl}\mathscr{E}_-, \\ -h'_-(x) &= (4\pi/c)j_+(x) = (4\pi/c)\hat{\sigma}_{tr}\mathscr{E}_+, \\ \mathscr{E}'_+(x) &= (i\omega/c)h_-(x), \\ \mathscr{E}'_-(x) &= (i\omega/c)h_+(x). \end{aligned} \tag{28}$$

Substituting now the expressions for h_- and h_+ from the third and fourth equations in (28) into the first and second equations, respectively, we obtain the following couple of equations for the symmetric and antisymmetric components of the electric field:

$$\frac{d^2\mathscr{E}_-}{dx^2} = -\frac{4\pi i\omega}{c^2}\hat{\sigma}_{fl}\mathscr{E}_-, \tag{29}$$

$$\frac{d^2\mathscr{E}_+}{dx^2} = -\frac{4\pi i\omega}{c^2}\hat{\sigma}_{tr}\mathscr{E}_+. \tag{30}$$

It can be seen that the penetration depth δ_- for the antisymmetric electric field component $\mathscr{E}_-$ is determined by the relatively low conductivity of flying electrons, while the penetration depth δ_+ for the symmetric electric field component $\mathscr{E}_+$ is determined by the high conductivity of trapped electrons. Since the penetration depth is inversely proportional to the square root of the conductivity, the following inequality holds:

$$\delta_+ \ll \delta_-. \tag{31}$$

The boundary conditions (18) imply that the right-hand sides of the third and fourth equations in (28) are of the same order of magnitude. Consequently, it follows from inequality (31) that the antisymmetric component of the electric field is much larger than the symmetric component; i.e., inequality (27) holds.

Thus, in the strong current mode, when $r \gg 1$, the electric field practically contains only the antisymmetric component. This means that the field of the skin layer is carried by trapped electrons to the opposite face of the plate without attenuation:

$$|\mathscr{E}(-d/2)| = |\mathscr{E}(d/2)|.$$

It is interesting to note that in this case, in accordance with Eq. (29), the penetration depth and the magnitude of the electric field are determined not by trapped, but by flying electrons. This is due to the fact that the conductivity operator (26) for trapped electrons, acting on the purely antisymmetric electric field $\mathscr{E}_-$, gives zero current density. Physically, this is a consequence of the fact that the trajectories of trapped electrons are symmetric relative to the middle of the sample. A trapped particle getting into the skin layer and acquiring a certain energy from the electric field encounters an electric field of opposite sign at the opposite face of the plate and gives away the entire energy acquired by it earlier.

In order to derive an integrodifferential equation for field $\mathscr{E}_-(x)$, we substitute the conductivity operator (25) for flying electrons into Eq. (29). In the kernel of operator (25), we retain only the component antisymmetric

---

[1] Generally speaking, the next terms in the expansion of the conductivity operator (14) for trapped electrons in parameter r^{-1} also possess the same property. However, an analysis shows that these terms have a higher order of smallness as compared to the conductivity operator (25) for flying electrons.

in x' since the symmetric component makes zero contribution to current density. This gives

$$\frac{d^2\mathscr{E}_-}{dx^2} = \frac{3i\pi\omega\sigma_0}{2lc^2}\int_{-d/2}^{d/2}\ln\frac{|x-x'|}{|x+x'|}\mathscr{E}_-(x')dx'. \tag{32}$$

In view of symmetry, we can obviously consider this equation on the interval (0, $d/2$). Let us now consider the boundary conditions (18), which can be written in the form

$$\begin{cases} h_+(-d/2) + h_-(-d/2) = h_m, \\ h_+(d/2) + h_-(d/2) = \mathscr{E}_+(d/2) + \mathscr{E}_-(d/2). \end{cases}$$

Using the fact that

$$\mathscr{E}_+ \ll \mathscr{E}_- \sim (\omega\delta_-/c)h_+,$$

we present this system in the form

$$\begin{cases} h_+(d/2) - h_-(d/2) = h_m, \\ h_+(d/2) + h_-(d/2) = O[(\omega\delta_-/c)h_+]. \end{cases}$$

It follows hence that

$$h_+(d/2) = -h_-(d/2) = h_m/2 \tag{33}$$

to within terms of the order of $(\omega\delta_-/c)h_m \ll h_m$. Using the last equation from (28), we can write this condition for the electric field $\mathscr{E}_-$:

$$\left.\frac{d\mathscr{E}_-}{dx}\right|_{x = d/2} = \frac{i\omega}{2c}h_m. \tag{34}$$

The second boundary condition for Eq. (32) is the condition of vanishing of the antisymmetric component $\mathscr{E}_-(x)$ at the origin $x = 0$:

$$\mathscr{E}_-(0) = 0. \tag{35}$$

Let us introduce the dimensionless coordinate ξ and the dimensionless field $f(\xi)$:

$$\xi = 2x/d, \quad \mathscr{E}_-(\xi) = f(\xi)\frac{i\omega d}{4c}h_m. \tag{36}$$

In the new notation, our boundary value problem assumes the form

$$\begin{gathered} f''(\xi) = -\beta\int_0^1 \ln\frac{|\xi-\xi'|}{\xi+\xi'}f(\xi')d\xi', \\ f(0) = 0, \quad f'(1) = 1. \end{gathered} \tag{37}$$

Here, parameter β is defined as

$$\beta = \frac{3i\pi\omega\sigma_0 d^3}{16c^2 l} \equiv \frac{d^3}{\delta^3} \gg 1. \tag{38}$$

The last equality defines the quantity

$$\delta = \left(\frac{16ic^2 l}{3\pi\omega\sigma_0}\right)^{1/3}. \tag{39}$$

This quantity obviously has the meaning of the complex penetration depth δ_- and coincides in order of magnitude with the result which will be obtained below from a qualitative analysis (see formula (55)).

5. QUALITATIVE MODEL OF THE ANOMALOUS PENETRATION EFFECT

The boundary value problem considered in Sections 3 and 4 is rather cumbersome and leads to a complex integrodifferential equation (37). In many problems of this type, a qualitative method based on the Pippard concept of noneffectiveness [10] was found to be quite useful. In accordance with this concept, the main contribution to the rf conductivity of a metal comes from electrons that stay in the skin layer for most of the time. In order to estimate the conductivity of the main groups of charge carriers, the so-called effective conductivity is introduced, which is calculated for each group of electrons by the formula

$$\sigma_{\text{eff}} = \frac{N_{\text{eff}}e^2\tau_{\text{eff}}}{m}. \tag{40}$$

Here, N_{eff} and τ_{eff} are the concentration of "effective" electrons and the time of their residence in the skin layer, respectively. The essence of the noneffectiveness concept is that the integral conductivity operators (25) and (26) for flying and trapped electrons are simply replaced by the operators of multiplication by the corresponding effective conductivity.

Let us begin by estimating the conductivity of flying carriers (i.e., estimating, in fact, the right-hand side of Eq. (29)). As in Section 3, we disregard the bending of the trajectories of flying electrons by the magnetic field by virtue of inequality (19). In this case, the group of electrons incident on the surface at small angles $\varphi \sim \delta_-/l$ and spending the entire mean free time in the skin layer proves to be effective. For the effective conductivity of such carriers, we obtain the following qualitative formula:

$$\sigma_{fl} = \sigma_0\frac{\delta_-}{l}, \tag{41}$$

where $\sigma_0 = Ne^2/m\nu$ is the conductivity of a bulk sample (we disregard here spatial dispersion, i.e., the complex correction $-i\omega$ to ν in view of relation (8)). Except for the logarithmic factor, this expression coincides with the conductivity estimate obtained under the conditions of the conventional anomalous skin effect in a semi-infinite sample in zero magnetic field.

Let us now estimate the conductivity of trapped charge carriers (i.e., the operator on the right-hand side of Eq. (30)). We assume that the magnetic field of the current is strong enough and the magnetodynamic nonlinearity is noticeable, $r \gg 1$. The angle at which trapped electrons enter the skin layer is equal to $\psi \sim (\delta_+/R)^{1/2}$ in order of magnitude. The length of the arc described by a trapped electron in the skin layer is $L \sim (R\delta_+)^{1/2}$. Consequently, the time spent by the electron in the skin layer over a period of its oscillatory motion is $(R\delta_+)^{1/2}/v$. We must now take into account multiple returns of electrons to the skin layer. Since the probability of a return is equal to $\exp[-2(Rd)^{1/2}/l]$, the effective time of the electron residence in the skin layer can be estimated as

$$\begin{aligned}\tau_{tr} &= \frac{(R\delta_+)^{1/2}}{v}\left(1 + 2\sum_{n=1}^{\infty}\exp[-2n(Rd)^{1/2}/l]\right)\\ &= \frac{(R\delta_+)^{1/2}}{v}\coth[(Rd)^{1/2}/l].\end{aligned} \tag{42}$$

The factor 2 appears in front of the sum in this relation due to the fact that the time of the first residence of an electron in the skin layer is on the average half as long as the next residence periods. Consequently, the following formula is valid for the effective conductivity of trapped carriers:

$$\begin{aligned}\sigma_{tr} &= \frac{Ne^2}{m}\left(\frac{\delta_+}{R}\right)^{1/2}\frac{(R\delta_+)^{1/2}}{v}\coth\left[\frac{(Rd)^{1/2}}{l}\right]\\ &= \sigma_0\frac{\delta_+}{l}\coth\left[\frac{(Rd)^{1/2}}{l}\right].\end{aligned} \tag{43}$$

In the strong-nonlinearity mode ($r = l/(Rd)^{1/2} \gg 1$), this expression can be written as

$$\sigma_{tr} = \sigma_0\frac{\delta_+}{(Rd)^{1/2}}. \tag{44}$$

Having obtained the estimates for the conductivities of flying and trapped electrons, we can now analyze qualitatively the Maxwell equations (28) together with the boundary conditions (18) and get correct order-of-magnitude estimates for the fields $\mathcal{E}_\pm$ and penetration depths $\delta_\pm$.

Let us begin with integrating the first two equations from (28) with respect to x from 0 to $d/2$, taking into account relations (33). This gives the following order-of-magnitude estimates:

$$(4\pi/c)j_-\delta_- = -h_m/2, \tag{45}$$

$$(4\pi/c)j_+\delta_+ = h_m/2, \tag{46}$$

where currents $j_\pm$ are defined as

$$j_+ = \sigma_{tr}\mathcal{E}_+, \tag{47}$$

$$j_- = \sigma_{fl}\mathcal{E}_-. \tag{48}$$

The effective conductivities σ_{fl} and σ_{tr} appearing in these formulas are defined by relations (41) and (44). The third and fourth equations in (28) assume the form

$$\frac{\mathcal{E}_+}{\delta_+} = \frac{i\omega}{c}h_-, \tag{49}$$

$$\frac{\mathcal{E}_-}{\delta_-} = \frac{i\omega}{c}h_+. \tag{50}$$

The six equations (45)–(50) and the two relations (33) form a system containing eight unknowns: $\delta_\pm$, $j_\pm$, $h_\pm$, and $\mathcal{E}_\pm$. Thus, the solution of this system makes it possible to estimate all the quantities we are interested in. For example, Eqs. (49), (50) and relations (33) lead to

$$\frac{\mathcal{E}_+}{\delta_+} = -\frac{\mathcal{E}_-}{\delta_-}. \tag{51}$$

Here, $\mathcal{E}_+$ and $\mathcal{E}_-$ have opposite signs. At the same time, Eqs. (45) and (46) imply that

$$j_-\delta_- = -j_+\delta_+.$$

Substituting into this relation the values of $j_\pm$ from Eqs. (47) and (48) with conductivities (41) and (44), we obtain

$$\sigma_0\frac{\delta_-^2}{l}\mathcal{E}_- = -\sigma_0\frac{\delta_+^2}{l}r\mathcal{E}_+.$$

Using relation (51), we now obtain

$$\begin{aligned}\delta_- &= \delta_+r^{1/3},\\ |\mathcal{E}_-| &= |\mathcal{E}_+|r^{1/3}.\end{aligned} \tag{52}$$

Thus, the qualitative analysis proves that the symmetric and antisymmetric electric field components have different scales and vary over different distances, which completely agrees with the results obtained in Section 4.

Further, from Eqs. (46), (47), and (49), we obtain the equality

$$\frac{4\pi}{c}\delta_+^2\frac{\sigma_0}{l}r\mathcal{E}_+ = -\frac{c}{i\omega}\frac{\mathcal{E}_+}{\delta_+}.$$

It follows hence that the penetration depth δ_+ is given by

$$\delta_+ = \left(\frac{c^2l}{4\pi i\omega\sigma_0 r}\right)^{1/3} \sim \delta_a r^{-1/3}, \tag{53}$$

where

$$\delta_a = \left(i\frac{c^2l}{3\pi^2\sigma_0\omega}\right)^{1/3} \tag{54}$$

is the skin depth in the conventional anomalous skin effect in the absence of transport current. Using formulas (52), we arrive at

$$\delta_- \sim \delta_a, \tag{55}$$

$$\mathcal{E}_+ \sim -\frac{i\omega}{c}\frac{h_m}{2}\delta_a r^{-1/3}, \quad \mathcal{E}_- \sim \frac{i\omega}{c}\frac{h_m}{2}\delta_a. \tag{56}$$

The ratios of the electric and magnetic field components (e.g., $\mathcal{E}_-/h_+$) can be obtained from Eqs. (28):

$$\left|\frac{\mathcal{E}_-}{h_+}\right| \sim \frac{\omega|\delta_-|}{c} \sim \left(\frac{\omega^2 l}{c\sigma_0}\right)^{1/3}. \tag{57}$$

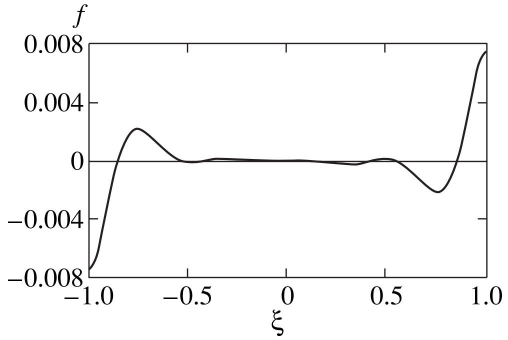


Fig. 2. Distribution of the dimensionless electric field $f(\xi)$ (see (36)) in a plate, which is obtained as a result of numerical solution of Eq. (58) for the ratio $d/|\delta_-| = 10$.

This ratio is always small owing to large values of the conductivity δ_0 for pure metals.

Formulas (56) solve, in principle, the formulated problem, providing a qualitative estimate for the electric field carried by trapped electrons from the skin layer to the opposite face of the sample. However, a more precise description of the spatial distribution of the electric field in a plate under the conditions of anomalous transparency requires a numerical solution of the problem, which will be given in the next section.

6. NUMERICAL SOLUTION OF THE EQUATION FOR SPATIAL DISTRIBUTION OF ELECTRIC FIELD

After certain transformations, the integrodifferential equation (37) can be reduced to Fredholm's integral equation of the second kind, in which the boundary conditions have already been taken into account:

$$f(\xi) - \beta\int_0^1 K(\xi, \xi')f(\xi')d\xi' = \xi, \tag{58}$$

where the kernel $K(\xi, \xi')$ is given by

$$\begin{aligned}K(\xi, \xi') &= \xi[(1-\xi')\ln(1-\xi') - (1+\xi')\ln(1+\xi')] \\ &+ \frac{1}{2}(\xi+\xi')^2\ln(\xi+\xi') - \frac{1}{2}(\xi-\xi')^2\ln|\xi-\xi'| - \xi\xi'.\end{aligned} \tag{59}$$

Since the magnitude of the spectral parameter β is much greater than unity, the series of successive approximations diverges. For this reason, the equation was solved by the discretization method. Figure 2 shows a typical distribution of the electric field $\mathcal{E}_-$ obtained as a result of numerical solution of Eq. (58). It can be seen that the order of magnitude of the skin depth coincides with the quantity $|\delta_-|$ defined by formula (39), which in turn coincides with that obtained in Section 5. The distributions of the symmetric and antisymmetric components of electric and magnetic fields in a metal as well as the current density components are shown schematically in Fig. 3.

Knowing the electric field distribution, we can calculate the quantity characterizing the electromagnetic response of the plate, viz., surface impedance. It is defined as

$$Z = \frac{4\pi}{c}\frac{\mathcal{E}(-d/2)}{h_m}. \tag{60}$$

In the main approximation in parameter r^{-1}, in which only the antisymmetric electric field component $\mathcal{E}_-$ is taken into account, we obtain the following formula for impedance using the dimensionless characteristics ξ and $f(\xi)$ (see formulas (36)):

$$Z = -\frac{4\pi}{c}\frac{\mathcal{E}_-(d/2)}{h_m} = -\frac{4\pi}{c}\frac{i\omega d}{c}f(\xi = 1). \tag{61}$$

For $I = 0$, we are dealing with the conventional anomalous skin effect in zero magnetic field. In this case, the complex penetration depth δ_a for the field of the wave is given by formula (54). The expression (39) for the skin depth implies that the penetration depth in the strong nonlinearity mode is of the same order of magnitude as in the absence of transport current.

The first equation from (17) shows that the surface impedance (60) can be presented in the form

$$Z = \frac{4\pi i\omega}{c^2}\frac{\mathcal{E}(-d/2)}{\mathcal{E}'(-d/2)}. \tag{62}$$

Consequently, the quantity Z can be estimated by order of magnitude as

$$Z \sim -4\pi i\omega\delta/c^2. \tag{63}$$

Thus, the surface impedance under the strong nonlinearity conditions (for strong currents) is of the same order of magnitude as for $I = 0$. However, in the intermediate range (for $r \sim 1$), the surface impedance as a function of current I may vary significantly.

7. CYCLOTRON RESONANCE AT TRAPPED ELECTRONS

In the previous sections, we considered the anomalous penetration of the electromagnetic field of a wave in a metal, associated with a group of trapped electrons. The phenomenon of anomalous penetration of a wave into a metal is closely related to another remarkable effect, viz., cyclotron resonance. In a uniform magnetic field, the cyclotron resonance emerges due to periodic returns of electrons in the skin layer provided that the wave frequency is a multiple of the Larmor frequency of an electron moving in a cyclotron orbit.

In the problem under investigation, trapped electrons experience a similar periodic return in the skin layer. They oscillate relative to the plane $x = 0$ with a period $2T$ (15). Consequently, if the period $2T$ of oscillations of a trapped electron is a multiple of the wave period $2\pi/\omega$, we can expect a resonance similar to the ordinary cyclotron resonance in a uniform magnetic field.

For the sake of simplicity, we assume that both faces of the sample are exposed symmetrically to a radiation in a magnetic field. In this case, the initial distribution of the electric field over the sample thickness is antisymmetric and localized near the plate boundaries. The resonance condition in such a formulation of the problem is that an electron must be accelerated after getting into each skin layer. This means that the half-period T of oscillations of a trapped electron must be equal to an odd number of half-periods of the wave, $(2n + 1)\pi/\omega$, since its phase must change by π during the motion of the electron through the plate in the case of the antisymmetric distribution of the electric field.

The cyclotron resonance at trapped electrons can be observed only if a number of important conditions are satisfied. The first of these conditions is that the half-period T of trapped electron oscillations must be much smaller than the charge carrier relaxation time τ. In other words, the characteristic length $(Rd)^{1/2}$ of an arc on the electron trajectory must be much smaller than the mean free path l. This requirement obviously coincides with the strong nonlinearity condition $r \gg 1$. The second important condition for observing the resonance is the requirement of strong time dispersion. This follows from the fact that the wave period $2\pi/\omega$ under the resonance conditions must be of the order of the period of oscillations of a trapped electron. It was mentioned above that the half-period of oscillations of trapped carriers in the strong nonlinearity mode is much shorter than the relaxation time τ. This means that, under cyclotron resonance conditions, the inequality $\omega \gg \nu$ opposite to inequality (8) must hold. The inclusion of time dispersion necessitates the replacement of the relaxation frequency ν by $\nu^* = \nu - i\omega$ in expression (14) for the conductivity operator. Since the imaginary components dominate in the complex-valued quantity ν^*, the asymptotic forms of the current density and field distributions in the plate under cyclotron resonance conditions differ significantly from those considered by us earlier.

We will analyze the predicted effect only by using a qualitative model. The conductivity of trapped electrons will be estimated using formula (40). The effective relaxation time is described by the same formula (42) in which, however, we must replace ν by ν^*. For the effective conductivity, we have formula (43) in which the quantity l is replaced by $l^* = V/(\nu - i\omega)$.

The complex penetration depth δ satisfies the equation

$$\delta^2 = \frac{ic^2}{4\pi\sigma_{tr}\omega}. \qquad (64)$$

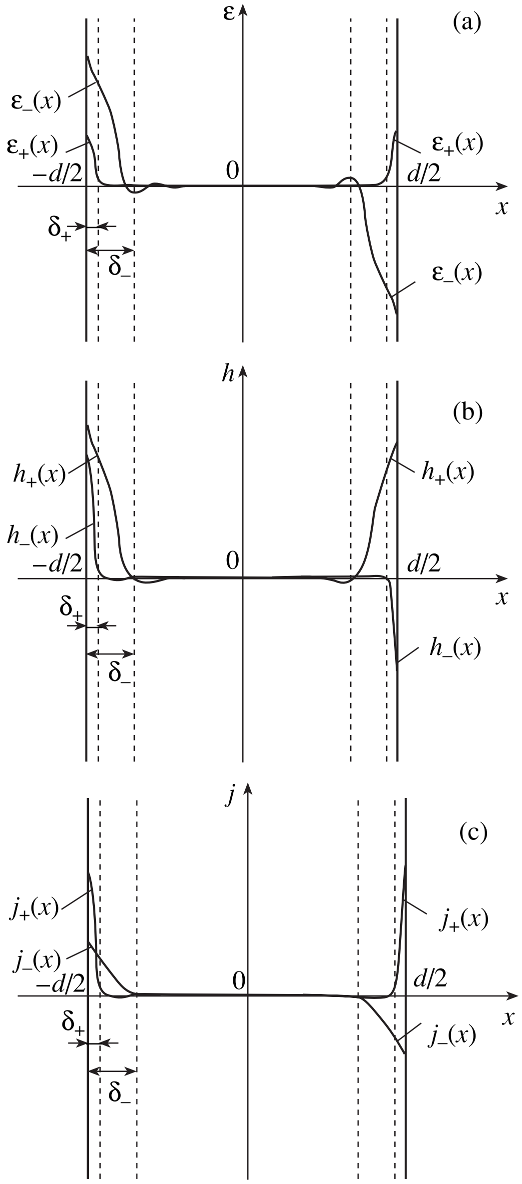


Fig. 3. Schematic diagram of the spatial distribution of even and odd components of (a) electric field, (b) magnetic field, and (c) current density.

This gives

$$\delta^3 = \frac{ic^2}{4\pi\sigma_0\omega} l^* \tanh\left[\frac{(Rd)^{1/2}}{l^*}\right] = \delta_n^2 l^* \tanh\left[\frac{(Rd)^{1/2}}{l^*}\right] = \delta_a^3 \frac{l^*}{l} \tanh\left[\frac{(Rd)^{1/2}}{l^*}\right], \qquad (65)$$

where

$$\delta_n = \left(\frac{c^2}{4\pi\sigma_0\omega}\right)^{1/2} \exp\left(\frac{i\pi}{4}\right), \quad \delta_a = (\delta_n^2 l)^{1/3}$$

is the penetration depth for the normal and anomalous skin effects in a semi-infinite sample in the absence of current (see formula (54)). In order to calculate the surface impedance, we will use formula (63) which is transformed, after the substitution of δ from relation (65), into

$$Z = Z_a\left(\frac{l^*}{l}\right)^{1/3}\tanh^{1/3}\left[\frac{(Rd)^{1/2}}{l^*}\right]. \qquad (66)$$

Here, Z_a is the surface impedance of a bulk metal under the conditions of the conventional anomalous skin effect. It was mentioned above that a strong time dispersion $\omega \gg \nu$ is required for observing the resonance. Taking this circumstance into account, we can write Eq. (66) in the form

$$Z = Z_a\left(\frac{i\nu}{\omega}\right)^{1/3} \times\left[\frac{\nu T\cos(\omega T/2) - i\sin(\omega T/2)}{\cos(\omega T/2) - i\nu T\sin(\omega T/2)}\right]^{1/3}. \qquad (67)$$

This formula provides a qualitatively correct description of the cyclotron resonance in metals with a cylindrical Fermi surface whose axis is directed along the magnetic field. In metals with a convex Fermi surface of the elliptical type, the resonance is blurred since the radius of curvature of electron trajectories and, hence, the frequency T^{-1} with which trapped electrons get into the skin layer are functions of the momentum component p_z. The main contribution to the conductivity of such metals is determined by the relatively small number of electrons located in the vicinity of the central cross section of the Fermi surface, for which the quantity T as a function of p_z has an extremum. As a result, the conductivity under the resonance conditions decreases by a factor of $(\nu T)^{-1/2} \sim r^{1/2} \gg 1$ as compared to the case of a cylindrical Fermi surface. Accordingly, the surface impedance (67) at the resonance acquires the additional factor $r^{1/6}$.

8. CONCLUSIONS

The anomalous penetration of an electromagnetic wave into a metallic current-carrying plate considered in this work is an example of the possible effect of the magnetodynamic nonlinearity on the transparency of metallic plates. This effect is similar in many respects to the well-known penetration of a field along a chain of electron trajectories in a constant uniform magnetic field. However, there are fundamental differences between these two effects.

For one, the anomalous penetration of a field into a plate in a uniform magnetic field leads to the so-called rf size effect (see, for example, [3]). If we denote by $2R$ the diameter of an electron orbit (for the sake of simplicity, we assume that the Fermi surface is cylindrical), the nth spike of the field occurs at a depth $2nR$ from the surface. If the resonance condition $2nR = d$ is satisfied, the spike appears at the opposite face of the plate and the sample becomes transparent to the incident wave. Naturally, if the Fermi surface is not cylindrical, the resonance is blurred and the role of $2R$ is played by the diameter of the orbit corresponding to the extremal cross section of the Fermi surface. However, the anomalous transparency which is in resonance in the magnetic field is preserved in this case also. In addition, the rf size effect is of a clearly manifested threshold type. Indeed, the necessary condition for transparency in the case of closed orbits in a uniform magnetic field is that at least one orbit must fit into the sample thickness, which means that the magnetic field much exceed a certain threshold value. In our situation of field penetration due to oscillating electron orbits, the sample is transparent for any values of current corresponding to $r \gtrsim 1$. The number of trapped electrons and, hence, the conductivity are smooth functions of the current irrespective of the geometry of the Fermi surface. Thus, in the case when the field is carried by trapped electrons, the sample is transparent for any values of the transport current and for any shape of the Fermi surface.

In order to observe the effect of anomalous penetration distinctly, the strong nonlinearity conditions (12) must be satisfied. For samples of thickness $d = 10^{-2}$ cm, width $D = 0.5$ cm, the mean free path of charge carriers $l = 10^{-1}$ cm, the electron concentration $N = 10^{23}$ cm^{-3}, and the Fermi momentum $p_F = 10^{-19}$ g cm/s, the strong nonlinearity mode is realized for currents $I \sim 10$ A. The required large values of mean free path are attained only in pure metals at helium temperatures.

It should be noted that the results of this work were obtained under the assumption of diffuse scattering of charge carriers at the sample surface. It is well known, however, that the electron scattering at the metal surface is close to specular reflection in many cases. This hampers the manifestations of a number of magnetodynamic nonlinearity effects. For example, the experimentally observed decrease in the plate resistance under the strong nonlinearity conditions amounts to only about ten percent [8], while the theory [7] predicts a considerable deviation of the IVC for a metallic sample with diffuse faces from Ohm's law. In the case of specular reflection of electrons, the group of surface particles starts playing a significant role. These carriers experience collisions only with one of the plate faces, and their contribution to conductivity, which is negligibly small in the case of diffuse scattering, might be of the order of the contribution from the remaining groups of electrons in the case of specular reflection. In contrast to the static IVC whose shape is determined by all electrons, the anomalous transparency of the plate predicted by us in this work is ensured only by trapped charge carriers which do not collide with the sample surface. For this reason, the mechanism of field penetration in the bulk of a metal must have a low sensitivity to the type of electron scattering at the surface. In all probability, the effect of anomalous field penetration and the cyclotron resonance at trapped electrons considered by us here will be distinguishable in experiments even in the case of a nearly specular reflection of electrons from the surface.

ACKNOWLEDGMENTS

The authors are grateful to L. M. Fisher for fruitful discussions.

REFERENCES

1. M. Ya. Azbel', Zh. Éksp. Teor. Fiz. **39**, 400 (1960) [Sov. Phys. JETP **12**, 283 (1960)].
2. É. A. Kaner, Zh. Éksp. Teor. Fiz. **44**, 1036 (1963) [Sov. Phys. JETP **17**, 700 (1963)].
3. É. A. Kaner and V. F. Gantmakher, Usp. Fiz. Nauk **94**, 193 (1968) [Sov. Phys. Usp. **11**, 81 (1968)].
4. V. T. Dolgopolov, Usp. Fiz. Nauk **130**, 241 (1980) [Sov. Phys. Usp. **23**, 134 (1980)].
5. N. M. Makarov and V. A. Yampol'skiĭ, Fiz. Nizk. Temp. **17**, 547 (1991) [Sov. J. Low Temp. Phys. **17**, 285 (1991)].
6. G. I. Babkin and V. T. Dolgopolov, Solid State Commun. **18**, 713 (1976).
7. É. A. Kaner, N. M. Makarov, I. B. Snapiro, and V. A. Yampol'skiĭ, Zh. Éksp. Teor. Fiz. **87**, 2166 (1984) [Sov. Phys. JETP **60**, 1252 (1984)].
8. I. F. Voloshin, S. V. Kravchenko, N. A. Podlevskikh, and L. M. Fisher, Zh. Éksp. Teor. Fiz. **89**, 233 (1985) [Sov. Phys. JETP **62**, 132 (1985)].
9. S. A. Derevyanko, G. B. Tkachev, and V. A. Yampol'skiĭ, Fiz. Nizk. Temp. **26**, 87 (2000) [Low Temp. Phys. **26**, 64 (2000)].
10. A. B. Pippard, Proc. R. Soc. London, Ser. A **191**, 385 (1947).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 642–648.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 731–739.
Original Russian Text Copyright © 2001 by Ofitserov, Edel'man.

SOLIDS
Electronic Properties

Investigation of the Spectrum of Surface States in Bismuth by Scanning Tunneling Spectroscopy

A. V. Ofitserov and V. S. Edel'man*
Kapitza Institute of Physical Problems, Russian Academy of Sciences, Moscow, 117334 Russia
**e-mail: edelman@kapitza.ras.ru*

Received April 19, 2001

Abstract—Scanning tunneling spectroscopy of trigonal (0001) and "quasitrigonal" surfaces of a twin interlayer on a cleaved face of bismuth is performed. It is found that both surfaces are characterized by surface electron states with spectra exhibiting clearly defined singularities, namely, relatively narrow maxima and minima of the density of states in the energy range of ∓ 1 eV. An analysis of the behavior of the current–voltage characteristics at low (of the order of tens of millivolts) voltages has revealed the existence on the bismuth surface of a two-dimensional layer, in which the density of states of electrons, unlike its anomalously small value in the bulk of bismuth, is of the order of magnitude typical of metals. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The electronic properties of bismuth have been investigated for many decades. The spectrum of conduction electrons in the bulk of metal has been studied in great detail (see review paper [1]). Considerable progress has been made toward the calculation of the band structure [2]: remarkable agreement has been attained with the experimental data in the most sensitive region, i.e., in the vicinity of the Fermi surface, which requires an accuracy of energy calculation within a millielectronvolt. The situation is less favorable in what regards the investigation of the electronic properties of surfaces. In the case of bismuth, from general considerations, one can expect a radical rearrangement of the electron spectrum in the vicinity of the surface. The reason is obvious: the concentration of current carriers in bismuth is low (of the order of 10^{-5} per atom), and their wavelength and Debye screening length exceed considerably the interatomic distance and reach hundreds of angstroms, so that a total rearrangement of the entire system of conduction electrons must occur at such distances from the surface. In view of this, a discussion has been under way for several decades of dimensional quantum phenomena in thin bismuth films and transition to the semiconductor state (see, for example, [3, 4]). For interpreting the results of experiments with thin epitaxial films oriented in the basal trigonal plane, one has to assume the presence of a surface charge, whose magnitude, in terms of volume with the layer thickness equal to the lattice constant along the trigonal axis (≈1.2 nm), exceeds the bulk concentration of both electrons and holes by at least two orders of magnitude. The assumption of the presence of a surface bending of energy bands of the order of 0.1 eV and of a surface charge must be invoked also in treating the reflection of current carriers from the surface during the investigation of their transverse focusing in a magnetic field [5].

The presence of surface states or resonances with an energy 0.4 eV below the Fermi level was evidenced by the results of precision investigations of the photoeffect [6]. However, these measurements are not quite precise (the resolution of the method is only 0.25 eV) and can produce information only about states below the Fermi level.

Therefore, it appears of interest to investigate tunneling spectra of bismuth. Tunneling measurements in sandwich structures were performed even when studying the spectrum of bismuth in the bulk, when the importance of the surface was not yet realized as clearly as it is at present. Note that the results obtained by different researchers are so varied (see the discussion of relevant studies in [1]) as to be hardly worth being treated here. Most likely, of decisive importance is the fact that the surface properties of film structures depend strongly on the technology of their treatment.

The technique of scanning tunneling spectroscopy (STS) consists in measuring the local current–voltage characteristics and involves the use of a scanning tunnel microscope; this technique enables one to investigate a concrete surface, especially, if the samples are prepared under high vacuum. Also, an advantage of the method is that it enables one to investigate the electron spectrum with space resolution up to atomic and combines naturally with scanning tunneling microscopy (STM), whereby its results may be compared with the actual surface structure. In so doing, it is desirable that the investigations be performed at low temperatures in order to avoid thermal broadening of spectral singularities at low energies and reduce the probability of tunneling with phonon emission. In the case of bismuth, this helps avoid complications associated with the thermal

1063-7761/01/9303-0642$21.00 © 2001 MAIK "Nauka/Interperiodica"

motion of the boundaries of atomic terraces [7]. It is very important that the preparation of samples by cleaving crystals *in situ* at low temperatures may be accompanied by the emergence of terraces with straight, almost atomically smooth, boundaries [8], macroscopic twins [8], and twin interlayers of quantized width with ideal boundaries [9, 10]. All of these objects are well defined physically; therefore, the possibilities for investigations are increased.

Some results of STS studies of atomically smooth regions of the trigonal surface of a low-temperature cleaved face of bismuth are given in [11]. It is demonstrated that the electron states at characteristic energies of the order of tens of millielectronvolts from the Fermi level are inhomogeneous along the surface on scales of the order of several interatomic distances, which is indicative of the generation of point defects during the cleaving of crystal.

We managed to perform the simultaneous spectroscopy of two crystal planes, namely, a trigonal (0001) plane (perpendicular to the Γ–T direction in conventional notation [2]) and a twin surface (perpendicular to one of three equivalent Γ–L directions). Their comparison made it possible to relate the singularities of the density of state determined from tunneling spectra to surface, rather than to volume, states. As a result, a direct proof of the existence of surface states was obtained, and the energy parameters of the respective electrons were determined.

2. EXPERIMENTAL SETUP AND PROCEDURE

Samples in the form of long sticks oriented along the [001] axis and sized approximately 1 × 2 × 5 mm were used for investigations. They were cut out in an electric-spark discharge machine from single crystals grown from the melt by the method of [12] using 99.99999% pure starting material. With this degree of purity, the concentration of impurity atoms on the cleavage surface must be of the order of ≈ 1 atom/μm^2. The dislocation density on the cleaved face of crystal, determined by the pits of etching in dilute nitric acid, was approximately 0.05 μm^{-2}. When preparing out a sample, a shallow notch was made in the middle, which defined the position of the cleavage surface. The samples were etched in nitric acid to remove the impure surface layer and rinsed in distilled water.

The investigations were performed in a low-temperature setup [13] using a scanning tunnel microscope [14]. The STM is characterized by the possibility of computer-controlled displacement of the tip in three directions with a step ranging from fractions of a micron to several microns, which enables one to select the surface region being investigated within typical sizes of samples of the order of 1–2 mm. Thereby, tens of regions may be investigated in a single cryogenic experiment, or, in fact, tens of different samples. Unlike the design described in [14], piezoinertial step motors were replaced by a single unit of three-coordinate displacement based on the effect of electrodynamic forces, which resulted in a higher vibration resistance and reduced noise due to mechanical shaking.

A sample placed in the microscope could be cleaved *in situ* at low temperatures under high vacuum or in a medium of heat-transfer gas such as helium and heated to a temperature of hundreds of kelvin. The measurements were performed at the sample temperature in the vicinity of the boiling point of helium.

STM images and current–voltage characteristics (CVC) were obtained under control of a computer using an ADC–DAC plate manufactured by the L-card company (Moscow, Russia) (this plate provides for digital feedback). In order to improve the signal-to-noise ratio, each CVC was recorded repeatedly and the results were averaged. Usually, 25 to 100 storing operations were performed, and the recording of a single CVC took from 5 to 20 s. The currents being measured were of the order of 1 nA, and the noise was of the order of 1–2%.

The step of displacement of the tip and the number of steps were preassigned when taking a set of CVCs along a line on the surface. At every step, the coordinate z (perpendicular to the surface) was recorded, and the voltage dependence of tunnel current was measured within the preassigned limits. As a result, an array of points was obtained, which reflected the surface relief along the selected line, and the respective CVCs. These data were subsequently processed using the ORIGIN computer codes which enabled one to perform mathematical processing of data, including the differentiation of the CVCs with smoothing. The resolution of the singularities of a CVC was defined by the number of points on the CVC curve (usually, 128 or 256) and by the smoothing parameter which was selected from the compromise between adequate suppression of high-frequency noise and slight broadening of narrow lines.

3. MEASUREMENT RESULTS

The observation of twins on a cleavage surface of bismuth using a scanning tunnel microscope was reported in [8–10]. A scheme of twinning is given in Fig. 1 [15]. When cleaving a crystal with twins (or during production of a twin in the process of cleaving), two nonequivalent planes are opened. One of these is a trigonal plane perpendicular to the Γ–T direction of the Brillouin band, and the other plane is perpendicular to one of three equivalent Γ–L directions. For simplicity, we will refer to such planes as quasitrigonal, because three out of four trigonal planes make a transition to these planes in the case of rhombohedral deformation transforming a cubic into bismuth lattice.

Twins occur quite rarely on STM images, one or two cases per hundred of examined frames sized 1 × 1 μm^2, and, apparently, in some cases they occur as a result of mechanical effect on the surface during incidental con-

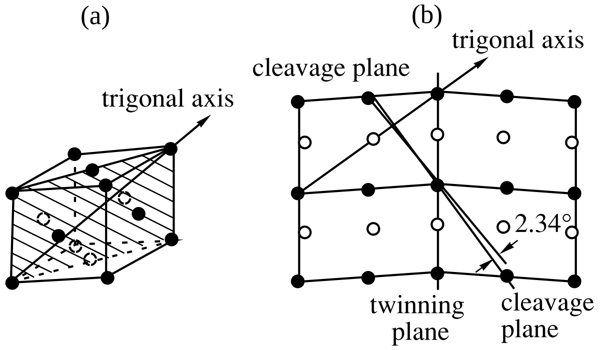


Fig. 1. (a) The scheme of one of sublattices of bismuth; (b) the scheme of arrangement of atoms in a plane of the type of the hatched plane in (a) upon twinning.

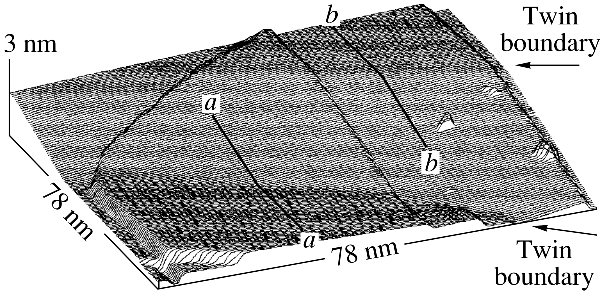


Fig. 2. An STM image of a region of a cleaved face of bismuth with a twin interlayer: (*a–a*) and (*b–b*) lines along which sets of current–voltage characteristics were registered. One can see terraces of diatomic height and (in the lower right-hand corner) point defects.

tact with the tip of the microscope. Otherwise, it would be hard to conceive how to reach the immediate vicinity of the point of termination of a twin interlayer of submicron width (see Fig. 10 in [8]). Apparently, this was the case in our study as well.

The twin interlayer was detected under conditions of heating to room temperature and cooling a sample cleaved *in situ* at helium temperature for the purpose of tracing the transformation of the terrace boundaries, after the tip touched the sample. The image of a surface with twin interlayer is shown in Fig. 2. Its appearance is quite classical: the horizontal region changes to a flat inclined one, and then to horizontal again, at a different level. The angle between the planes differs from 180° by ±2.4°, with an accuracy of approximately 0.2° in accordance with the scheme of Fig. 1. The interlayer is wedge-shaped, with the wedge angle in projection onto the plane being 10°–12°. Both boundaries are inclined ±(5°–6°) to the directions of the terrace boundaries, which coincide with the direction of atomic series on the surface [8]. Therefore, neither of them is coherent.

One can find out which planes are planes of the (0001) type by comparing the heights of steps on the terrace boundaries that are well seen in Fig. 2. Their values, according to crystallographic data, must be equal to 0.395 and 0.374 nm on the trigonal and quasitrigonal planes, respectively. The average values of height, measured for different steps at several points, were 0.395 × 0.01 nm for the horizontal regions in Fig. 2 and 0.375 × 0.01 nm for the inclined plane region. Thereby, the orientation of the interlayer may be regarded to be reliably determined. (Note that it was by the height of step on the (0001) plane that the absolute calibration with respect to size was performed; however, this cannot affect the determination of orientation, because the knowledge of relative values is sufficient for this purpose.)

Arrays of CVC data were taken during the motion of the Pt + 6% Rh tip of the microscope along the *a–a* and *b–b* lines plotted in Fig. 2. Identical measurements were repeated twice. Three series of measurements were performed, namely, with the voltage U between the sample and tip varying within ±0.65, ±1.2, and ±2.4 V. The tunneling gap was stabilized at a voltage that was maximal for each range; after that, the STM feedback was "frozen," and the tunnel current was measured with decreasing voltage. After reaching the minimal value, the voltage returned abruptly to the initial value, the feedback operation was resumed, and the procedure was repeated.

It turned out that, when removed 2–4 nm from the twin boundary, the CVCs and their derivatives in both trigonal regions coincided with one another within 1–2% and approximately 5–10%, respectively. The same was true of the interlayer. This enabled one to average the CVCs (over approximately 40 curves at low voltages and 20 curves at high voltages) and, thereby, reduce the noise several times over. The thus obtained CVCs and their derivatives are given in Figs. 3 and 4. One can see that they differ considerably for crystallographically different planes, especially, in the voltage range from –0.5 to +1 V. At high voltages, when they become comparable with the work function $W \approx 4$ eV, the current increases rapidly with voltage. In this region, only two diffuse singularities are observed at $U = -1.2$ and +1.7 V, and the CVCs for different surface regions converge with one another.

One can trace how the CVCs transform during transition from the trigonal to quasitrigonal plane by constructing the curves of dependence of tunnel current on the surface coordinate y for several fixed values of voltage. Two such diagrams, in which the $z(y)$ dependences are also plotted, are given in Fig. 5. In constructing the dependences, the values of current for 16 close values of voltage were averaged for reducing the level of high-frequency noise.

Note two interesting facts.

Firstly, the widths of the transition region both for $z(y)$ and for $I(U, y)$ for the concave and convex boundaries differ somewhat and are 2–3 and 4–5 nm, respectively. In the experiments described herein, we failed to attain atomic resolution: apparently, the end of the tip

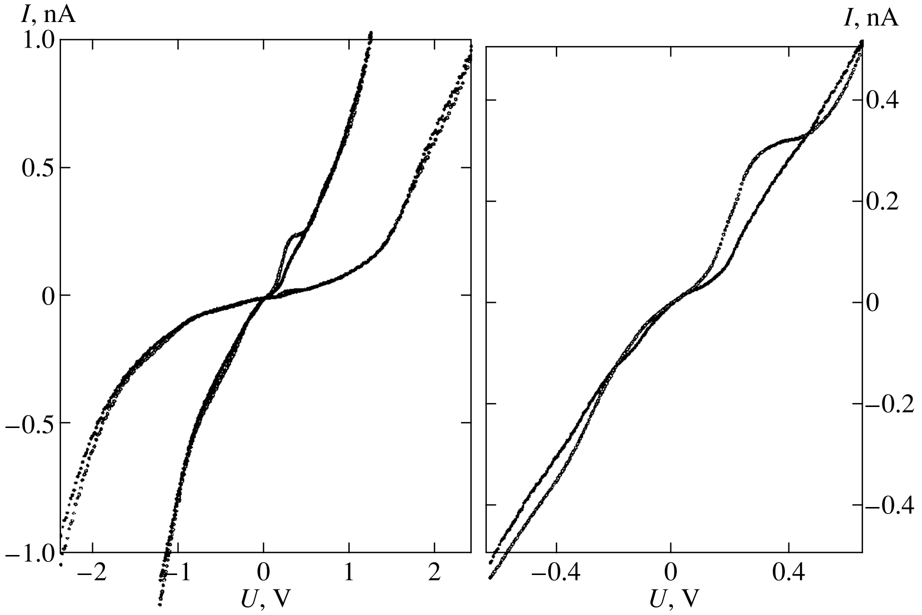


Fig. 3. The averaged current–voltage characteristics of tunneling gap between a sample and the Pt + 6% Rh tip of a microscope, recorded in different ranges of voltage variation. Dark circles indicate a trigonal plane, and light circles indicate a quasitrigonal plane.

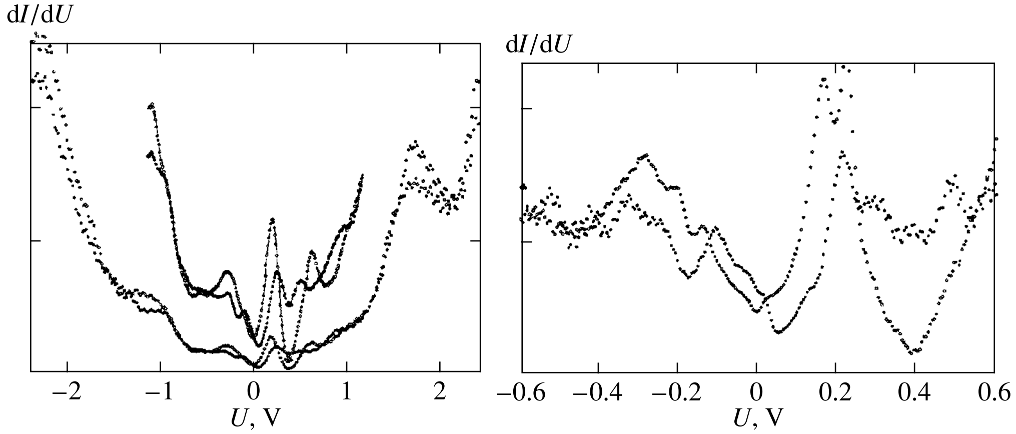


Fig. 4. Differential CVCs obtained by differentiation with smoothing over seven points of the dependences similar to those given in Fig. 3. Dark circles indicate a trigonal plane, and light circles indicate a quasitrigonal plane. Note that, because of the smoothing procedure, the double peak in the right drawing at $U = +0.2$ V appears as the single peak in the left drawing.

from which the tunneling occurred was made up of several atoms. However, point defects were observed on the surface with a width of approximately 0.5 nm at half-height (some of them are seen in Fig. 2). Therefore, the numbers given are not strongly distorted owing to the finite size of the tip. Note further that the width of the region of transition from one flat face to another in observing twin interlayers of quantized width in [9, 10] with atomic resolution on the concave boundary was approximately 1.5 nm, i.e., a value close to that obtained by us.

Secondly, the transition region for currents is shifted to the region of quasitrigonal surface, as one can see especially clearly when studying the convex boundary. In fact, the value of current characteristic of a trigonal plane is maintained until the middle of the transition region, and the value characteristic of a quasitrigonal plane is obtained after shifting to the flat part of the surface.

The spectra for the concave and convex boundaries differ from one another both quantitatively and qualitatively. In the case of a convex boundary on differential CVCs, note the behavior of the peaks of a doublet characteristic of a quasitrigonal plane (Fig. 6a). Its component corresponding to $U = +0.22$ V broadens gradually and decreases with respect to amplitude to shift towards

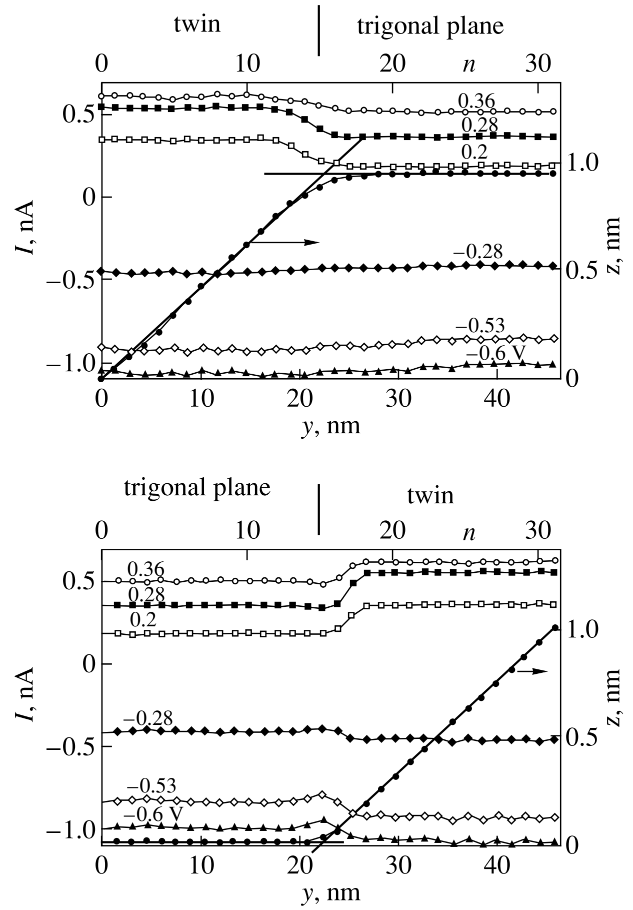


Fig. 5. The tip position *z* and the tunnel current as functions of the coordinate along the surface for different values of voltage (the currents are averaged over 16 close values of voltage). The top horizontal axis indicates the numbers of points along the *y* axis and of the respective CVCs.

higher voltage at a rate of approximately 50 mV/nm and, apparently, changes to a poorly resolved maximum at $U = +(0.32–0.34)$ V for a trigonal plane. At the same time, the component corresponding to the value of +0.17 V also broadens and decreases with respect to amplitude to shift toward a value of +0.2 V (on curve *16*), and changes to a component at $U = +0.22$ V which predominates for a trigonal plane. The peak at 0.7 V changes to a peak at 0.48 V and shifts in the opposite direction at a rate of about –0.1 V/nm.

This is not the end: it is only in the transition region that a small peak emerges at $U = +24$ mV, and a strong rise occurs of the amplitude of the peak at $U = -0.13$ V which is observed on the CVC for the quasitrigonal plane. It appears that this rise is associated with the shift to this voltage of the peak at $U = -0.22$ V which is replaced by a pronounced minimum. After this, one can see the tendency of this peak to shift toward a peak of similar shape on the trigonal plane at $U = -0.11$ V.

The differential CVCs (Fig. 6b) taken in the transition region for the concave boundary exhibit no clearly defined singularities which would not be known to appear on the characteristics away from the boundary, and these differential CVCs appear rather as the sum of CVCs on the trigonal and quasitrigonal planes in some or other proportion.

In conclusion of this section, note that the measurements described above were performed under conditions of true vacuum tunneling, as evidenced by the extent of the shift of the tip of the microscope during the variation of current from 0.1 to 1 nA. The shift with respect to *z*, averaged over a frame of 64 × 64 points, amounted to 0.125 ± 0.02 nm. The expected value with the work function of approximately 4 eV is 0.115 nm, i.e., coincides with the measured value within the error of measurement.

4. DISCUSSION OF THE RESULTS

In analyzing the current–voltage characteristics of a tunneling contact, we will use the formula for tunnel current obtained in [16] and reflecting the qualitative singularities of the effect,

$$I(U) \propto \int_{E_F}^{E_F + eU} \rho(E) T(E, U) dE, \tag{1}$$

where $\rho(E)$ and $T(E, U)$ denote the density of electron states and the transparency of the tunneling barrier, respectively, and E_F is the Fermi energy. At $U \ll W$, the transparency of the barrier does not depend on the applied voltage and, in this region, $dI/dU \propto \rho(E_F + eU)$.

Formula (1) assumes that the density of states for the tip in the energy range being treated exhibits no singularities. It is usually rather difficult to validate this assumption, especially, in view of the fact that the properties of the tip reflect its structure, which is not known. Fortunately, in our case, one can state with assurance that no singularities due to the electron spectrum of the tip are observed on tunneling spectra. This is true because, with the tip being unvaried, the appearances of the differential CVCs for the trigonal and quasitrigonal planes are totally different: either the minima and maxima do not coincide as regards their positions, or they differ considerably in relative magnitude. For example, the value of dI/dU at a low minimum at $U = +0.4$ V for a quasitrigonal plane is approximately four times less than at $U = 0$, and for a trigonal plane it is approximately half as much again.

The most interesting question arising during the investigation of the bismuth surface is whether it possesses metallic conductivity or the surface states are localized. The finite density of states $\rho(E) = dN/dE \propto dI/dU$ on the Fermi level, i.e., at $U = 0$, corresponds to the former situation. For the latter case, $dN/dE = 0$. One can see in Fig. 4 that the derivative of CVC at $U = 0$ for both trigonal and quasitrigonal planes has a value of the same order of magnitude as in the case of other voltages. One must make clear, however, whether this results from the fact that bismuth in the bulk is a metal.

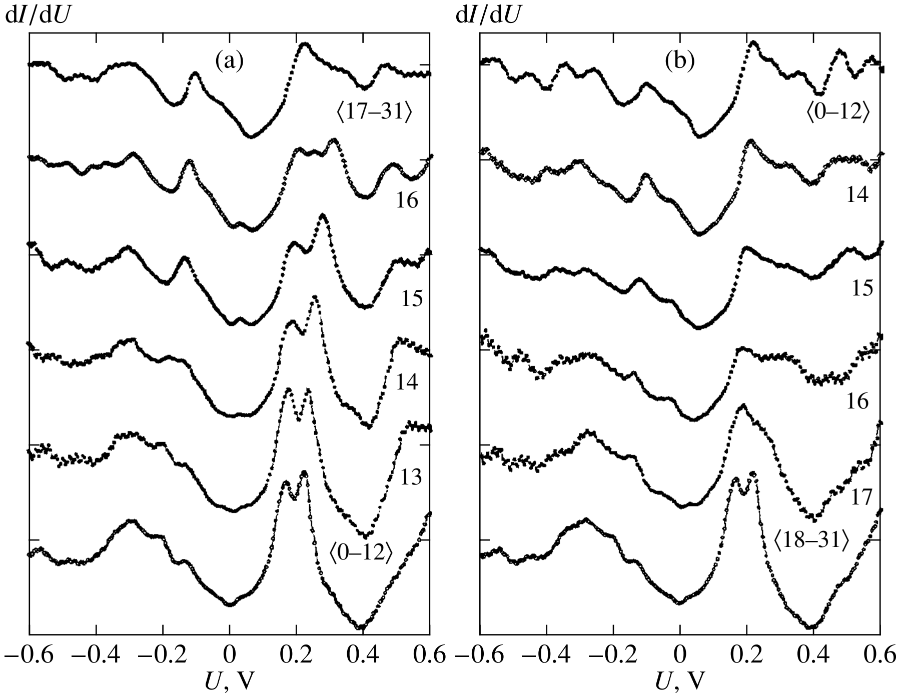


Fig. 6. Differential CVCs corresponding to (a) the top and (b) the bottom part of Fig. 5. The numerals by the curves correspond to those in Fig. 5, and the angle brackets indicate the averaged CVCs with the respective values of n.

We will examine the possible behavior of the density of states under conditions of the well-known spectrum of bismuth electrons in the vicinity of the Fermi level [1]. Bismuth has holes located in the neighborhood of point T of the Brillouin zone with a spectrum that is very close to quadratic,

$$E_{h0} - E_h = P_\perp^2/2m_\perp + P_\parallel^2/2m_\parallel, \qquad (2)$$

where E_{h0} = 12 meV is the energy above the Fermi level; the symbols $\perp$ and $\parallel$ indicate the directions perpendicular and parallel to the trigonal axis, respectively; P denotes components of momentum; and the respective masses are equal to 0.064 and 0.70 of the mass of a free electron. With this spectrum, the density of states must tend to zero at E_{h0} as $(E_{h0} - E_h)^{1/2}$; i.e., a kink must emerge on dI/dU curve at $U = +12$ mV. The derivative of current must increase with decreasing U and, at $U = 0$, reach a value corresponding to the density of states of 1.3×10^{-3}/atom eV.

The conduction electrons are localized in the vicinity of points L of the Brillouin zone, and their Fermi surface consists of three ellipsoids (in a first approximation); however, their spectrum is strongly nonquadratic. In the two-band model of Lax [1], which is quite adequate in the case being treated, this spectrum is written in the principal axes of each of the ellipsoids in the form

$$\frac{E^2 - E_g^2/4}{E_g} = \sum \frac{P_1^2}{2m_i}. \qquad (3)$$

Here, the energy is counted off from the value of –35 meV from the Fermi level (i.e., in our case, from the voltage of –35 mV), and the gap energy is $E_g \approx 13$ meV (the values of masses are not given and can be found in [1]). Therefore, for the electron contribution, one can expect root singularities at $U = -29$ and –41 mV. The current derivative must increase for higher and lower values of voltage; in so doing, at energies in the vicinity of the Fermi level, the density of states varies as $\propto E^2$, because $E_g^2/4 \ll E^2$, and at the Fermi level the density of states is equal to 0.93×10^{-3}/atom eV.

On turning to Fig. 4, one can see that a kink in the derivative is indeed observed in the vicinity of the voltage of +12 mV; however, as the voltage decreases, a decrease in the derivative is observed with U tending to zero, instead of an increase. No singularities are observed in the case of negative values of voltage in the vicinity of $U = -29$ and –41 mV. Therefore, the behavior of a differential CVC is not what one should expect for bulk electrons.

We will now turn to the quantitative aspect. As is seen in Figs. 3 and 4, both the current and its derivative increase rapidly at high voltages. This increase is associated primarily with the variation in the transparency of the tunneling barrier. If we assume that the barrier is trapezoidal in shape, i.e., varies in the direction from the sample to the tip as $(W - eUz/z_0)$, where z_0 is the width of the tunneling gap, it is easy to estimate that

$$T(U) \propto \int\left(W - eU\frac{z}{z_0}\right)^{1/2} dz \approx T(U = 0)\left(1 - \frac{eU}{2W}\right). \quad (4)$$

We substitute reasonable values of the vacuum tunneling gap $z_0 \approx 1$ nm and of the work function $W \approx 4$ eV to find that, at $U = \pm 2$ V, an increase in the transparency of the tunneling barrier leads to a current increase by a factor of 10 to 15. If this is taken into account, one can see in Figs. 3 and 4 that the density of states in the entire investigated voltage range does not differ too much compared with its value at $U = 0$. Because the special smallness of the density of states for bismuth (of the order of 10^{-3} electron/atom eV) is associated with the singularities of the spectrum of bulk electrons only in the vicinity of the Fermi level, while for all other values of energy the density of states must be of an ordinary magnitude of 0.1 to 1 electron/atom eV, we come to a conclusion that the density of states observed by us has a regular order of magnitude characteristic of metals even at $U = 0$. Therefore, in the vicinity of $U = 0$, the bulk electrons must not make any appreciable contribution to the tunnel current.

The foregoing reasoning demonstrates that the observed finite density of states at the Fermi level is associated only with surface states which form a two-dimensional metal with an ordinary electron density. It would be interesting to see their contribution directly to the conduction of a thin plate of bismuth; however, this is not simple to do, because the electron mobility must be several orders of magnitude lower than that of bulk electrons. Firstly, their mass must be of an ordinary magnitude and two orders of magnitude less than that of conduction electrons. Secondly, their free path at low temperatures must be limited by scattering from the terrace boundaries and amount to fractions of a micrometer in contrast to millimeters for bulk electrons.

One can see in Fig. 4 that numerous singularities are observed in the tunneling spectrum. Unfortunately, we are not aware of theoretical calculations devoted to surface states in bismuth. However, it is safe to say that all singularities in the voltage range of approximately –0.5 to +0.5 V are associated with surface states. This statement is based both on the difference of spectra for the trigonal and quasitrigonal planes and on the continuous transition from one to the other on the twin boundary (Fig. 6). Apparently, the maximum at $U = -0.35$ V corresponds to the peak at $U = -0.4$ V observed in [6].

As to the singularities at voltages of 1–2 V, they may correspond to bulk electrons and, in this case, the quantitative difference in the density of states for different planes may be due to the variation in the overlapping of the electron wave functions of the sample and tip with changing orientation. According to Liu and Allen [2], extrema are observed at points Γ and T at $U \approx 1.2$ eV; the plateau observed on the curves in Fig. 4 may correspond to these extrema. The maximum at $U \approx +1.7$ V may have two close bands at this energy at point Γ corresponding to it. However, further away from the Fermi level and with increasing spatial localization of the respective electrons, the difference between the sample bulk and surface must be smoothed out.

Note yet another intriguing result. According to Fig. 6, a continuous shift of parameters is observed for some singularities when crossing the twin boundary. In this case, the interplanar spacing apparently varies (this spacing for a trigonal plane exceeds that for a quasitrigonal plane by 5.6%), and a uniaxial "tension" of the lattice occurs normally to the twinning boundary by 5.8% with the disappearance of third-order symmetry. There is reason to believe that it will be possible to identify the effect of each one of these factors individually, after spectral calculations for this situation are performed.

ACKNOWLEDGMENTS

We are grateful to A.F. Andreev for his interest in our work. This study received partial financial support from the State contract 106-10(00)-P and from the Russian Foundation for Basic Research (project no. 01-02-16711).

REFERENCES

1. V. S. Édel'man, Usp. Fiz. Nauk **123**, 257 (1977) [Sov. Phys. Usp. **20**, 819 (1977)].
2. Y. Lin and R. E. Allen, Phys. Rev. B **52**, 1566 (1995).
3. Yu. F. Komnik, U. I. Bukhshtab, Yu. V. Nikitenko, and V. V. Andrievskiĭ, Zh. Éksp. Teor. Fiz. **60**, 669 (1971) [Sov. Phys. JETP **33**, 364 (1971)].
4. C. A. Hoffman, J. R. Meyer, and E. J. Bartoli, Phys. Rev. B **48**, 11431 (1993).
5. V. S. Tsoĭ and I. I. Razgonov, Pis'ma Zh. Éksp. Teor. Fiz. **23**, 107 (1975) [JETP Lett. **23**, 92 (1976)].
6. G. Jezequel, J. Thomas, and I. Pollini, Phys. Rev. B **56**, 6620 (1997).
7. A. M. Troyanovskiĭ and V. S. Édel'man, Pis'ma Zh. Éksp. Teor. Fiz. **60**, 104 (1994) [JETP Lett. **60**, 111 (1994)].
8. A. M. Troyanovskiĭ and V. S. Édel'man, Kristallografiya **44**, 336 (1999) [Crystallogr. Rep. **44**, 300 (1999)].
9. V. S. Edel'man, Phys. Lett. A **210**, 105 (1996).
10. V. S. Edelman, D. Yu. Sharvin, I. N. Khlyustikov, and A. M. Troyanovskiĭ, Europhys. Lett. **34**, 115 (1996).
11. A. M. Troyanovskiĭ and V. S. Édel'man, Zh. Éksp. Teor. Fiz. **115**, 2214 (1999) [JETP **88**, 1212 (1999)].
12. M. S. Khaĭkin, S. M. Cheremisin, and V. S. Édel'man, Prib. Tekh. Éksp., No. 4, 225 (1970).
13. I. N. Khlyustikov and V. S. Édel'man, Prib. Tekh. Éksp., No. 1, 158 (1996).
14. V. S. Edel'man, A. M. Troyanovskii, M. S. Khaikin, *et al.*, J. Vac. Sci. Technol. B **9** (2), 618 (1991).
15. M. V. Klassen-Neklyudova, *Mechanical Twinning of Crystals* (Nauka, Moscow, 1960; Consultants Bureau, New York, 1964).
16. A. Selloni, P. Carnevali, E. Tosatti, and C. D. Chen, Phys. Rev. B **31**, 2602 (1985).

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 649–656.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 740–748.
Original Russian Text Copyright © 2001 by Vecheslavov, Chirikov.

Separatrix Conservation Mechanism for Nonlinear Resonance in Strong Chaos

V. V. Vecheslavov* and B. V. Chirikov**

Budker Institute of Nuclear Physics, Siberian Division, Russian Academy of Sciences,
pr. akademika Lavrent'eva 11, Novosibirsk, 630090 Russia

**e-mail: vecheslavov@inp.nsk.su*

***e-mail: chirikov@inp.nsk.su*

Received August 21, 2000

Abstract—We propose a simple, approximate theory of the fairly general mechanism for the separatrix conservation of nonlinear resonance, which leads to the complete suppression of global diffusion despite the strong local chaos of motion. This theory allows the separatrix splitting angle to be plotted against system parameters and, in particular, yields their values at which the separatrix remains unsplit. We present the results of our numerical experiments confirming theoretical conclusions for a certain class of dynamical Hamiltonian systems. New features of chaos suppression have been found in such systems. In conclusion, we discuss the range of application of the proposed theory. © *2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION: UNEXPECTED STABILITY OF THE SEPARATRIX OF NONLINEAR RESONANCE

The dynamics of nonlinear Hamiltonian systems is governed by the interaction of nonlinear resonances, with each of them, in contrast to a linear resonance, occupying a relatively small region of phase space bounded by the so-called separatrix under a small perturbation (see, e.g., [1-4]). For a single resonance, the separatrix is the trajectory (in general, a surface) that separates phase oscillations (inside the resonance) from phase rotation (outside the resonance).[1] In fact, these are two spatially coincident branches corresponding to going forward and backward in time, respectively. Each branch is a continuous trajectory with an infinite period of motion that goes out of an unstable equilibrium position (saddle) and then asymptotically approaches it. In a typical (i.e., nonintegrable) Hamiltonian system, any arbitrarily small perturbation, for example, from other (at least one) nonlinear resonances, causes the separatrix to split up into two intersecting branches, which go out of the saddle toward each other as before but no longer return to it. The free ends of the branches of the split separatrix form an infinite number of loops with a limitlessly increasing length; these loops fill a narrow region near the unperturbed separatrix to form the so-called chaotic layer. Overlapping of the chaotic layers of all system resonances gives rise to global chaos and, in particular, to diffusion bounded only by the exact integrals of motion, for example, by a surface of constant energy.

The conditions for the formation of global chaos depend on both the magnitude and smoothness of the perturbation (in phase). The latter is characterized by the rate of decrease in its Fourier amplitudes. For an analytic perturbation, the decrease is exponential. In this case, there is always such a critical perturbation magnitude $\epsilon_{\rm cr}$ that global diffusion emerges only at $\epsilon \gtrsim \epsilon_{\rm cr}$. If, alternatively, $\epsilon \lesssim \epsilon_{\rm cr}$, chaos is localized in relatively narrow chaotic layers that are formed at any $\epsilon > 0$. For $N > 2$ degrees of freedom, global diffusion is still possible, but only for special initial conditions and with a very low rate (the so-called Arnold diffusion [2]). When $\epsilon \longrightarrow 0$, both the rate of diffusion and the measure of its range decrease exponentially in parameter $1/\epsilon$.

The pattern of motion significantly changes for a smooth Hamiltonian perturbation, which has only a finite number γ of continuous derivatives (see, e.g., [5] and references therein). In this case, there is such a critical smoothness $\gamma_{\rm cr}$ that global diffusion is suppressed under a sufficiently small perturbation only at $\gamma > \gamma_{\rm cr}$ [6]. Significantly, the converse is generally not true; i.e., at $\gamma < \gamma_{\rm cr}$, global diffusion is commonly observed in numerical experiments, but we know examples when the trajectory remained localized in a part of phase space over the entire long computational time (see, e.g., [7, 8]).

Recently, Ovsyannikov [9] has found a relatively simple, exactly solvable example [see (2.1) below] for which he managed to prove the theorem on the conservation of a single (unsplit) separatrix at special values of the perturbation parameter. This theorem is given in its entirety in [10] (Appendix). Intensive studies of

---

[1] Below, we use the canonical action-phase variables.

1063-7761/01/9303-0649$21.00 © 2001 MAIK "Nauka/Interperiodica"

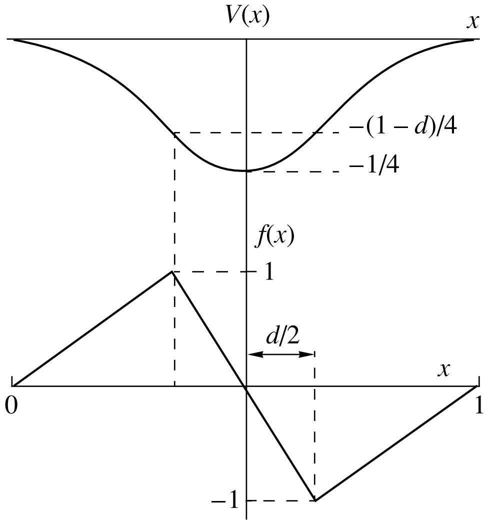


Fig. 1. A scheme of the potential $V(x)$ and force $f(x) = -dV/dx$ with a period of 1 for the family of models (2.2) with parameter d.

model (2.1), which we call below a symmetric piecewise linear mapping, immediately showed that the conservation of the separatrix in strong chaos rather than in the exceptional case of a completely integrable system without any chaos whatsoever turned out to be the most important and unexpected thing in this theorem. Moreover, at the special values of the perturbation parameter found both by Ovsyannikov and by one of us [10–12], the separatrices of nonlinear resonances not only remain unsplit, but form impenetrable barriers for other trajectories; i.e., they completely suppress global diffusion. This takes place despite the fact that the perturbation smoothness in the model of a symmetric piecewise linear mapping is considerably smaller than its critical value, and one might expect global diffusion at any perturbation magnitude.

Meanwhile, an examination of the literature has shown that the same model was mathematically analyzed in detail by Bullett [13] (see also [14]) well before. Although Ovsyannikov proved his theorem independently, this coincidence of the models is not fortuitous, because the solution of a linear (even if piecewise) mapping considerably simplifies the problem. Note that even Ovsyannikov's linear mapping can be completely solved only when the separatrix is conserved, because, otherwise, the two branches of the split separatrix form random trajectories. For the same reason, the model of a symmetric piecewise linear mapping cannot be simplified to a purely linear Arnold-type mapping, in which the separatrices of nonlinear resonances are always split (see Sect. 3). Bullett's and Ovsyannikov's mathematical analyses are therefore restricted to the invariant curves of a new type (with a rational rotation number ν, including the separatrices) themselves, whose first examples were given in [14].

In this paper (as in our previous papers on this subject [10-12]), we rely mainly on numerical experiments, which also allow us to investigate the vicinities of the invariant curves both under various initial conditions of motion and for various model parameters. In this regard, our approach is similar to the study of Hénon and Wisdom [8] for a different model.

2. THE MODEL

Ovsyannikov considered a difference equation that was equivalent to the following two-dimensional mapping in canonical variables "action p–phase x":

$$\bar{p} = p + Kf(x), \quad \bar{x} = x + \bar{p}. \qquad (2.1)$$

Here, $K = \epsilon > 0$ is the perturbation parameter (not necessarily small), and the "force" $f(x)$ has the form of an antisymmetric ($f(-y) = -f(y)$, $y = x - 1/2$) piecewise linear "saw" with a period of 1 [see (2.2) below].

Perhaps, the most unexpected thing in this example is that the smoothness of the Hamiltonian (generating function) for mapping (2.1), $\gamma = 1 < \gamma_{\mathrm{cr}} \le 4$ [6], is considerably smaller than its critical value. In other words, for a certain countable set of special values, $K = K_m$, the unsplit separatrix is "immersed in the sea" of strong chaos; nevertheless, it is conserved and blocks global diffusion [10, 11]!

Since an exact K cannot be specified on a computer, the next crucial step was to analyze the behavior of the (split) separatrix and other trajectories for a small deviation, $|K - K_m| \longrightarrow 0$, which is possible only in numerical experiments. Even the first studies [12] showed that the separatrix splitting angle changed sign with difference $K - K_m$; this angle smoothly passed through zero at odd m and abruptly changed sign at even m (see Fig. 1 from [11] and Fig. 2 below). First, this allowed a set of other special K_m at which the separatrix was conserved to be found immediately and easily. At the same time, such an unusual behavior of the splitting angle also suggested a dynamical mechanism for the conservation of the separatrix, which is the main subject of our discussion here.

It is convenient to simultaneously consider the whole family of sawtooth perturbations specified by the force[2]

$$f(x) = \begin{cases} 2x/(1-d), & \text{if } |x| \le (1-d)/2 \\ -2y/d, & \text{if } |y| \le d/2, \end{cases} \qquad (2.2)$$

where $y = x - 1/2$ and $d < 1$ is the distance between the saw teeth $|f(x)| = 1$ at points $y = y_\pm = \pm d/2$. The best studied special case of a symmetric piecewise linear mapping corresponds to $d = 1/2$. At these two points, the

---

[2] A similar family is briefly mentioned in [13].

force has a singularity, a discontinuity of the first derivative $f' = df/dx$:

$$\Delta f' = \pm\frac{2}{d(1-d)}. \tag{2.3}$$

The original idea behind the separatrix conservation mechanism is that the perturbation (force) has two singularities, which can interfere between themselves and, in particular, can compensate each other or cancel out, a term introduced in [8] where such a mechanism was apparently first proposed and used. Our approach is peculiar in that we are interested in the action of this mechanism (and give a theory) directly for the separatrix of nonlinear resonance, whereas in [8] such cancellations were determined for arbitrary trajectories and used as a heuristic consideration to search for possible invariant curves (not separatrices) among them by numerical experiments.

To construct a theory, it is convenient to pass from the initial mapping (2.1) to a continuous system with a Hamiltonian that explicitly depends on time (see [2–4, 10]):

$$\begin{aligned} H(x, p, t) &= \frac{p^2}{2} + KV(x)\delta_1(t) \\ &= H_0(x, p) + H_1(x, t), \end{aligned} \tag{2.4}$$

where $\delta_1(t)$ denotes the δ function of period 1. The unperturbed Hamiltonian

$$H_0 = \frac{p^2}{2} + KV(x) \tag{2.5}$$

describes the main (integer) resonance in (2.1), and

$$H_1(x, t) = KV(x)(\delta_1(t) - 1) \tag{2.6}$$

describes its perturbation (with the same period $T_1 = 1$ and frequency $\Omega = 2\pi/T_1 = 2\pi$) from all the remaining integer resonances.

The potential of force (2.2) is

$$\begin{aligned} V(x) &= -\int f(x)dx \\ &= \begin{cases} -x^2/(1-d), & \text{if } |x| \le (1-d)/2 \\ (4y^2-d)/4d, & \text{if } |y| \le d/2. \end{cases} \end{aligned} \tag{2.7}$$

The maximum potential $V_{\max} = 0$ determines the unperturbed separatrix of the main resonance:

$$p_s(x) = \pm\sqrt{-2KV(x)}, \tag{2.8}$$

while its minimum $V_{\min} = -1/4$ gives the total depth U of the unperturbed potential well:

$$U = K(V_{\max} - V_{\min}) = \frac{K}{4}. \tag{2.9}$$

Perturbation (2.6) is peculiar in that it is of the order of the unperturbed Hamiltonian, irrespective of the perturbation parameter $K \longrightarrow 0$. Nevertheless, the perturbation theory is generally applicable if the other perturbation parameter is large,

$$\lambda = \frac{\Omega}{\omega_0} \gg 1. \tag{2.10}$$

Here, $\omega_0 = \sqrt{2K/d}$ is the frequency of small oscillations at the main resonance (2.5) and $\Omega = 2\pi$ is the frequency of the external perturbation. This adiabaticity parameter governs the separatrix splitting. Using the term adiabaticity in this case emphasizes that the effect of a high-frequency perturbation is qualitatively the same as that of a low-frequency perturbation.

To investigate the motion near the separatrix, let us first determine the change in unperturbed Hamiltonian (2.5) for the period of motion in a close vicinity of the unperturbed separatrix. Following [10], we obtain

$$\begin{aligned} \Delta H_0 &= \int_{-\infty}^{\infty} dt\{H_1, H_0\} \\ &\approx K\int_{-\infty}^{\infty} dt p_s f(x_s)(\delta_1(t) - 1). \end{aligned} \tag{2.11}$$

In the latter expression, the motion along a trajectory close to the separatrix was approximated by the motion along the unperturbed separatrix (hence the infinite integration limits).

Since the force $f(x)$ has two singularities (2.3) at points $y_\pm = \pm d/2$, we integrate (2.11) twice by parts, so that

$$\frac{d^2 f(t)}{dt^2} \approx \frac{d^2 f(y)}{dy^2}p^2 = p^2\Delta f'\delta_1(y - y_\pm), \tag{2.12}$$

where $p = dx/dt$, and only the principal term with the δ function was retained. As a result, we obtain

$$\Delta H_0 \approx Kp_\pm^2\Delta f'[\psi(t_+) - \psi(t_-)] = \frac{K^2}{d}\Delta\psi, \tag{2.13}$$

where $t_\pm$ are the passage times of the singularities at points $y_\pm$, and the function $\psi(t)$ is given by

$$\ddot{\psi} = \frac{d^2\psi}{dt^2} = \delta_1(t) - 1. \tag{2.14}$$

To calculate the difference $\Delta\psi = \psi(t_+) - \psi(t_-)$, we change over to new variables Δ and t_0, where

$$\begin{aligned} \Delta &= \frac{t_+ - t_-}{2} = \sqrt{\frac{d}{2K}}\arcsin\sqrt{d} \\ &= \frac{\lambda}{2\pi}\arcsin\sqrt{d} \end{aligned} \tag{2.15}$$

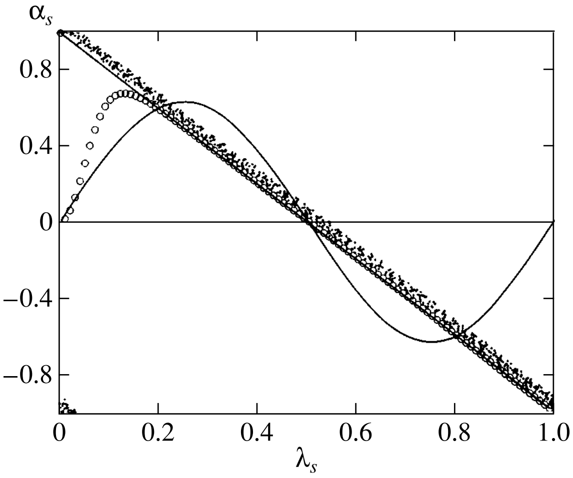


Fig. 2. Periodic dependence of the separatrix splitting angle α on parameters K and d in normalized variables $\alpha_s(\lambda_s)$ (2.20): $d = 0.25, 0.5, 0.75, 0.999$ (dots), $d = 0.01$ (circles) as constructed from our numerical calculations. The solid straight line and the curve represent, respectively, theory (2.19) and the first approximation $\delta_1(t) - 1 \approx 2\cos(2\pi t)$ [see (2.11)]. The argument λ_s is taken modulo 1, so all points (but not circles!) represent many periods of the dependence $\alpha(K, d)$ (see text).

is half the time of motion between the singularities, and

$$t_0 = \frac{t_+ + t_-}{2} \tag{2.16}$$

is the time when the potential minimum is passed ($y = 0$), where the intersection of the branches of the split separatrix is usually analyzed.

In general, $\Delta\psi$ is not factorized in these variables, but this is possible under an additional constraint: $|t_0| \le \Delta$ (> 0). In that case,

$$\Delta\psi = t_0(1 - 2\Delta). \tag{2.17}$$

The angle $\alpha \ll 1$ between the branches of the split separatrix is given by an approximate formula (see [12, 15]),

$$\tan\alpha \approx \alpha \approx \frac{dp}{dy} \approx \frac{dH_0}{p_0^2 dt_0} \approx \frac{2K}{d}(1 - 2\Delta). \tag{2.18}$$

Here, we use the following relations: $dp = dH_0/p$, $dy = pdt$, and

$$\frac{dH_0}{dt} = \Delta\dot{\psi} = 1 - 2\Delta,$$

where all quantities are taken at the point of intersection of the separatrix branches ($y = 0$). The dependence $\alpha(K, d)$ assumes an extremely simple form:

$$\alpha_s \approx (1 - 2\lambda_s) \tag{2.19}$$

in the transformed variables

$$\alpha_s = \alpha\frac{d}{2K} = \alpha\frac{\lambda^2}{4\pi^2},$$
$$\lambda_s = \Delta = \frac{\lambda}{2\pi}\arcsin\sqrt{d} \quad (\text{mod } 1). \tag{2.20}$$

We emphasize that the oscillations of $\alpha(K)$ resulted from the two interfering singularities in the Hamiltonian.

Relation (2.19) is the main result of our analysis. It explains and describes the new phenomenon of the suppression of separatrix chaos and, hence, global diffusion in a certain class of Hamiltonian systems.

A comparison with the results of our numerical experiments is shown in Fig. 2. The very high accuracy of the simple theory (for $K \ll 1$) is limited by a small shift in critical values $K = K_m$. In Ovsyannikov's example (a symmetric piecewise linear mapping, $d = 1/2$), it can be derived even without numerical experiments by using exact expressions for K_m, both predicted in [9] and found later in [12]:

$$K_m = \frac{\pi^2}{16m^2}\left(1 - \frac{\pi^2}{48m^2} + \dots\right) \approx \frac{\pi^2}{16m^2}. \tag{2.21}$$

The last term represents our theory, and a first-order correction is given in parentheses. The theory also explains the unexpected discontinuity of the function $\alpha(\lambda_s)$ at $\lambda_s = 0$ (mod 1) (but not at $\lambda_s = 1/2$), which was found and discussed from a different point of view in [12].

Figure 2 also shows an even simpler approximation with only the first term of the Fourier expansion $\delta_1(t) - 1 \approx 2\cos(2\pi t)$ being retained; it represents the critical values K_m equally well, but does not reproduce the discontinuity in the function of angle α.

The simple relation (2.19) does not give a full picture for the entire family (2.2) either, as demonstrated by the example with a small value of $d = 0.01$ in Fig. 2 (circles). The separatrix splitting is thus seen to be nonsymmetric when $d \longrightarrow 0$ and $d \longrightarrow 1$. On the other hand, it follows from expression (2.2) for the force that the symmetry is preserved when both parameters of the mapping family change: $d \longrightarrow 1 - d$ and $K \longrightarrow -K$. Consequently, changing the sign of K also causes the behavior of the separatrix to change qualitatively. The symmetry is preserved only in the special case of $d = 1/2$, i.e., for the model of a symmetric piecewise linear mapping.

3. THE LIMIT $d \longrightarrow 0$: DISCONTINUITY IN FORCE

Let us first consider the limiting case $d = 0$, $K > 0$, where the force function $f(x)$ experiences a discontinuity (see Fig. 1). The limit differs qualitatively in that the two singularities of the potential at $d > 0$ now merge into one. Consequently, according to our theory, the separatrix splitting angle does not change sign; i.e., the separatrix splits up at any $K > 0$.

Figure 3 shows the results of numerical experiments both in the limit $d = 0$ itself (circles) and in its close vicinity $d = 0.001$ (triangles) and $d = 0.01$ (dots). The dependence $\alpha(K)$ proper is given, because the adiabaticity parameter $\lambda = \pi\sqrt{2d/K} = 0$ loses its meaning at $d = 0$. First, the passage to the limit $d \longrightarrow 0$ in model (2.2) is seen to be continuous with an empirical boundary [by maximum $\alpha(K)$] at

$$K \sim K_B \sim 7d. \tag{3.1}$$

The physical reason why relation (2.18) becomes inapplicable for $K \gtrsim K_B$ is that in deriving it, we ignored the change in velocity between the two singularities through the action of the first of them and the change in transit time Δ between them [see Eqs. (2.13) and (2.15)]. In the previous variables, the transition between the two modes is also shown in Fig. 2 for $d = 0.01$ (circles). We emphasize that there is a deviation from (2.18) only for $K \gtrsim K_B$ and that it is not repeated periodically as dependence (2.18) (see the circle in the upper left corner of Fig. 2). Thus, in the limit $d = 0$, the function of splitting angle actually does not change sign, and, consequently, the separatrix always splits up.

The same method as for $d \neq 0$ (Sect. 2) may be used for a quantitative analysis. The only difference is that the force itself now has a discontinuity, $\Delta f(x) = -2$, and it will therefore suffice to integrate (2.11) by parts only once. We have

$$\begin{aligned}\Delta H_0 &\approx K\int_{-\infty}^{\infty} dt p_s f(x_s)(\delta_1(t) - 1)\\ &\approx K p_0 \Delta f \dot{\psi}(t_0) \approx -\sqrt{2}K^{3/2}\dot{\psi}(t_0).\end{aligned} \tag{3.2}$$

Here, $p_0 \approx \sqrt{K/2}$ as before and

$$\dot{\psi}(t) = \frac{1}{2} - t \quad (\text{mod } 1) \tag{3.3}$$

[see (2.14)]. However, the simple expression (2.18) in the small-angle approximation is now no longer applicable, because the following singularity arises when differentiating with respect to t_0:

$$\frac{dH_0}{dt} = -Kp_0(\delta_1(t) - 1), \tag{3.4}$$

the derivative is taken at two values: $t = t_0 = 0$ and 1/2 when $\Delta H_0 = 0$ [the intersection of the separatrix branches, formula (3.3)]. Each of these values determines the inclination of the corresponding separatrix branch with respect to the x axis (see Fig. 4). The singularity arises at $t = 0$ and corresponds to $\alpha_0 = \pi/2$ ($\tan\alpha_0 = \infty$). The angle of the other branch β is given by [cf. (2.18)]

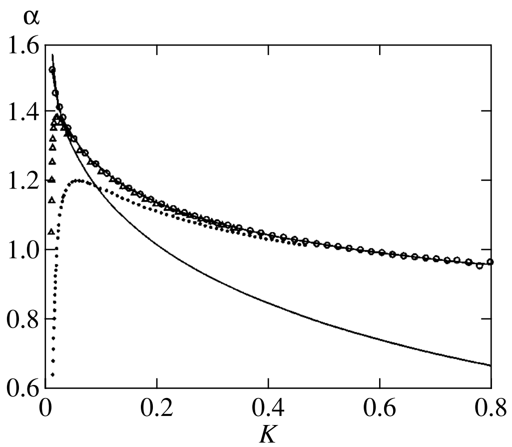


Fig. 3. A plot of separatrix splitting angle α versus parameter K for $d = 0.01$ (dots), 0.001 (triangles), and $d = 0$ (circles) as constructed from our numerical calculations. The lower and upper curves represent the approximate theory (3.6) and the exact theory (3.10), respectively.

$$\tan\beta = \frac{dp}{dx} \approx \frac{1}{2}\frac{dH_0}{p_0^2 dt_0} \approx \frac{K}{p_0} \approx \sqrt{2K}. \tag{3.5}$$

The factor 1/2 of the derivative emerges because ΔH_0 is calculated relative to the unperturbed separatrix (the solid broken line in Fig. 4), while the angle β (as α_0) is taken relative to the x axis. The signs of the angles are determined with the consideration that the branch with α_0 corresponds to moving forward in time, while the other branch corresponds to moving backward in time (see [11]). Finally, we obtain for the angle between the separatrix branches

$$\alpha(K) = \alpha_0 - \beta \approx \frac{\pi}{2} - \sqrt{2K}. \tag{3.6}$$

This simple dependence is indicated in Fig. 3 by the solid line. It agrees well with our numerical experiments for small K; however, the error increases with K, reaching about 40% at $K \approx 0.6$. At larger K, the entire simple pattern of the separatrix splitting for an individual resonance loses its meaning because many resonances overlap (see below).

The error at large K is attributable to the approximate use of the unperturbed separatrix [see (2.8) and (2.7)]

$$p_s(x) = \pm p_0(1 - 2|y|) \tag{3.7}$$

with an amplitude $p_0 \approx \sqrt{K/2}$ when calculating integral (3.2). An interesting feature of the system under consideration is that the unperturbed separatrix (two straight

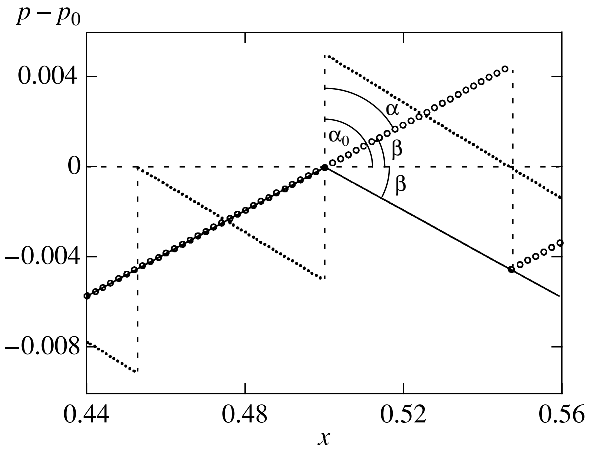


Fig. 4. An example of the separatrix splitting at $K = 0.005$ and $d = 0$: the solid line broken at $x = 0.5$ indicates the unperturbed separatrix (3.7); the separatrix branches are represented by dots (forward in time) and circles (backward in time); the breaks in the branches are connected by the dotted line showing a sequence of points; $p_0 \approx 0.04756$ is the ordinate of the point of intersection (3.8).

lines) retains its shape under the action of a perturbation (Fig. 4). This allows an exact value of $p_0(K)$ to be calculated for any K from the eigenvectors of initial mapping (2.1) at unstable fixed point $x = p = 0$. As a result, we obtain

$$p_0(K) = \frac{K}{\sqrt{2K + K^2} + K}. \qquad (3.8)$$

However, when this expression is substituted into (3.5), the agreement even worsens rather than improving:

$$\tan\beta \approx \frac{K}{p_0(K)} > \sqrt{2K}. \qquad (3.9)$$

The reason lies in another approximation of integral (3.2), retaining the contribution from the jump in force $\Delta f(x)$ alone. Again, because of the peculiar singularity of the unperturbed separatrix at $d = 0$, the angle β can be calculated exactly without any integration directly from the results in Fig. 4: $\tan\beta = 2p_0(K)$. We thus obtain

$$\alpha(K) = \frac{\pi}{2} - \arctan\left(\frac{2}{1 + \sqrt{1 + 2/K}}\right). \qquad (3.10)$$

This expression represents the most accurate result of our theory (the upper solid line in Fig. 3), which is in excellent agreement with the numerical experiment (circles) up to the point at which the resonances begin to overlap. Since mapping (2.1) is periodic not only in x but also in p (and with the same period of 1), there is an infinite set of integer resonances at $p(0) = n$ and $H_0 = n^2/2$, where n is any integer, positive, negative, or zero. The latter special case is considered in this paper. The separatrices of adjacent integer resonances begin to overlap at $p_0 = 1/2$, which destroys their structure. Formally, this occurs only in the limit $K \longrightarrow \infty$ (3.8). Actually, however, this destruction begins much earlier because of the overlapping with intermediate fractional resonances (see [11]). Note also that, formally, the splitting angle is always $\alpha > \pi/4 \approx 0.785$ [see (3.10)]; however, actually and for the same reason, the regular dependence $\alpha(K)$ abruptly cuts off even at $\alpha \approx 0.96 \approx 55°$, $K = K_{cr} \approx 0.8$, and $p_0(K_{cr}) \approx 1/3$ (Fig. 3). For $K > K_{cr}$, the separatrix branches become so unstable that the splitting angle cannot be reliably measured. Interestingly, the deviation of the ordinate of the point at which the separatrix branches intersect $p_0(K)$, according to (3.8), in this range (up to $K = 1.24$) does not exceed 1%. However, this is enough for a strong and irregular distortion of the separatrix branches.

Thus, even at relatively large $K < 0.8$, the separatrix splitting angle is far from reaching zero, let alone changing sign; hence, the separatrix always splits up.

In [13], the limit $d \longrightarrow 0$ was also considered briefly but only for $K < 0$ (in our notation). The inverse limit, $d \longrightarrow 1$, loses its meaning in [13], because the family of mappings is defined there in such a way that in this case, the force $f(x) \longrightarrow 0$ vanishes. In our case, dependence (2.19) is preserved, at least up to $d = 0.999$ (Fig. 2). It should be noted, however, that the pattern of motion qualitatively changes in the limit itself ($d = 1$), because the motion along the unperturbed separatrix represents simple harmonic oscillations [see (2.7) and Fig. 1]

4. CONCLUSION: HOW TYPICAL IS THE CONSERVATION OF THE SEPARATRIX?

We have proposed and tested a simple theory of a new unexpected phenomenon—the conservation of the separatrix of nonlinear resonance in strong chaos on most of the phase plane of a dynamical system [9–13]. The mechanism of this phenomenon is based on a simple idea of the interference (in particular, cancellation [8]) of several singularities in the Hamiltonian of a dynamical system, which govern the separatrix splitting for nonlinear resonance. Numerical experiments and theoretical analysis were carried out for the family of 2D mappings (2.1) in the simplest case of two singularities, which also included the first example of a symmetric piecewise linear mapping [9, 11, 13]. Our study not only confirmed and explained this mechanism, but also allowed a simple theory to be developed to calculate both special values of $K = K_m$ and the dependence of the separatrix splitting angle $\alpha(K)$ over wide ranges of K and d [see (2.19), (3.6), and (3.10)].

We separately considered the passage to the limit $d \longrightarrow 0$, in which the separatrix splits up at any $K > 0$ (Sect. 3). In the opposite case $d \longrightarrow 1$, relation (2.19) remains valid, at least up to $d = 0.999$ (Fig. 2). It should be noted, however, that the pattern of motion qualitatively changes in the limit itself ($d = 1$). First, the motion along the unperturbed separatrix represents simple harmonic oscillations with a finite period, $T_1 = \pi\sqrt{2/K}$ [see (2.7) and Fig. 1]. In addition, the trajectory inside the resonance ($H_0 < 0$) does not reach the singularity of the potential at point $x = 0$ (mod 1) at all. Finally, our preliminary numerical experiments in this limit strongly suggest that the measure of the chaotic component decreases rapidly with decreasing K. It would be of great interest to continue the analysis of this special dynamical system.

We have considered only the integer resonances with $p(0) = n$, where n is any integer. The fractional resonances with $p(0) \approx n/q$ are known (see, e.g., [2, 3, 11]) to have the same structure in appropriate variables. Therefore, one might expect such a mechanism and its simple theory to be also applicable to fractional resonances. If confirmed, we hope in the immediate future, this would allow the conservation of the separatrix for fractional resonances predicted in [13] and revealed by numerical experiments [10–12] to be explained. Although the set of all fractional resonances on the p axis is dense everywhere, the set of all special values of $K = K_{qn}$ at which the separatrix is conserved is not dense [13]. However, its mean density is large enough, and one might expect the strong (although incomplete) suppression of global diffusion at any K. This hypothesis is additionally confirmed by the large number of periodic invariant curves predicted in [13], both ordinary ones with an irrational rotation number ν and new ones with a rational $\nu \neq 0$. We surmise that the emergence of the latter can be interpreted as the suppression of the resonances themselves together with their separatrices. One such strange case for $K = 1/4$ with $\nu = 1/3$ was observed in [11], but its further analysis was postponed to the future.

Of interest and importance is the following question: How typical are the conservation of the separatrix in general and its specific mechanism in particular? It is well known that a large number of examples and even whole families of the so-called completely integrable nonlinear dynamical systems have been “constructed” to date (see, e.g., [16]). There is absolutely no chaos in such systems. However, they are definitely not typical but, in a sense, form a set of measure zero in the space of all possible dynamical systems. From this viewpoint, the new phenomenon of the separatrix conservation in a chaotic system seems more typical, despite the very limited number of examples at present.

The condition for the existence of several singularities in the potential that we used as the basis for our study is neither necessary nor sufficient for the separatrix conservation per se. On the one hand, a preliminary analysis of other examples shows that the presence of several singularities in the force does not yet guarantee the separatrix conservation. For instance, if we simply extend force (2.2) with $d = 0$ by another period, so that two singularities will formally appear, the separatrix will break up as before at any K. In the case under consideration, this takes place because the potential assumes the form of two conjugate wells with an unstable fixed point exactly at the boundary between them. As a result, the unperturbed separatrix always proves to be localized at one of them (depending on initial conditions) with the only singularity.

On the other hand, for an analytic potential, the singularities that determine the separatrix splitting can be located not on the real time axis but in the complex plane. Such a situation appears to have been actually observed in a completely different problem of charged-particle confinement in Cohen’s long magnetic trap [17] (see also [18]).

Finally, the passage of the separatrix splitting angle through zero depending on the system parameter and, hence, the separatrix conservation at certain values of this parameter are also possible in principle for a special form of the potential with no singularities whatsoever. This all undoubtedly deserves a further study.

In conclusion, note that even though the new effect of the separatrix conservation in chaos is not universal (a favorite term in current studies of dynamic chaos), nevertheless, we hope that the criterion for the interference of singularities and the theory developed on its basis (which can be easily generalized to an arbitrary number of singularities) can significantly help in studies of a wide class of Hamiltonian dynamical systems.

ACKNOWLEDGMENTS

This study was supported in part by the Russian Foundation for Basic Research (project no. 01-02-16836).

REFERENCES

1. V. I. Arnold and A. Avez, *Ergodic Problems of Classical Mechanics* (Benjamin, New York, 1968; RCD, Izhevsk, 1999).
2. B. V. Chirikov, Phys. Rep. **52**, 263 (1979).
3. A. Lichtenberg and M. Lieberman, *Regular and Chaotic Dynamics* (Springer-Verlag, New York, 1992; Mir, Moscow, 1984).
4. G. M. Zaslavsky and R. Z. Sagdeev, *An Introduction to the Nonlinear Physics* (Nauka, Moscow, 1988).
5. B. V. Chirikov, Chaos, Solitons and Fractals **1**, 79 (1991).
6. J. Moser, *Stable and Random Motions in Dynamical Systems* (Princeton Univ. Press, Princeton, 1973).
7. B. V. Chirikov, E. Keil, and A. Sessler, J. Stat. Phys. **3**, 307 (1971).
8. M. Hénon and J. Wisdom, Physica D (Amsterdam) **8**, 157 (1983).
9. L. V. Ovsyannikov, private communication (1999).
10. V. V. Vecheslavov, Zh. Éksp. Teor. Fiz. **119**, 853 (2001) [JETP **92**, 744 (2001)].

11. V. V. Vecheslavov, Preprint No. 2000-27, IYaF RAN (Budker Institute of Nuclear Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 2000); nlin.CD/0005048.
12. V. V. Vecheslavov, Preprint No. 99-69, IYaF RAN (Budker Institute of Nuclear Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 1999).
13. S. Bullett, Commun. Math. Phys. **107**, 241 (1986).
14. M. Wojtkowski, Commun. Math. Phys. **80**, 453 (1981); Ergod. Theory Dyn. Syst. **2**, 525 (1982).
15. V. V. Vecheslavov and B. V. Chirikov, Zh. Éksp. Teor. Fiz. **114**, 1516 (1998) [JETP **87**, 823 (1998)].
16. B. G. Konopelchenko, *Nonlinear Integrable Equations* (Springer-Verlag, Berlin, 1987).
17. R. Cohen, in *Intrinsic Stochasticity in Plasmas*, Ed. by G. Laval and D. Gresillon (Editions de Physique, Orsay, 1979).
18. B. V. Chirikov, in *Reviews of Plasma Physics*, Ed. by B. B. Kadomtsev (Énergoatomizdat, Moscow, 1984; Consultants Bureau, New York, 1987), Vol. 13.

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 93, No. 3, 2001, pp. 657–670.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 3, 2001, pp. 749–764.
Original Russian Text Copyright © 2001 by Zabolotskiĭ.

NONLINEAR PHYSICS

Soliton Dynamics in a Two-Level Medium with Pumping

A. A. Zabolotskiĭ
Institute of Automatics and Electrometry, Siberian Division, Russian Academy of Sciences,
Universitetskiĭ pr. 1, Novosibirsk, 630090 Russia
e-mail: zabolotskii@iae.nsk.su

Received March 5, 2001

Abstract—A perturbation theory is developed for a system of evolution equations close to integrable systems in the method of inverse scattering with a spectral parameter depending on variables. This theory is used for analyzing the evolution features for soliton light pulses in a two-level medium with the upper level pumping taking into account linear and nonlinear losses as well as dispersion. Various modes of soliton evolution (including the random mode) in such a system are investigated numerically. It is shown that anomalies in the dependence of soliton parameters on the length of the laser medium appear in the presence of nonuniform broadening in the case when the upper level pumping rate and inverse population losses are functions of detuning. The contribution from the radiative component of the solution is also analyzed taking into account perturbations; it is shown that this contribution can be disregarded in a certain range of parameters. *© 2001 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Dynamics of solitons in nonlinear optics has been studied by using several approaches, including completely integrable models (see, for example, [1] and review [2]). The application of such models imposes, as a rule, a number of physical limitations; however, these models provide the most comprehensive description of the evolution of ultrashort pulses in nonlinear media. If the evolution equations differ from integrable equations in small terms, these terms can be taken into account in the perturbation theory which has been developed for quasi-integrable systems [3, 4] (see also review [5]). This perturbation theory was constructed for nearly integrable systems on the basis of the method of inverse scattering problem (MISP) [1] with a spectral parameter independent of variables. However, each integrable equation with constant coefficients to which the algorithm of the inverse problem is applicable corresponds to a whole class of equations with variable coefficients. These equations are integrable in the MISP with a variable spectral parameter and are referred to as "deformations" of the initial equations, or deformed integrable models [6]. A detailed description of such models and the method of their exact solution is given by Burtsev *et al.* [6]. This class of deformed integrable models includes a number of important physical models some of which can be used in nonlinear optics [7, 8].

In this paper, the apparatus of perturbation theory developed in [3, 4] for systems of equations close to integrable equations in the MISP with a spectral parameter independent of variables is generalized for systems close to deformed integrable models. This generalization makes it possible to study analytically the strongly nonlinear stage of field evolution for new ranges of physical parameters.

The system of Maxwell–Bloch evolution equations with external pumping is used for simulating processes in gaseous media, solids, and dyes [9, 10]. In such lasing media, three- or four-level pumping modes are often used. Under certain conditions, these modes can be reduced to the system of Maxwell–Bloch equations for a two-level system with an additional term in the dynamic equation for the difference in the populations of levels, which describes the pumping rate (see Eq. (9.106) in [9]).

The MISP apparatus was used earlier [12] for studying the characteristics of a two-level laser amplifier [11]. For the initial and boundary conditions corresponding to the amplification of a small initial pulse or fluctuations of the medium in an extended laser amplifier, it was proved that the (quasi)radiative asymptotic form of the field is completely determined by the continuous real spectrum in the Zakharov–Shabat problem [12]. There is no external pumping in the laser amplifier, and the system is initially assumed to be in the inverted state. Medium fluctuations or a small initial pulse of the field are amplified due to the conversion of energy stored in the medium at the initial instant into the energy of the light field. A formal analogy with the Chu–Scott model for Raman scattering [13] is worth noting. On the basis of this model, it is shown [14] that the initial conditions corresponding to a partial initial population inversion in an extended nonlinear medium also lead to a quasi-radiative (nonsoliton) asymptotic form.

At the same time, the model of a two-level lasing transition with external pumping [9] is used much more frequently for describing laser generation in nonlinear optics. In this connection, an analysis of qualitatively new features of nonlinear modes of pulse evolution in such models is of practical importance.

The integrability of the model of a two-level laser with a constant pumping of levels in the MISP with a variable spectral parameter was proved in [6, 8]. Under the action of pumping, the upper level of the medium is populated and lasing starts when population inversion is created. For small initial pulses, laser generation can be described asymptotically by a (quasi-)radiative solution associated with the real continuous spectrum of the Zakharov–Shabat problem as in the case of an extended laser amplifier [12]. The results of analytic [15–17] and numerical [18] analyses proved that this solution consists of nonlinear oscillations with an amplitude increasing monotonically with coordinate z. However, in contrast to a laser amplifier, the field amplitude averaged over oscillations does not tend to zero, but increases in proportion to $\sqrt{z}$.

A two-level lasing medium with pumping can also be used for amplifying soliton pulses with a carrier frequency close to the frequency of the transition. The dynamics of a soliton pulse which is associated with the isolated eigenvalue of the Zakharov–Shabat problem with a positive imaginary component differs significantly from the behavior of the radiative solution. Partial soliton solutions of the Maxwell–Bloch integrable system of equations with pumping were obtained in [6–8, 15–19] for a pumping rate independent of z and of frequency detuning.

In the present work, the developed apparatus of perturbation theory is applied for studying the soliton dynamics in a model close to the deformed integrable system of Maxwell–Bloch equations with pumping. A distinguishing feature of this work as compared to [6–8, 15–19] is that nonuniform broadening and the dependence of pumping rate and inverse population losses on the coordinate and frequency detuning are additionally taken into account. The inclusion of this dependence is essential in the presence of nonuniform broadening of the laser transition which can be associated with the motion of atoms, the interaction of atom with the ambient, and nonuniformity of electric or magnetic fields. This formally simple generalization leads to a number of new physical properties of the amplifying medium. The additional inclusion of quadratic dispersion, linear and nonlinear losses, and driving force using the perturbation theory apparatus developed by us here leads to a number of new features in the dependence of soliton parameters on the laser medium length.

In the perturbation theory describing the nonlinear dynamics of isolated pulses, robust soliton solutions of the unperturbed integrable system of equations with parameters independent of variables are used as nonlinear modes [5]. In the present work, we choose exact “deformed” soliton solutions with variable parameters for such nonlinear modes. The inclusion of perturbations in the framework of the developed perturbation theory leads to systems of equations including the contribution of these deformations and perturbations. Numerical and theoretical analysis of the obtained equation makes it possible to reveal some peculiarities in soliton enhancement in such a system, which may be accompanied by an abrupt switching of amplification modes, bifurcation, and randomization.

The soliton dynamics is affected by the radiative solution whose contribution cannot be disregarded in the amplifying medium. The soliton amplitudes determined from the solution of the equations in perturbation theory are compared with that obtained from the numerical solution of quasi-self-similar solution of the initial system of evolution equations taking into account weak perturbations. The effects of perturbations on soliton dynamics and on the radiative solution are qualitatively different. This allows us to discover a range of parameters in which the amplitude of nonlinear oscillations in the radiative solution is much smaller than the amplitude of a soliton and, hence, the contribution of the radiative solution can be neglected.

The paper has the following structure. In Section 2, the equations of the perturbation theory for a model close to integrable are presented using the MISP apparatus with the help of the Zakharov–Shabat spectral problem with a variable spectral parameter. The model of a two-level laser with a frequency-dependent pumping and its derivation are described in Section 3. In Section 4, the regimes of soliton propagation are determined for various conditions of pumping and perturbations. The role of pumping (including periodically modulated one) in the emergence of random dynamics is analyzed numerically in Section 5. The last section is devoted to an analysis of the contribution from the radiative component of the solution to the general solution of the model for various types of perturbations.

2. PERTURBATION THEORY FOR SYSTEMS CLOSE TO DEFORMED INTEGRABLE MODELS

The perturbation theory for nonlinear systems of equations close to integrable was constructed by Kaup [3] and Karpman and Maslov [4] on the basis of the Zakharov–Shabat spectral problem. We will generalize the results obtained by these authors to the case of deformed integrable models.

We assume that an exactly integrable system of equations can be presented in the form of the compatibility condition for the following linear systems of equations:

$$\Phi_z = U\Phi, \quad \Phi_\tau = V\Phi. \tag{1}$$

Here, U, V, and Φ are matrix functions of τ, z, and the spectral parameter λ. In the general case, U and V, which are rational functions of λ, have the form

$$U(z, \tau; \lambda) = u_0 + \sum_{n=1}^{N_1} \frac{u_n(\tau, z)}{\lambda - \lambda_n}, \qquad V(z, \tau; \lambda) = v_0 + \sum_{n=1}^{N_2} \frac{v_n(\tau, z)}{\lambda - \mu_n}, \tag{2}$$

where the simple poles λ_n and μ_m do not coincide. The compatibility condition has the form

$$U_z - V_\tau + [U, V] = 0. \tag{3}$$

The dependence of the spectral parameter on the variables in the general case is determined by the following equations:

$$D_z\lambda \equiv [\partial_z - F(\lambda)\partial_\lambda]\lambda = 0, \qquad D_\tau\lambda \equiv [\partial_\tau - G(\lambda)\partial_\lambda]\lambda = 0. \tag{4}$$

Taking these expressions into account, we can formally write the compatibility condition (3) so that it does not contain an explicit dependence of λ on z and τ:

$$D_zU - D_\tau V + [U, V] + [D_z, D_\tau] = 0. \tag{5}$$

Exact integrability requires that the last commutator on the left-hand side of Eq. (5) be equal to zero, i.e.,

$$F_\tau - GF_\lambda = G_z - FG_\lambda. \tag{6}$$

Let the functions F and G have the form

$$F = -\sum_{m=1}^{N_1} \frac{c_m}{\lambda - \lambda_m}, \quad G = -\sum_{m=1}^{N_2} \frac{b_m}{\lambda - \mu_m}. \tag{7}$$

In this case, the evolution equations satisfying conditions (5) and (6) can be reduced, using a gauge transformation, to the form [6]

$$\begin{gathered} \partial_z u_0 - \partial_\tau v_0 + [u_0, v_0] = 0, \\ \frac{\partial u_n}{\partial z} + \left[u_n, \sum_{k=1}^{N_2} \frac{v_k}{\lambda_n - \mu_k}\right] = \sum_{m=1}^{N_2} \frac{b_m u_n + c_n v_m}{(\lambda_n - \mu_m)^2}, \\ \frac{\partial v_n}{\partial \tau} + \left[v_n, \sum_{k=1}^{N_1} \frac{u_k}{\mu_n - \lambda_k}\right] = \sum_{m=1}^{N_1} \frac{c_m v_n + b_n u_m}{(\lambda_m - \mu_n)^2}. \end{gathered} \tag{8}$$

In the first approximation in the small parameter ϵ of the perturbation theory which will be proposed below, it is sufficient to require that

$$[D_z, D_\tau] \sim O(\epsilon^2). \tag{9}$$

This condition is satisfied, for example, if F and G are small and are slowly varying functions of one of the variables:

$$F = \epsilon\mathcal{F}(\epsilon\tau, z; \lambda), \quad G = \epsilon\mathcal{G}(\tau, \epsilon z; \lambda) \tag{10}$$

or

$$G\frac{\partial}{\partial\lambda}F \lesssim O(\epsilon^2), \quad \frac{\partial}{\partial\tau}F \lesssim O(\epsilon^2), \qquad \frac{\partial}{\partial z}G \lesssim O(\epsilon^2), \quad \frac{\partial}{\partial\lambda}G \lesssim O(\epsilon^2). \tag{11}$$

For the system of equations (8), conditions (10) indicate that

$$c_n = \epsilon C_n(\epsilon\tau, z), \quad b_n = \epsilon\mathcal{B}_n(\tau, \epsilon z), \qquad \lambda_m = \lambda_m(\epsilon\tau, z), \quad \mu_m = \mu_m(\tau, \epsilon z), \tag{12}$$

i.e., the right-hand sides of the second and third equations in system (8) are of the order of $O(\epsilon)$. Conditions (11) are satisfied for

$$b_n \lesssim O(\epsilon^2), \quad \frac{\partial}{\partial z}b_n \lesssim O(\epsilon^2), \tag{13}$$

where the functions c_n, λ_m, and μ_m as well as their derivatives can be $\gtrsim O(1)$.

Let us suppose that the evolution equation has the form

$$u_z = \mathcal{N}[u] + \epsilon P[u]. \tag{14}$$

We assume that this equation is exactly integrable in the MISP in the absence of a small perturbation $\epsilon P[u]$, $\epsilon \ll 1$. Taking into account this perturbation, the compatibility condition (5) assumes the form

$$D_\tau U - D_z V + [U, V] = [D_\tau, D_z] + \epsilon\hat{P}[u], \tag{15}$$

where $\hat{P}$ is a certain operator whose form depends on the type of equation and perturbation.

In the case of the Zakharov–Shabat spectral problem [1], the equations in perturbation theory for a system of equations close to the deformed integrable model can be obtained by generalizing the known equations from [4]. In this case, $u(\tau, 0)$ is the potential (see Eq. (14)).

It should be noted that, in the present work, the space (z) and time (τ) variables are transposed as compared to similar equations in [4]. This is dictated by the specific physical examples considered by us here.

We assume that the evolution of scattering data $\mathcal{S}$ is described by the following equation:

$$\frac{\partial}{\partial z}\mathcal{S} = \Omega(\lambda)\mathcal{S} + \mathcal{F}(\tau, z; \lambda). \tag{16}$$

Here, $\Omega(\lambda)$ is an analogue of the dispersion relation. In contrast to the similar expression presented in [4], Eq. (16) is supplemented with the term $\mathcal{F}$. For the spectral data corresponding to the continuous and discrete spectra of the problem $\Phi_z = U\Phi$, we have $\mathcal{F} \equiv F$. For other spectral data, $\mathcal{F} \equiv 0$. Repeating the derivation carried out in [4] and taking into account this term, we arrive at the following system of equations for the evolution of scattering data in the first order in the small parameter ϵ characterizing the perturbation:

$$\begin{aligned}\frac{\partial a(\lambda, z)}{\partial z} &= \epsilon \int_{-\infty}^{+\infty} d\tau P[u]\psi^{(1)}(\tau, \lambda)\varphi^{(2)}(\tau, \lambda)\\ &+ \epsilon \int_{-\infty}^{+\infty} d\tau P^*[u]\psi^{(2)}(\tau, \lambda)\varphi^{(1)}(\tau, \lambda),\end{aligned} \tag{17}$$

$$\begin{aligned}\frac{\partial b(\lambda, z)}{\partial z} &= 2i\Omega(\lambda) + \epsilon \int_{-\infty}^{+\infty} d\tau P[u]\psi^{(2)*}(\tau, \lambda)\varphi^{(1)}(\tau, \lambda)\\ &- \epsilon^* \int_{-\infty}^{+\infty} d\tau P^*[u]\psi^{(1)*}(\tau, \lambda)\varphi^{(2)}(\tau, \lambda),\end{aligned} \tag{18}$$

$$\begin{aligned}\frac{\partial b_n(\lambda_n, z)}{\partial z} &= 2i\Omega(\lambda_n) + \frac{b_n}{a'(\lambda_n)} \int_{-\infty}^{+\infty} d\tau\\ &\times \{\epsilon P[u]Q^{(1)}(\tau, \lambda_n) + \epsilon^* P^*[u]Q^{(2)}(\tau, \lambda_n)\},\\ Q^{(m)}(\tau, \lambda_n) &= \frac{\partial}{\partial\lambda}[\psi^{(m)}(\tau, \lambda)\varphi^{(m)}(\tau, \lambda)\\ &- \psi^{(m)}(\tau, \lambda)\varphi^{(m)}(\tau, \lambda)]\Big|_{\lambda = \lambda_n}, \quad m = 1, 2,\end{aligned} \tag{19}$$

$$\begin{aligned}\frac{\partial \lambda_n(z, \tau)}{\partial z} &= F(z, \tau; \lambda_n) - \frac{1}{a'(\lambda_n)}\\ &\times \int_{-\infty}^{+\infty} d\tau \{\epsilon P[u]\psi^{(1)}(\tau, \lambda_n)\varphi^{(2)}(\tau, \lambda_n)\\ &+ e^* P^*[u]\psi^{(2)}(\tau, \lambda_n)\varphi^{(1)}(\tau, \lambda_n)\}\\ &= F(Z, \tau; \lambda_n) - Q(z, \tau; \lambda_n).\end{aligned} \tag{20}$$

The asterisk indicates complex conjugation. It should be noted that Eqs. (18) and (19) were derived in [3, 4] for trivial boundary conditions ($z = 0, \infty$) corresponding to the evolution of pulses against the background of the stable state.

In the case when λ is a function of τ, the system of equations (17)–(20) should be supplemented with the equation

$$\frac{\partial \lambda_n(z, \tau)}{\partial \tau} = G(z, \tau; \lambda_n). \tag{21}$$

This equation is a consequence of the "deformation" of the unperturbed model and does not change in the first order of perturbation theory.

While deriving the system of equations (17)–(20), we used standard assumptions concerning the applicability of perturbation theory [4]. As compared to the known equations in [4], the right-hand side of Eq. (20) is supplemented with the term $F(\lambda_n)$, and Eq. (21) is also added. It should be recalled that Eq. (20) without the second term on the right-hand side and Eq. (21) describe the dependence of the spectral parameter on variables without resorting to perturbation theory.

In the general case, system (17)–(19) must be supplemented with the equations

$$\partial_z \lambda = F(z, \tau; \lambda), \quad \partial_\tau \lambda = G(z, \tau; \lambda), \tag{22}$$

where λ belongs to the continuous real spectrum $\Phi_z = U\Phi$ of the problem. In the present work, we study the dynamics of the discrete soliton spectrum without taking into account the contribution of the continuous spectrum; i.e., the last pair of equations is not taken into account. Below (in Section 5), we will analyze the contribution from the continuous real spectrum and determine the conditions under which it can be neglected.

In the case when λ_n are functions of variable τ by virtue of Eq. (21), this dependence must be taken into account while deriving the functions $\psi^{(k)}(\tau; \lambda_n)$ and $\varphi^{(k)}(\tau; \lambda_n)$ appearing on the right-hand sides of Eqs. (17)–(20). This constitutes one of the main distinctions of the perturbation theory for systems close to deformed integrable models from the known theory [4]. Since λ_n in the general case satisfy two equations (20) and (21), we must find the compatibility condition for these equations. In the case of perturbations, the dependence of λ_n on z associated with this perturbation must also be taken into consideration. Taking into account Eqs. (20) and (21), we can write the approximate commutativity condition for the derivatives of λ in the first order of perturbation theory in the form

$$\begin{aligned}&\frac{\partial}{\partial\tau}F - G\frac{\partial}{\partial\lambda}(F - Q) - \frac{\partial}{\partial z}G\\ &+ (F - Q)\frac{\partial}{\partial\lambda}G \sim O(\epsilon^2).\end{aligned} \tag{23}$$

For $Q \sim O(\epsilon)$, condition (23) holds, in particular, if conditions (11) are valid. It should be noted that, when condition (23) is satisfied and the commutator is small, $[D_z, D_\tau] \sim O(\epsilon)$, the latter can be included in the perturbation:

$$\epsilon\hat{P}[u] \longrightarrow \epsilon\hat{P}[u] + [D_\tau, D_z].$$

In this case, the fulfillment of condition (9) is not required.

Relations following from the condition $[D_z, D_\tau] = 0$ were obtained in [6] for a number of exactly integrable models. The limitations imposed on the physical parameters of models cannot be realized in most cases. In the perturbed problem, condition (23) includes the contribution of perturbation Q. In the first approximation in ϵ, the difference between F and Q is only that Q appears with a small parameter. This fact enables us to use in the perturbation theory the results obtained in the exact theory developed for systems with a variable spectral parameter and vice versa. For example, it follows from (23) that in some cases this condition might be satisfied approximately for a special choice of the form of perturbation and the conditions imposed on the functions $F(z, \tau; \lambda)$, $G(z, \tau; \lambda)$, and $Q(z; \lambda)$.

3. MAXWELL–BLOCH EQUATIONS FOR A TWO-LEVEL SYSTEM WITH PUMPING

We apply the perturbation theory constructed by us for describing the dynamics of the soliton solution of the system of Maxwell–Bloch equations for a two-level medium with the upper level pumping. Some idealized three- and four-level models of lasing in gas, ionic, and other lasers can be reduced to this system of equations [9].

We assume that the diagram of laser pumping has the form of a four-level medium with energy levels 0, 1, 2, 3 (Fig. 1) and corresponding energies W_i such that

$$W_3 > W_1 > W_2 > W_0.$$

There exists a certain mechanism of pumping according to which electrons are transferred from the ground level 0 to level 3 with energy W_3, thus populating it.

As a result of a fast transition of electrons from level 3 to the lower lying level 1, the difference in the populations of level 1 and level 2, having a lower energy, increases; we assume that level 2 lies above the ground level 0 and is empty at the initial instant. It is assumed that the energy difference $W_1 - W_0$ is much larger than $k_B T_k$ (k_B is the Boltzmann constant and T_k is the temperature). This condition enables us to disregard the thermal mechanism of populating level 2. The population inversion is maintained in the case of rapid transition of electrons from level 2 to the ground level 0. In such a system, we can obtain population inversion of levels 1 and 2 and generation of a field with amplitude E_{12} in this transition at frequency $\omega_{12} = (W_1 - W_2)/\hbar$.

In order to describe the generation of ultrashort pulses, many authors employ the system of Maxwell–Bloch equations for a two-level medium, e.g., for describing self-induced transparency [9]. The contribution of pumping is taken into account by adding a term proportional to the difference in the pumping rates for levels 1 and 2 to the equation describing the dynamics of the difference in the populations of levels 1 and 2. The pumping of the latter level can be executed by electron transfer between levels 3 and 2 (see Fig. 1). The parity prohibition for such transitions can be removed due to the interaction with surroundings in the solid matrix or as a result of a cascade transition through an intermediate level. This transition reduces the population inversion like the electron transfer from level 0 to level 2 and from level 1 to level i characterized by an energy $W_i < W_1$. If the relaxation times at these transitions are much shorter than the characteristic duration of pulses, the contribution of these transitions can also be taken into account in the model described below. In a three-level system, the lower laser level 2 coincides with the ground level [9]. In the idealized model which will be solved below, these lasing diagrams differ only in the initial conditions for the difference in the populations of the levels involved in the 1 ⟷ 2 lasing transition.

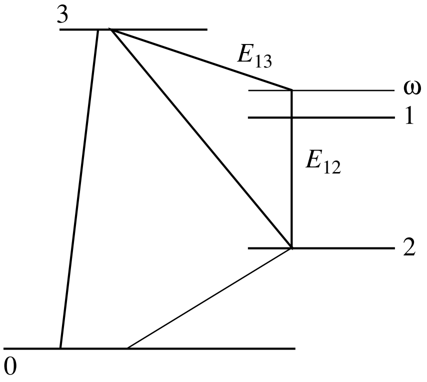


Fig. 1. Diagram of a four-level laser: E_{12} is the field amplitude on the 1 ⟷ 2 laser transition, and ω is the detuning. Horizontal lines are energy levels. Inclined lines show possible transitions between the levels.

In the slow envelope approximation, the system of Maxwell–Bloch equations can be presented in the form

$$\begin{aligned} &\partial_z E(z, \tau) \\ &= \int_{-\infty}^{\infty} g(z; \omega) R_+(z, \tau; \omega) d\omega + \epsilon P(z, \tau), \\ &\partial_\tau R_+ = 2R_3 E(z, \tau) - \gamma_2 R_+ + 2i\omega R_+, \\ &\partial_\tau R_3 = -[R_+^* E(z, \tau) + R_+ E^*(z, \tau)] \\ &\qquad + \tilde{h}(z; \omega) - \gamma_1 R_3. \end{aligned} \tag{24}$$

Here, $E = (2d/\hbar)E_{12}$, where E_{12} is the electric field amplitude at the lasing transition;

$$z = \frac{1}{2}\kappa_0 \int_0^x N_0(s) ds, \quad \tau = t - \frac{x}{c}, \quad \kappa_0 = \frac{2\pi\omega_{12}\mu_{12}^2}{\hbar c},$$

$$R_3(z, \tau; \omega) = \frac{N_3(z, \tau; \omega)}{N_0(z)},$$

$$R_+(z, \tau; \omega) = \frac{R(z, \tau; \omega)}{N_0(z)}, \quad R_- = R_+^*;$$

N_0, N_3, and R are the density of active atoms, the difference in the populations of the levels, and the polarizability of the transition in the two-level medium, respectively; ω_{12} is the transition frequency; μ_{12} is the dipole moment, x and t are the space and time variables; ϵP is perturbation with a small parameter ϵ; $\gamma_{1,2} = 1/T_{1,2}$ are the relaxation constants; $g(z; \omega)$ is the function describing nonuniform broadening, i.e., the distribution of atoms over the frequency shift ω relative to the transition frequency; $\tilde{h}(z;\omega)$ is the pumping rate for $\tilde{h} > 0$ or the inverse population loss rate if $\tilde{h} < 0$.

We assume that $E(0, \tau) = 0$ for all $\tau \longrightarrow \pm\infty$ and $R_+(z, 0) = 0$, $R_3(z, 0) = -n_0 < 0$ for all z; i.e., the field is absent at infinities, and the medium is in the ground state at the instant $\tau = 0$: $n_0 = 1$ for a three-level lasing system and $n_0 = 0$ for a four-level system.

In the absence of perturbation ($\epsilon P = 0$) and for characteristic times much shorter than $T_{1,2}$ ($\gamma_{1,2} = 0$), the system of equations (24) can be presented in the form of compatibility condition (3) for the following linear systems [8]:

$$\partial_\tau \Phi = \begin{pmatrix} -i\lambda & E \\ -E^* & i\lambda \end{pmatrix} \Phi = U\Phi, \tag{25}$$

$$\partial_z \Phi = \frac{i}{2}\int_{-\infty}^{\infty} \frac{g(\omega)}{\lambda - \omega} \times \begin{pmatrix} R_3 & -R_+ \\ -R_- & -R_3 \end{pmatrix}(z, \tau, \omega)\, d\omega \Phi = V\Phi. \tag{26}$$

Here, λ is the spectral parameter and $\Phi(z, t\,\lambda)$ is a two-component function. The dependence of the spectral parameter on variable z is determined by an additional condition. This dependence, which was obtained in [6, 16] in the form

$$\partial_z \lambda = \mathrm{const}/\lambda,$$

corresponds to pumping at a frequency exactly coinciding with the transition frequency (disregarding the nonuniform broadening of the transition) for $\tilde{h}(z;\omega) \equiv \mathrm{const}$. The MISP apparatus was developed in [8] in the Maxwell–Bloch model for a medium with pumping taking into account nonuniform broadening. In the present work, we propose the following new generalization of this dependence:

$$\frac{\partial}{\partial z}\lambda = \int_{-\infty}^{\infty} \frac{h(z,\nu)g(z,\nu)d\nu}{\nu - \lambda} \equiv F(z;\lambda), \tag{27}$$

where $g(z;\nu)$ and $h(z;\nu) = \tilde{h}(z;\nu)/2$ are arbitrary real-valued functions (in the general case). The dependence $h(z;\nu)$ on the frequency detuning ν may be associated with the physical origin of pumping and the structure of the atomic system and emerges, for example, due to different relaxation rates of atomic transitions, the spectral dependence of electron transport mechanisms, the frequency dependence of the absorption cross section, the effect of the motion of atoms and atomic collisions, and the nonuniformity of the pumping field.

We will demonstrate the frequency dependence of pumping using the following simple qualitative example. Let us consider the four-level lasing system depicted in Fig. 1. We assume that the dipole transitions $3 \longleftrightarrow 1$ and $1 \longleftrightarrow 2$ between the energy levels are possible. The dynamics of population ρ_{11} of level 1 and of the polarization ρ_{13} of the $1 \longleftrightarrow 3$ transition are described by the following equations from the Bloch system:

$$\partial_\tau \rho_{11} = \frac{2i\mu_{12}}{\hbar}\rho_{12}E_{12}^* + \frac{2i\mu_{12}}{\hbar}\rho_{13}E_{13}^* + \mathrm{c.c.}, \tag{28}$$

$$\partial_\tau \rho_{13} + (i\omega + \gamma_{13})\rho_{13} = \frac{2i\mu_{13}}{\hbar}(\rho_{11} - \rho_{33})E_{13}. \tag{29}$$

Here, ρ_{ij} are the elements of the density matrix of a multilevel system, μ_{ij} are the dipole moments of the transitions, ω is the frequency detuning, $E_{12}(z, \tau)$ is the field amplitude on the lasing transition, and $E_{13}(z)$ is the constant preset field which transfers electrons from level 3 to level 1. Level 3 is permanently populated by electrons from the ground level 0. In an ion laser, for example, this process may be induced by the strong current flowing through the gas. We assume that the relaxation time γ_{13}^{-1} at the $1 \longleftrightarrow 3$ transition is much shorter than the characteristic times of the pulses being generated and $\rho_{33}(z, 0) \gg \rho_{11}(z, \tau)$ $\forall z, \tau$. In this case, we obtain from the system of equations (28), (29)

$$\begin{aligned} \partial_\tau \rho_{11} &= \frac{2i\mu_{12}}{\hbar}\rho_{12}E_{12}^* - \frac{2i\mu_{12}^*}{\hbar}\rho_{12}^*E_{12} \\ &+ |\mu_{13}E_{13}(z)|^2 \frac{8\rho_{33}(z,0)\gamma_{13}}{\hbar^2(\gamma_{13}^2 + \omega^2)}. \end{aligned} \tag{30}$$

The last term on the right-hand side of this equation is the pumping rate $\tilde{h}(z, \omega)$. A similar expression (but with opposite sign) corresponding to the population inversion loss rate appears if the transition $3 \longleftrightarrow 2$ and (or) the transition from level 1 to an intermediate level i with energy W_i such that $W_2 < W_i < W_1$ and with short relaxation times can occur.

In the presence of nonuniform broadening, the dependence of pumping rate on detuning ω may lead to peculiarities and anomalies in the dependence of the parameters of solitons propagating in a two-level medium with pumping. The remaining part of this article is devoted to an analysis of such features using a model close to the deformed integrable system of Maxwell–Bloch equations.

4. SOLITON DYNAMICS IN A TWO-LEVEL MEDIUM WITH PUMPING

In the example considered below, the function $G(z, \tau; \lambda) \equiv 0$ for $\gamma_{1,2} = 0$. In this case, F is an arbitrary function of z and λ and $Q = Q(z, \lambda) \sim O(\epsilon)$.

Let us assume that the frequency distribution describing nonuniform broadening is of Lorentzian shape and is normalized to unity:

$$g(z;\omega) = \frac{\Gamma_0}{\pi}\frac{1}{\omega^2+\Gamma_0^2}, \tag{31}$$

and $h(z;\omega) \equiv h_1(z)$. The solution to Eq. (27) corresponding to an isolated value of the spectral parameter $\lambda = i\eta + \xi$ has the form

$$\eta(z) = \left\{\frac{y_0-\xi_0+H}{2} + \frac{[(y_0-\xi_0+H)^2+4\xi_0^2 y_0]^{1/2}}{2}\right\}^{1/2} - \Gamma_0, \tag{32}$$

$$\xi(z) = \xi_0\frac{\eta(0)+\Gamma_0}{\eta(z)+\Gamma_0}.$$

Here,

$$y_0 = [\eta(0)+\Gamma_0]^2, \quad H(z) = 2\int_0^z h_1(z)dz > 0,$$

$$\xi_0 = \xi(0).$$

The signs are chosen taking into account the fact that η is real and positive.

For $\xi(0) = 0$ and $h_1(z) \equiv h = \text{const}$, solution (32) can be simplified and the soliton solution of the unperturbed model (24) ($P = \gamma_{1,2} = 0$) has the form

$$E_s(z,\tau) = \frac{2\eta(z)}{\cosh[2(\eta(z)-\Gamma_0)(\tau-1/h)+2\eta_0/h]}. \tag{33}$$

This solution describes the nonmonotonic increase in the soliton amplitude with increasing value of $hz > 0$, which tends to infinity for $z \longrightarrow \infty$ In this case, the soliton duration tends to zero and the group velocity of the soliton tends to the velocity of light.

When small perturbations are taken into account, the amplification dynamics may change qualitatively. For example, a steady-state or attenuating asymptotic form may appear for $\eta(z)$, and soliton parameters become qualitatively dependent on the spectral characteristics of pumping and losses. These singularities can be described in the perturbation theory developed by us here.

Let a soliton associated with the spectral parameter $\lambda_0 = \xi(0) + i\eta(0)$ and having the form

$$E_0(\tau) = \frac{2\eta(0)\exp[-2i\xi(0)\tau - i\phi_0]}{\cosh[2\eta(0)\tau]} \tag{34}$$

at point $z = 0$ be injected into the medium at this point. Further, we consider the time scale on which relaxation can be neglected: $\gamma_{1,2} \longrightarrow 0$. In the perturbation theory developed above, we analyze the evolution of soliton (34) in z, which can be described by the system of equations (24) with pumping and perturbation ϵP. We assume that the latter includes losses to the first and third powers in the field amplitude and linear dispersion of the second order; i.e., the perturbation has the form

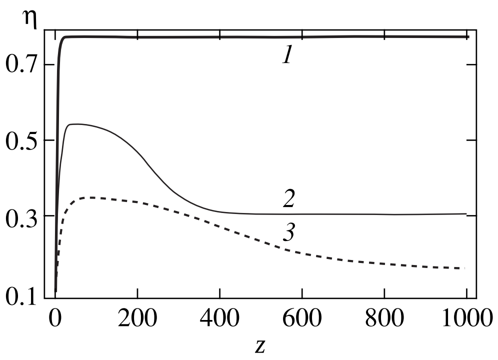


Fig. 2. Numerical solution of system (36): dependence of η on the length of the medium z; $\alpha_1 = 0.02$, $\alpha_2 = -0.02$, $\alpha_3 = 0.03$, $h_1(z) = 0.1$; $\Gamma_0 = 0.5$ (*1*), 2 (*2*) and 5 (*3*).

$$\epsilon P = -\alpha_1 E + \alpha_2 E_{zz} - \alpha_3 |E|^2 E, \tag{35}$$

where α_i are real coefficients. Solutions $\psi^{(m)}(\tau, \lambda_0)$ and $\varphi^{(m)}(\tau, \lambda_0)$ of the Zakharov–Shabat spectral problem for potential (34) are known (see, for example, [4, 5]). Substituting these expressions into Eq. (20) and integrating with respect to variable τ, we obtain the following system of equations for the imaginary and real components of the spectral parameter in the first approximation of perturbation theory:

$$\begin{aligned}\frac{\partial}{\partial z}\eta &= -2\eta\left(\alpha_1 + \frac{4}{3}\alpha_2\eta^2 + \frac{8}{3}\alpha_3\eta^3\right) \\ &\quad - \frac{16}{3}\alpha_2\eta\xi^2 + \frac{h_1(z)(\eta+\Gamma_0)}{(\eta+\Gamma_0)^2+\xi^2}, \\ \frac{\partial}{\partial z}\xi &= -\frac{16}{3}\eta^2\xi + \frac{h_1(z)\xi}{(\eta+\Gamma_0)^2+\xi^2}.\end{aligned} \tag{36}$$

The results of numerical solution of this system of equations are presented in Fig. 2.

It turns out that a change in the spectral width (equal to $2\Gamma_0$) might change qualitatively the dependence of the soliton amplitude on the amplification length. For example, this curve acquires a hump for $\alpha_1 = 0.01$, $\alpha_2 = 0.01$, $\alpha_3 = 0.02$, $h_1 = 0.1$ and for a value of Γ_0 lying in the interval 0.6–5 (see Fig. 2). For $\Gamma_0 \gtrsim 8.7$, the soliton amplitude asymptotically tends to zero. For smaller spectral widths, the soliton amplitude increases to a nonzero constant value upon an increase in z. The amplification efficiency decreases with increasing spectral width. Since the value of h depends on frequency, this effect is associated with losses and dispersion effects included in the perturbation P.

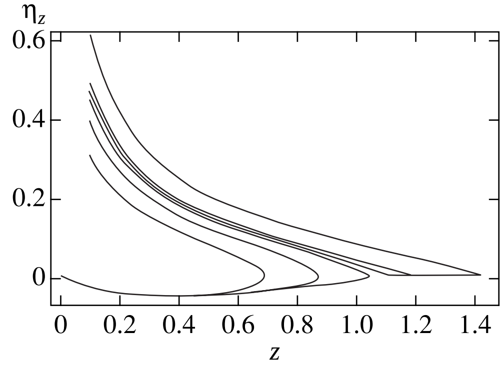


Fig. 3. Phase portrait: dependence of η_z on η for various values of $\xi(0)$; $\eta(0) = 0.1$.

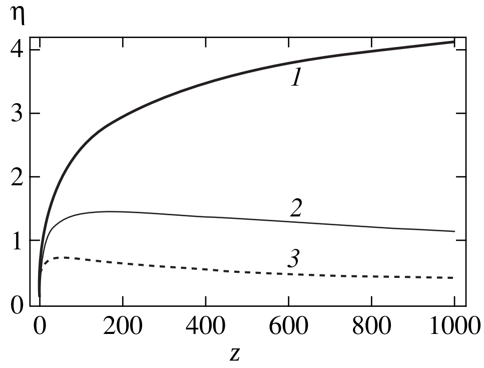


Fig. 4. Numerical solution of system (38): dependence of the soliton amplitude on the amplification length in the absence of perturbations; $\Gamma_0 = 3.0$, $\Gamma_1 = 3.0$, $\Gamma_2 = 3.1$, $h_1 = 33$, $h_2 = 33$ (*1*), 35 (*2*), and 38 (*3*).

The asymptotic form of the solution of system (36) is determined by the boundary values of the parameters $\eta(0)$ and $\xi(0)$ since the system has more than one stable focus in the phase space in the general case. The numerical analysis carried out for $h = 0.125$, $\Gamma_0 = 0.1$, $\alpha_2 = 0.01$, $\alpha_{1,3} = 0$, and various values of $\xi(0)$ revealed the existence of two stable foci $\eta_1 = 0$, $\eta_2 \approx 1.425$, and $(\eta_{1,2})_z = 0$ (see Fig. 3). It turned out that the phase trajectories may diverge in the vicinity of $\xi_c(\eta(0)) \approx -0.115$; i.e., for small variations of the boundary (for $z = 0$) detuning $\xi(0)$ of a soliton in the vicinity of ξ_c, the soliton amplitude tends either to $2\eta \approx 2.85$ or to 0 upon an increase in z. It was found using numerical methods that a decrease in the initial value of the soliton amplitude $\eta(0)$ leads to a decrease in the critical value $|\xi_c|$.

In real resonance media, processes leading to population inversion losses during the propagation of a soliton in the medium are possible apart from the upper level pumping. These losses may be caused by collisions of atoms in the gas, by collisions of atoms in the gas, and by the population of the lower level. In the general case, the rates of pumping and losses are characterized by different frequency dependences, which is due to different absorption cross sections, relaxation constants, and other physical factors [9]. For this reason, the frequency dependence of not only the pumping rate but also the loss rate (cf. Eq. (30)) must be taken into account. In the system of Maxwell–Bloch equations under investigation, the losses in the population of level 1 are equivalent to the pumping of the lower lasing level (see Fig. 1). This pumping may be due to transition from level 3 to level 2 and may depend on frequency just like the pumping of the upper laser level.

The numerical results presented below indicate that the inclusion of the frequency dependence of pumping rate leads to a number of peculiarities in soliton amplification. For example, in the presence of nonuniform broadening, soliton amplification is also possible in the case when the integral rates of pumping and losses in the inverse population of a laser transition (i.e., the rates integrated over the entire frequency spectrum) are identical. Under certain conditions, soliton amplification is possible on a bounded segment when the integrated losses prevail over the integrated pumping.

We choose the Lorentzian form of the frequency dependence of rates of pumping and population inversion losses (cf. (30)):

$$h(z;\,\omega) = \frac{h_1(z)}{\pi}\frac{\Gamma_1}{\omega^2+\Gamma_1^2} - \frac{h_2(z)}{\pi}\frac{\Gamma_2}{\omega^2+\Gamma_2^2}. \qquad (37)$$

Here, $h_1 > 0$ corresponds to the pumping of the upper level, and $h_2 > 0$ corresponds to the loss of the population of the upper laser level 1 or the population of the lower level 2.

Equations (36) for the same perturbation (35) assume the form

$$\frac{\partial}{\partial z}\eta = -2\eta\left(\alpha_1 + \frac{4}{3}\alpha_2\eta^2 + \frac{8}{3}\alpha_3\eta^3\right) - \frac{16}{3}\alpha_2\eta\xi^2$$

$$- \operatorname{Im}[\mathcal{H}_1(\eta, \xi) + \mathcal{H}_2(\eta, \xi)],$$

$$\frac{\partial}{\partial z}\xi = -\frac{16}{3}\alpha_2\eta^2\xi - \operatorname{Re}[\mathcal{H}_1(\eta, \xi) + \mathcal{H}_2(\eta, \xi)], \qquad (38)$$

$$\mathcal{H}_m(\eta, \xi) = \frac{h_m(z)}{\pi}$$

$$\times\left[\frac{i(\Gamma_0+\Gamma_m) + i\eta + \xi}{(\Gamma_m+\Gamma_0)(i\Gamma_0 + i\eta + \xi)(i\Gamma_m + i\eta + \xi)}\right].$$

Since the rates of pumping and losses depend on detuning, their frequency characteristics play a significant role. A numerical analysis of system (38) proved that the soliton dynamics depends qualitatively on the relation between quantities h_1 and h_2 and the spectral widths $2\Gamma_{0,1,2}$. The results of numerical solution of Eqs. (38) obtained without taking into account perturbations ($\alpha_{1,2,3} = 0$) are presented in Fig. 4 for various coefficients h_1 and h_2. For $\Gamma_2 < \Gamma_1$ and $h_1 = h_2$, the soliton amplitude vanishes at a distance $z \sim (h_1)^{-1}$. It can be seen from Fig. 4 that soliton amplification is possible in the case when $h_1 = h_2$. Moreover, a numerical analysis of Eqs. (38) revealed that an increase in the soliton amplitude at a finite initial interval is also possible for $h_1 < h_2$ provided that $\Gamma_1 < \Gamma_2$ and the laser transition is broadened nonuniformly.

The inclusion of perturbation leads to a qualitative change in the dependence of soliton parameters on z. For $\alpha_2 > 0$, the soliton amplitude increases at the initial stage and then tends to a constant asymptotic value. For $\Gamma_2 < \Gamma_1$, the asymptotic forms change for a relatively small variation of h_1 ($|\Delta h_1| \ll h_1$) and a fixed h_2. The soliton amplitude tend to a constant nonzero value for $h_1 > h_2$ and tends to zero upon an increase in z for $h_2 < h_2$ (see Fig. 5).

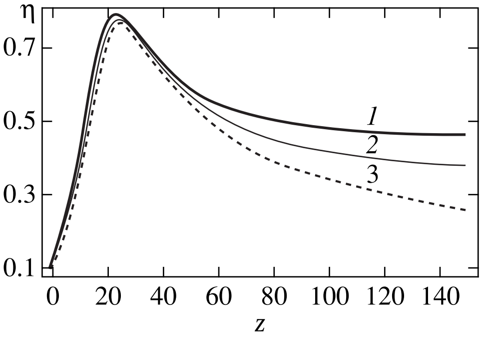


Fig. 5. The same as in Fig. 4; $\Gamma_0 = 3.0$, $\Gamma_1 = 3.0$, $\Gamma_2 = 3.1$, $\alpha_1 = \alpha_3 = 0$, $\alpha_2 = 0.02$, $h_1 = 3.1$, $h_2 = 3.0$ (*1*), $h_1 = h_2 = 3.1$ (*2*), and $h_1 = 3.0$, $h_2 = 3.1$ (*3*).

5. RANDOMIZATION IN AN AMPLIFYING MEDIUM

The systems of equations including the Maxwell–Bloch equations, linear and nonlinear dispersion, and also cubic terms in the field amplitude are employed for analyzing the effects of propagation of ultrashort solitons in optical fibers [20]. Stable field pulses in an optical fiber appear as a result of balance between dispersion and nonlinearity. In order to compensate the losses emerging during the propagation of a soliton in an optical fiber, various mechanisms are used, including those associated with the amplification of pulses over finite intervals of the fiber, arranged periodically over its length. Soliton amplification takes place, for example, when a soliton passes through a segment of the fiber with implanted atoms of rare-earth metals with a resonance two-level transition. Soliton dynamics in such a medium can be described in a model close to the integrable system of Maxwell–Bloch equations taking into account pumping depending on the coordinate and frequency. The dispersion and Kerr nonlinearity in the system of Maxwell–Bloch equations can be taken into account either in perturbation theory or by using a model combining this system with the nonlinear Schrödinger equation. The integrability of such a combined model requires that a stringent condition be imposed on physical parameters [2, 21, 22]. However, the application of the combined model is expedient for determining the qualitative features of the strongly nonlinear soliton dynamics, including the random behavior of ultrashort pulses.

Nozaki and Bekki [23] investigated numerically the emergence of random dynamics under the action of a small driving force in the form of the sum of two harmonics and quadratic dispersion for the soliton solution of the nonlinear Schrödinger equation, which has the following form in our notation:

$$E(z, \tau) = \frac{2\eta \exp(i\phi)}{\cosh(2\eta\tau)}. \tag{39}$$

In our variables, this perturbation has the form

$$\epsilon P_{NB} = \epsilon_1 \exp(ik_1 z) + \epsilon_2 \exp(ik_2 z) - \alpha_2 E_{\tau\tau}.$$

The authors of [23] solved equations for soliton parameters which emerge in the first order of perturbation theory and found that period-doubling bifurcations take place upon an increase in the values of parameters, followed by the emergence of a random dependence of these parameters on z. Under pumping conditions, we can expect a change in the conditions for the emergence of random dynamics in the model combining system (24) and the nonlinear Schrödinger equation.

On account of the interaction with two-level atoms, the evolution of ultrashort optical pulse in an optical fiber can be described by the following Maxwell equation for a slow amplitude U of the light field:

$$U_t = U_x + i\frac{k''}{2}U_{tt} + i\chi_3|U|^2U = L_0\langle R_+\rangle. \tag{40}$$

Here, k'' and χ_3 are the coefficients characterizing the dispersion and Kerr nonlinearity, respectively, and L_0 is the resonance absorption length. The physical values of these coefficients are given, for example, in [21, 22].

The integrable combination of the Maxwell–Bloch equations and Eq. (40) was obtained in [21, 22] in the absence of pumping. In the variables used in system (24), the combined model has the form

$$\partial_z E = i\chi_0\left(\frac{1}{2}E_{\tau\tau} + E|E|^2\right) + \langle R_+\rangle, \tag{41}$$

$$\partial_\tau R_+ = 2R_3 E + 2i\omega R_+, \tag{42}$$

$$\partial_\tau R_3 = -[R_+^* E + R_+ E^*] + \tilde{h}(z; \omega). \tag{43}$$

The integrability of this combined system requires the fulfillment of a certain additional condition imposed on physical parameters, which emerges when the system of Maxwell–Bloch equations and Eq. (40) is reduced to the form (41)–(43). The physical situations in which this condition holds were proposed in [21, 22]. Since

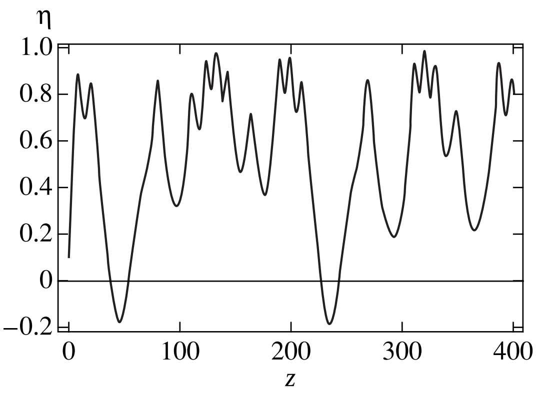


Fig. 6. Numerical solution of system (47): dependence of η on z for $\beta = 1$, $h_0 = 0.125$, $\alpha_3 = 0.03$, $\Gamma_0 = 1$, $\epsilon_1 = 0.04$, $\epsilon_2 = 0$, $\chi_0 = 0.225$, and $k_0 = k_1 = 0.5$.

the aim of this section is an analysis of qualitative characteristics of soliton amplification in the presence of pumping, we will not specify the physical conditions. It should be noted that the existence of a soliton solution of the combined model was proved numerically and experimentally in [24, 25].

The Lax representation for system (41)–(43) has the form

$$\frac{\partial}{\partial\tau}\Psi = \begin{pmatrix} -i\lambda & E \\ -E^* & i\lambda \end{pmatrix}\Psi, \tag{44}$$

$$\begin{aligned}\frac{\partial}{\partial z}\Psi &= \frac{1}{2}\left\langle \frac{i}{\lambda+\omega}\begin{pmatrix}\rho_{11} & -\rho_{12}\\ -\rho_{21} & -\rho_{11}\end{pmatrix}\right\rangle\Psi \\ &+ \chi_0\begin{pmatrix} -i\lambda^2 - \frac{i}{2}|E|^2 & \lambda E + \frac{i}{2}E_\tau \\ -\lambda E^* + \frac{i}{2}E^*_\tau & i\lambda^2 + \frac{i}{2}|E|^2 \end{pmatrix}\Psi.\end{aligned} \tag{45}$$

The fundamental difference of this Lax representation from the similar representation given in [21, 22] lies only in that the spectral parameter satisfies additionally Eq. (27).

We assume that the dependence of the pumping rate $\tilde{h}$ on the coordinate and frequency is given by Eq. (31). Let us suppose that perturbation ϵP has the form

$$\epsilon P_c = \epsilon_1\exp(ik_1 z) + \epsilon_2\exp(ik_2 z) - \frac{1}{2}\alpha_3|E|^2E. \tag{46}$$

This perturbation may describe the action of a constant external field (or a combination of fields) with wave detunings k_1 and k_2 from the wave vector of field E and cubic nonlinear losses. In the framework of the perturbation theory developed above, we obtain for perturbation (46) the following analog of the Nozaki–Bekki equations [23] for the soliton parameters (39):

$$\begin{aligned}\frac{\partial}{\partial z}\eta &= -\frac{8}{3}\alpha_3\eta^3 - \frac{\pi}{2}[\epsilon_1\sin(k_1 z - \varphi) \\ &+ \epsilon_2\sin(k_2 z - \varphi)] + \frac{h_1(z)}{\eta + \Gamma_0},\end{aligned} \tag{47}$$

$$\frac{\partial}{\partial z}\varphi = 4\chi_0\eta^2.$$

We choose the periodic dependence of the pumping rate on z:

$$h_1(z) = h_0[1 - \beta + \beta\cos(k_0 z)]^2 > 0.$$

For the above expression (30) for the pumping rate, this dependence corresponds to a standing wave of the field:

$$E_{13}(z) = E_{13}(0)[1 - \beta + \beta\cos(k_0 z)].$$

The periodic coordinate dependence of the pumping rate leads to a qualitative difference of the conditions for a transition to random dynamics from the corresponding conditions given in [23]. In particular, for $\beta \neq 0$, one harmonic of the external force ($\epsilon_1 \neq 0$, $\epsilon_2 = 0$) and the coincidence of the wave vectors ($k_0 = k_1$) are sufficient for the emergence of a random attractor. For $h_0 \approx 0.12$, random dynamics was discovered for $\epsilon_1 = \epsilon_2 = 0$, $k_0 = 0.1$, $\beta = 1$, and $\alpha_3 = 0.01$. It should be recalled that, in the model analyzed in [23], a biharmonic driving force is required for the emergence of random dynamics.

Figure 6 shows the results of the numerical solution of system (47) for the soliton parameter η. A phase portrait demonstrating a random attractor for the same values of parameters as in Fig. 6 is presented in Fig. 7.

The numerical analysis proved that, for the pumping rate $h_0 \gtrsim 0.1$, period-doubling bifurcations and random dynamics appear for values of $\epsilon_{1,2} \neq 0$ two orders of magnitude lower than in the absence of pumping. It was found that, in the presence of pumping independent of z, the amplitudes of random modulation do not increase. On the contrary, the amplitudes of fluctuations increase significantly (by a factor of several units) when periodic coordinate dependence of the pumping rate is taken into account.

For $h_0(z) = \text{const}$, $\alpha_3 > 0$, and $\epsilon_{1,2} > 0.01$, the soliton amplitude for large values of z attains a quasi-stationary regime, i.e., oscillates randomly about its mean value. The results of numerical analysis proved that quasi-random dynamics is also exhibited in the absence of perturbation of the form $-\alpha_i|E|^3E$ since the pumping rate is a nonlinear function of parameter η. When linear losses $-\alpha_1 E$ are taken into account, the amplitude of random pulsations about the mean value decreases, but does not attenuate upon an increase in z.

Concluding the section, we note that a similar random dynamics may also be observed when the first and second terms on the right-hand side of Eq. (41) are

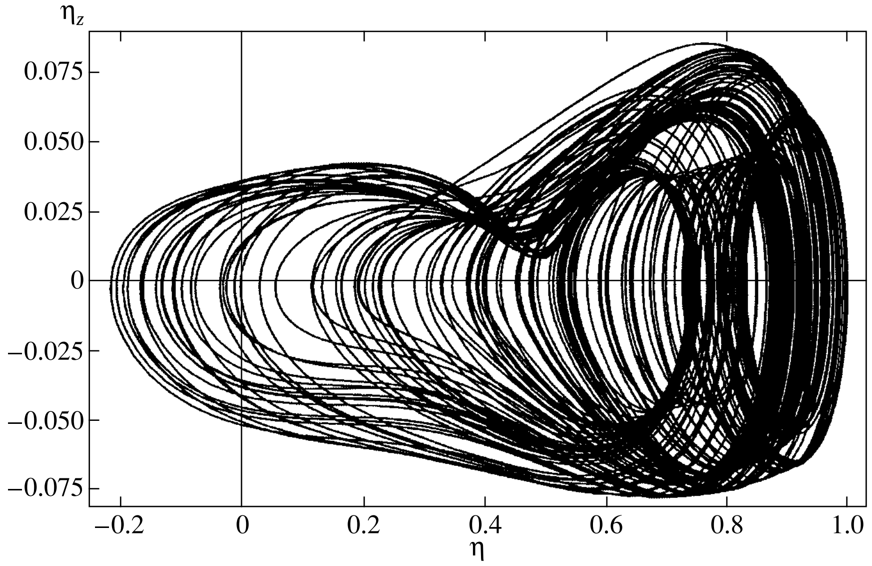


Fig. 7. Phase portrait: dependence of η_z on η. The values of parameters are the same as in Fig. 6.

taken into account in the form of perturbations for the small coefficients χ_0 and a sufficiently high pumping rate ($\sim h_0$).

6. CONTRIBUTION OF RADIATIVE SOLUTION

In the course of field evolution in lasing media, nonlinear pulsations constituting the radiative component of the solution are amplified along with solitons. In the case of completely integrable model (24), this solution is associated with the continuous real spectrum of the Zakharov–Shabat problem (25) and may be initiated by fluctuations of the medium. For the model (24) investigated by us here, the radiative solution can make a significant (and even dominant) contribution to the general solution. In this connection, the analysis of the radiative component of the solution is required when the dynamics of solitons being amplified is investigated. It should be noted that this case is similar to that of a long laser amplifier in which, according to Manakov [12], the asymptotic form of the solution in the model of field amplitude under certain initial conditions is described by the Painlevé III transcendent and is proportional to z.

One of the solutions of the exactly integrable version of the Maxwell–Bloch equations with pumping under certain initial conditions in the case when the relations

$$\begin{gathered} \gamma_{1,2} = \epsilon = 0, \quad g(\omega) = \delta(\omega), \\ \tilde{h}(z;\omega) = c = \text{const}, \end{gathered} \tag{48}$$

hold is proportional to the self-similar solution which can be expressed in terms of the Painlevé V transcendent [15–17].

Following [15, 16], we introduce the self-similar variables

$$\zeta = \tau\sqrt{2cz}$$

and the functions

$$\begin{gathered} E = \sqrt{2cz}\mathscr{E}(\zeta), \quad R = \sqrt{c/2z}\mathscr{R}(\zeta), \\ N = \sqrt{c/2z}\mathscr{N}(\zeta). \end{gathered} \tag{49}$$

This transforms system (24) to the following system of equations:

$$\begin{gathered} \frac{\partial}{\partial\zeta}(\zeta\mathscr{E}(\zeta)) = \mathscr{R}(\zeta) + \epsilon\sqrt{\frac{2z}{c}}P(\sqrt{2cz}\mathscr{E}(\zeta)), \\ \frac{\partial}{\partial\zeta}\mathscr{R}(\zeta) = \mathscr{N}(\zeta)\mathscr{E}(\zeta) - \frac{\gamma_2}{\sqrt{2cz}}\mathscr{R}(\zeta), \\ \frac{\partial}{\partial\zeta}\mathscr{N}(\zeta) = -\mathscr{R}(\zeta)\mathscr{E}(\zeta) + 1 - \frac{\gamma_1}{\sqrt{2cz}}\mathscr{N}(\zeta). \end{gathered} \tag{50}$$

We assume that perturbation ϵP is a function of the field amplitude and its derivatives.

Let us prove that, for small $\gamma_{1,2}$ and ϵ, the initial and boundary conditions

$$|R(z,0)| \ll 1, \quad N(z,0) = -n_0, \quad E(0,\tau) = 0, \tag{51}$$

and large values of ζ, the radiative solution of system (50) completely determines the asymptotic form.

System (24) has the following linear solutions:

$$E = R = 0, \quad N = -n_0 + c\tau, \tag{52}$$

where $-n_0 < 0$ is the initial difference in the populations of levels 1 and 2, and

$$E = \sqrt{2cz}, \quad R = \sqrt{c/2z}, \quad N = 0. \qquad (53)$$

The radiative solution is initiated by medium fluctuations and describes the nonlinear stage of the development of instability. For small fluctuations (as in the case of a laser amplifier [12]), the emerging shape of the leading front and the characteristic delay can be determined from the linear solution for small deviations of the field amplitude and polarization from state (52). For the linear stage of development of instability and for $g(\omega) = \delta(\omega)$, $\epsilon = 0$, we obtain from system (24)

$$\partial_z \tilde{E} = \tilde{R}, \qquad (54)$$

$$\partial_\tau \partial_z \tilde{E} = \left[-n_0 + \frac{c}{\gamma_1}(1 - e^{-\gamma_1 \tau})\right]\tilde{E} + f(z, \tau)e^{\gamma_2 \tau}. \qquad (55)$$

Here,

$$\tilde{E} = Ee^{-\gamma_2 \tau}, \quad \tilde{R} = Re^{-\gamma_2 \tau},$$

and f is a function describing fluctuations of the polarization of the medium. In analogy with the theory of laser amplifiers developed using the semiclassical theory [26], the initiation of the process can be described by simulating the initial polarization of the stochastic function with a small amplitude of fluctuations.

Using the substitution

$$\tilde{\tau} = -n_0 \tau + \frac{c}{\gamma_1}\left[\tau - \frac{1}{\gamma_1}(1 - e^{-\gamma_1 \tau})\right],$$

we can reduce Eq. (55) to the telegraph equation. For $\gamma_1 \longrightarrow 0$ and $\tau \gg n_0$, the solution to this equation has the form

$$E(z, \tau) = \int_0^{\chi}\int_0^{\tau} f(z', \tau') \times I_0(\sqrt{2c(\tau^2 - \tau'^2)(z - z')})dz'd\tau'. \qquad (56)$$

For a very slow function $f(z, \tau)$ (which increases not faster than according to a power law), solution (56) describes the leading front of the quasi-self-similar solution of the initial model (24). For $\zeta \gg 1$, we obtain

$$E \approx \frac{z \operatorname{sgn} f_0}{\zeta} e^{\zeta - \zeta_0}\left[1 + O\left(\frac{1}{\sqrt{\zeta}}\right)\right]. \qquad (57)$$

Here, $\zeta_0 = \ln|f_0|$ and f_0 is a certain mean value of the amplitude of fluctuations. A more rigorous computation of stochastic averages leads to the same result.

For $\mathcal{R}(0) \ll 1$ and $\epsilon = \gamma_1 = 0$, system (50) in the linear limit describing small deviations from state (52) can be reduced to the equation

$$\zeta \partial_\zeta^2 \mathcal{E} + \partial_\zeta \mathcal{E} = (\zeta - n_0)\mathcal{E}, \qquad (58)$$

which can be transformed to a degenerate hypergeometric equation. The solution to this equation for small ζ and large values of n_0 consists of oscillations attenuating upon an increase in ζ, which are transformed, for $\zeta > \zeta_c$, into a rapidly increasing solution proportional to the Bessel function $I_0(\zeta)$ (ζ_c is the point of joining of the attenuating and increasing solutions). For $\zeta > \zeta_c \gg 1$, we obtain the following solution:

$$E \approx \frac{z \operatorname{sgn} \mathcal{E}_c}{\zeta} e^{\zeta - \zeta_c}\left[1 + O\left(\frac{1}{\sqrt{\zeta}}\right)\right]. \qquad (59)$$

The value of $\operatorname{sgn} \mathcal{E}_c$ is determined by the joining of the increasing component with the attenuating initial component of the solution to Eq. (58). Solutions (57) and (59) coincide accurate to the delay; i.e., the asymptotic form of the solution to Eqs. (24) for the initial and boundary conditions differing insignificantly from conditions (52) is joined with the quasi-self-similar solution to Eq. (50). Thus, when conditions (48) are satisfied, the solution of system (24) in the nonlinear regime is described by the nonlinear solution (50) which can be reduced to the solution to the following Painlevé V equation:

$$u_{\xi\xi} = (u_\xi)^2 \frac{3u - 1}{2u(u - 1)} - \frac{1}{\xi}u_\xi + \frac{\gamma u}{\xi} - \frac{u(u + 1)}{2(u - 1)}, \qquad (60)$$

where $\xi = 2\zeta$ and

$$\mathcal{E} = \frac{u - u_\xi}{\sqrt{u}(u - 1)} e^{i\varphi}. \qquad (61)$$

Using the linear solution (52), we choose

$$\mathcal{N}(0) = 0, \quad \mathcal{E}(0) = \mathcal{R}(0) = \mathcal{R}_0, \quad \operatorname{Im}\mathcal{R}_0 = 0.$$

The initial conditions for u and the parameters of the equation are determined by joining with this linear solution and with the conditions of regularity for $\xi = 0$; in fact [15–18], these conditions have the form

$$u(0) = -1, \quad \varphi = \pi/2,$$
$$u_\xi(0) = 1 + 2\mathcal{R}_0, \quad \gamma = -1 - 2\mathcal{R}_0. \qquad (62)$$

It should be noted that solution (59) coincides with that for a laser amplifier [12] correct to the self-similar variable and the phase factor. In particular, this solution shows that for small $\mathcal{R}_0$, the amplification is delayed as in the case of a laser amplifier. At the same time, an additional delay in the generation of the radiative solution is observed in the case of a three-level laser with pumping due to the initial population of the lower laser level:

$$R_3(z, 0) = -n_0 < 0.$$

Let us now analyze the role of a small perturbation in the formation of the self-similar radiative asymptotic form and compare these results with those for the soliton amplitude. The numerical solution of systems (36) and (50) for $\epsilon = \gamma_{1,2} = 0$ is illustrated in Fig. 8. It can be seen that there exists a certain interval of values $[0, z_d \approx 16]$, in which the contribution of the continuous spectrum can be neglected. The length z_d of this interval increases with n_0 and upon a decrease in $|\mathcal{R}_0|$. In the general case, in the absence of a perturbation, for $\gamma_{1,2} = 0$, and for large values of z, the soliton solution starting with a small initial amplitude is of the same order as the pulses of the radiative solution. It can be seen from Fig. 8 that the contribution of the radiative solution to the soliton dynamics for $z > z_d$ cannot be disregarded in the general case. At the same time, according to the results of numerical analysis, the behaviors of the soliton and radiative solutions in the presence of perturbation and pumping differ qualitatively. This enables us to determine the conditions under which the amplitude of the soliton solution is much larger than the amplitude of pulses constituting the radiative solution for all values of z. In this case, the contribution of the radiative solution can be disregarded while studying the evolution of solitons.

The inclusion of small perturbations of type (35) changes qualitatively the radiative asymptotic form. We fix point ζ_1 at the leading front of a pulse from the radiative solution and carry out a transition to the reference frame attached to this point. It can be seen from system (50) that the perturbation in this reference frame increases with z in proportion to the field amplitude, while the contribution from relaxation terms decreases in proportion to $T_{1,2}^{-1}$. Numerical calculations revealed that perturbations containing losses in the form of odd powers of the field amplitude, e.g.,

$$\epsilon\mathcal{P} = -\alpha_1 E - \alpha_3 |E|^2 E - \alpha_4 |E|^4 E, \quad \alpha_i > 0,$$

lead to a decrease in the radiative contribution for large z. This suppression is the stronger, the larger the delay, i.e., the larger the initial population of the lower lasing level and the smaller $|\mathcal{R}_0|$.

At the same time, it follows from the previous section that small perturbations may lead to a qualitative change in the asymptotic form of the soliton solution. Figure 9 shows the results of numerical analysis of system (24) for the soliton mode and of system (50) for the quasi-radiative mode, for which the perturbation has the form

$$\epsilon P = -\alpha_1 E - \alpha_3 |E|^2 E.$$

The initial and boundary conditions in this case are given by

$$2\eta(0) = 0.1, \quad R_0 = 0.001, \quad n_0 = 1.$$

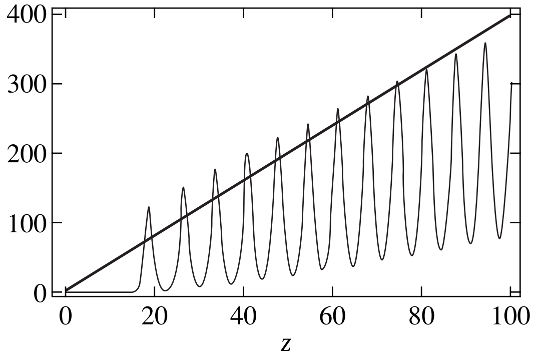


Fig. 8. Numerical solution of systems (36) and (50) for $\epsilon = 0$, $c = 1$, $n_0 = -1$, and $\alpha_{1,2,3} = 0$. Dependences of the maximum soliton intensity (bold line) and of the intensity $|E|^2$ of the radiative solution (fine curve) on z.

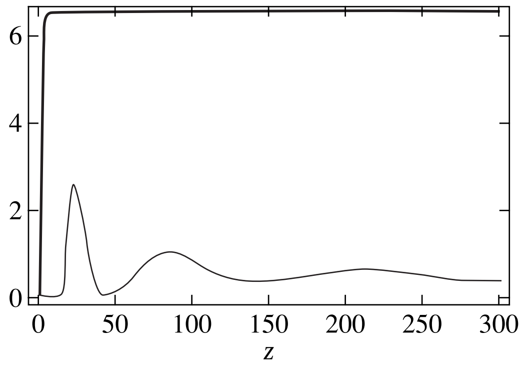


Fig. 9. Numerical solution of systems (24) and (50) for $\alpha_1 = 0.01$, $\alpha_2 = 0$, and $\alpha_3 = 0.05$.

The numerical analysis was carried out for various values of $\alpha_{1,3}$ and h. It was found that smaller values of $\alpha_{1,3} > 0$ correspond to relatively weaker suppression of the radiative contribution. The change in h considerably affects the radiative solution. For $\alpha_{1,3} \sim 10^{-1}$–$10^{-4}$, $c \sim 1$, and large values of z, the amplitude of radiative solution pulsations does not exceed 10% of the soliton amplitude for all z. This fraction decreases with decreasing c. As the value of c increases and $\alpha_{1,3} \neq 0$ decreases, the ratio of maximum amplitudes of radiative solution pulsations to the soliton amplitude increases and may exceed unity on a certain interval of z (cf. Fig. 8).

The above estimates of the radiative solution delay show that the three-level scheme of laser pumping in which the lower level of the laser transition coincides with the ground level is preferable for observing the soliton dynamics since the radiative solution delay in this case may exceed that for the four-level scheme of pumping considerably (by an order of magnitude). At the same time, relaxation effects as well as linear and nonlinear losses lead to strong suppression of the radiative component of the solution. Conversely, in the case of zero initial population of the lower laser level and higher values of the pumping coefficient $c \gg 1$ and in the absence of losses, the contribution of the continuous spectrum of the problem must be taken into account in an analysis of the soliton dynamics.

7. CONCLUSION

The solution of systems of equations close to integrable makes it possible to study the nonlinear stage of the evolution of solitons and other nonlinear modes. The perturbation theory apparatus presented by us here supplements the list of such models with the models close to deformed integrable systems of equations. Some of the integrable deformed models have important physical applications (e.g., the Heisenberg cylindrical equation and the equations of the principal chiral field, which are used in gravitation theory [6]).

The Maxwell–Bloch equations close to system (24) and describing the interaction of a light field with a resonance medium under pumping are used quite frequently [9]. The solutions of the idealized integrable version of this model may be regarded as nonlinear modes (solitons) strongly "deformed" during the evolution since the spectral parameter depends on variables. The inclusion of small perturbations leads to the emergence of additional terms in the equations describing the dependence of soliton parameters on the variables. It has been demonstrated by us here that a combination of these properties leads to a qualitatively new dynamics of nonlinear modes. In particular, soliton amplification modes appear, which cannot be described by using the standard perturbation theory [4, 5].

On the other hand, conditions of type (9) in the perturbed model can be satisfied approximately for a special choice of perturbation. This enables us to extend the physical range of applicability of models integrable in the framework of the MISP with a variable spectral parameter.

The qualitative effects observed by us and manifested during the propagation of a soliton in a medium with pumping may be applied in an analysis of lasing and for choosing lasing conditions.

ACKNOWLEDGMENTS

This work was supported financially by the Russian Foundation for Basic Research (project no. 98-02-17904).

REFERENCES

1. S. P. Novikov, S. V. Manakov, L. P. Pitaevskii, and V. E. Zakharov, *Theory of Solitons: the Inverse Scattering Method* (Nauka, Moscow, 1980; Consultants Bureau, New York, 1984).
2. A. I. Maimistov and A. M. Basharov, *Nonlinear Optical Waves* (Kluwer, Dordrecht, 1999).
3. D. J. Kaup, Stud. Appl. Math. **31**, 121 (1976).
4. V. I. Karpman and E. M. Maslov, Zh. Éksp. Teor. Fiz. **73**, 537 (1977) [Sov. Phys. JETP **46**, 281 (1977)].
5. Yu. S. Kivshar and B. A. Malomed, Rev. Mod. Phys. **61**, 763 (1989).
6. S. P. Burtsev, A. V. Mikhaĭlov, and V. E. Zakhapov, Teor. Mat. Fiz. **70**, 323 (1987).
7. S. P. Burtsev and I. R. Gabitov, Phys. Rev. A **49**, 2065 (1994).
8. S. P. Burtsev, I. R. Gabitov, and V. E. Zakharov, in *Plasma Theory and Nonlinear and Turbulent Processes in Physics* (World Scientific, Singapore, 1988), p. 897.
9. O. Svelto, *Principles of Lasers* (Plenum, New York, 1976; Mir, Moscow, 1979).
10. A. A. Apolonsky, A. A. Zabolotskii, V. P. Drachev, and E. I. Zinin, Proc. SPIE **2041**, 385 (1994).
11. Q. H. F. Vrehen, in *Laser Spectroscopy*, Ed. by H. Walther and K. W. Rothe (Springer-Verlag, Berlin, 1979), Vol. 4, p. 79.
12. S. V. Manakov, Zh. Éksp. Teor. Fiz. **83**, 68 (1982) [Sov. Phys. JETP **56**, 37 (1982)].
13. F. Y. F. Chu and A. C. Scott, Phys. Rev. A **12**, 2060 (1975).
14. A. A. Zabolotskiĭ, Zh. Éksp. Teor. Fiz. **115**, 1168 (1999) [JETP **88**, 642 (1999)].
15. *Painlevé Transcendents. Their Asymptotics and Physical Applications*, Ed. by P. Winternitz and D. Levi (Plenum, New York, 1992), NATO ASI Ser., Ser. B **278** (1992).
16. S. P. Burtsev, Phys. Lett. A **177**, 341 (1993).
17. A. V. Kitaev, A. V. Rybin, and J. Timonen, J. Phys. A **26**, 3583 (1993).
18. H. Steudel and I. Leonhardt, Opt. Commun. **107**, 88 (1994).
19. A. V. Rybin, J. Phys. A **24**, 5235 (1991).
20. *Cambridge Studies in Modern Optics*, Vol. 10: *Optical Solitons—Theory and Experiment*, Ed. by J. R. Taylor (Cambridge Univ. Press, Cambridge, 1992).
21. A. I. Maĭmistov and É. A. Manykin, Zh. Éksp. Teor. Fiz. **85**, 1177 (1983) [Sov. Phys. JETP **58**, 685 (1983)].
22. E. A. Doktorov and R. A. Vlasov, Opt. Acta **30**, 2 (1983).
23. K. Nozaki and N. Bekki, Phys. Lett. A **102A**, 383 (1984).
24. M. Nakazawa, E. Yamada, and H. Kubota, Phys. Rev. Lett. **66**, 2625 (1991).
25. M. Nakazawa, Y. Kimura, K. Kurokawa, and K. Suzuki, Phys. Rev. A **45**, 23 (1992).
26. F. Haake, H. King, G. Schröder, *et al.*, Phys. Rev. A **20**, 2047 (1979).

Translated by N. Wadhwa