Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 1002–1005.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1144–1148.

SOLIDS
Structure

Possible Use of Nuclear Magnetic Resonance for Studying Magnetoelectric Effects in the Mn_2Sb Ferromagnet

M. I. Kurkin, I. F. Mirsaev, and E. A. Turov

Institute of Metal Physics, Ural Division, Russian Academy of Sciences, ul. S. Kovalevskoĭ 18, Yekaterinburg, 620219 Russia

e-mail: kurkin@imp.uran.ru; mirsaev@imp.uran.ru; turov@imp.uran.ru

Received October 8, 2003

Abstract—The possibilities of resonance excitation of nuclear spins by an alternating electric field (nuclear magnetoelectric resonance) in the Mn_2Sb ferromagnet are analyzed as applied to the studying of magnetoelectric effects in this compound. *© 2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A symmetry description of the magnetoelectric effect, which manifests itself in the magnetization of a substance by electric field **E** and its electric polarization by magnetic field **H** [1], uses invariant (with respect to the symmetry operations of a crystal) combinations of the ferromagnetism **M**, antiferromagnetism **L**, and electric polarization **P** vectors [2], which enter into the equation for the thermodynamic potential in the form of summands,

$$\Phi_{MP} = -\gamma^{\alpha\beta\nu} M^\alpha L^\beta P^\nu, \quad (1)$$

where $\gamma^{\alpha\beta\nu}$ are the magnetoelectric tensor components, relations between which are determined by crystal symmetry [1]. As follows from (1), polarization in electric field **E**, $P^\nu = \kappa^{\nu\delta} E^\delta$ (where $\kappa^{\nu\delta}$ is the electric susceptibility tensor components), results in the appearance of the magnetization

$$H^\alpha_{\text{eff}} = -\frac{\partial\Phi}{\partial M^\alpha} \propto L^\beta P^\nu = L^\beta \kappa^{\nu\delta} E^\delta,$$

$$M^\lambda = \chi^{\lambda\alpha} H^\alpha_{\text{eff}} \propto \chi^{\lambda\alpha} L^\beta \kappa^{\nu\delta} E^\delta, \quad (2)$$

where $\chi^{\lambda\alpha}$ is the magnetic susceptibility tensor components. Accordingly, magnetization in a magnetic field, $M^\alpha = \chi^{\alpha\delta} H^\delta$, results in the appearance of the electric polarization

$$E^\nu_{\text{eff}} = -\frac{\partial\Phi}{\partial P^\nu} \propto M^\alpha L^\beta = \chi^{\alpha\delta} H^\delta L^\beta,$$

$$P^\lambda = \kappa^{\lambda\nu} E^\nu_{\text{eff}} \propto \kappa^{\lambda\nu} \chi^{\alpha\delta} H^\delta L^\beta. \quad (3)$$

It follows from (2) and (3) that both values, M^λ and P^λ, are only nonzero at $L^\beta \neq 0$. For this reason, the magnetoelectric effect is easiest to study in antiferromagnets [2], which are characterized by large **L** and zero **M** vectors. Conversely, in ferromagnets with two or more sublattices, the **M** vector is nonzero, whereas **L** = 0. As a result, the magnetoelectric effect in these substances is related to **L** vector changes. These changes cannot be recorded in static magnetic measurements, and the magnetoelectric effect in ferromagnets cannot therefore be studied by the existing experimental methods. This work is concerned with the Mn_2Sb compound, which is a four-sublattice ferromagnet [3] and whose magnetic structure is therefore characterized by both vectors (**M** and **L**). Nevertheless, we will show that experimental studies of the magnetoelectric effect in this compound involve the same difficulties as with ferromagnets.

One of the approaches to solving this problem is through using alternating electric **E**(*t*) and magnetic **H**(*t*) fields with near magnetic resonance frequencies. The feasibility of using field **E**(*t*) at the antimagnon resonance frequency, when pure antiferromagnetism vectors **L** are only excited, whereas the total magnetization vector **M** remains constant, was discussed in [3]. The antimagnon resonance frequency range is determined by exchange interaction, which corresponds to 10^2–10^4 GHz [3]. In this work, we also discuss the feasibility of using field **E**(*t*) at the nuclear magnetic resonance frequency. In [4], this type of resonances was called nuclear magnetoelectric (NMER). Its frequency range, like that of NMR [5], is 10^2–10^3 MHz. The use of NMER for studying the magnetoelectric effect in Mn_2Sb is favored by the presence of only one manganese isotope Mn^{55} with a large nuclear magnetic moment of $3.5\mu_n$, where μ_n is the nuclear magneton [5].

2. THE MAGNETIC STRUCTURE AND THERMODYNAMIC POTENTIAL OF Mn_2Sb

The Mn_2Sb compound has a Curie point $T_C = 550$ K, easy-axis magnetic anisotropy above 240 K, and easy-plane anisotropy at $T < 240$ K. Its crystallographic symmetry is described by the space group $D_{4h}^7 \equiv P_4/nmm$. The manganese ions occupy two different twofold sites [3], MnI (site *a*) and MnII (site *c*), with different atomic magnetic moments, $\mu_I = (2.13 \pm 0.20)\mu_B$ and $\mu_{II} = (3.87 \pm 0.40)\mu_B$. The atomic moments are ordered ferromagnetically within the sublattices of sites *a* and *c* and antiferromagnetically between themselves. This corresponds to a ferrimagnetic exchange structure whose unit cell contains four ions with the ferromagnetism vectors $\mathbf{M}_{a1}$, $\mathbf{M}_{a2}$ and $\mathbf{M}_{c1}$, $\mathbf{M}_{c2}$. As sites *a* and *c* are not related by symmetry operations, this structure can conveniently be described in terms of two ferromagnetism (**M**) and two antiferromagnetism (**L**) vectors,

$$\mathbf{M}_a = \mathbf{M}_{a1} + \mathbf{M}_{a2}, \quad \mathbf{L}_a = \mathbf{M}_{a1} - \mathbf{M}_{a2}, \tag{4}$$

$$\mathbf{M}_c = \mathbf{M}_{c1} + \mathbf{M}_{c2}, \quad \mathbf{L}_c = \mathbf{M}_{c1} - \mathbf{M}_{c2}, \tag{5}$$

with the conditions $\mathbf{M}_a \downarrow\uparrow \mathbf{M}_c$ (in addition, $M_c > M_a$) and

$$\mathbf{M}_\xi^2 + \mathbf{L}_\xi^2 = 4M_{\xi 0}^2, \quad \mathbf{M}_\xi \cdot \mathbf{L}_\xi = 0, \quad \xi = a, c, \tag{6}$$

where $M_{a0} = M_{a1} = M_{a2}$ and $M_{c0} = M_{c1} = M_{c2}$ are the magnetizations of the sublattices. The total magnetization of Mn_2Sb is determined by the difference $\mathbf{M} = \mathbf{M}_c - \mathbf{M}_a$. To be specific, let the ferromagnetism vector in the ground state $\mathbf{M} \equiv \mathbf{M}_0$ be parallel to the *z* axis (easy axis); that is, $M_0^z = M_0$. In what follows, we only consider uniform oscillations of the $\mathbf{L}_\xi$ variables, because ferromagnetic oscillations of the $\mathbf{M}_\xi$ vectors are not excited by an electric field.

The magnetoelectric effects in Mn_2Sb are described in this work using the thermodynamic potential of the form

$$\Phi = \Phi_{\text{ex}} + \Phi_{ME} + \Phi_{HF}, \tag{7}$$

where

$$\Phi_{\text{ex}} = \frac{1}{2}J_a L_a^2 + \frac{1}{2}J_c L_c^2 \tag{8}$$

is the exchange part of the thermodynamic potential in the approximation quadratic in $\mathbf{L}_\xi$. Such a form was used in [3] to describe spin waves in Mn_2Sb excited by an alternating electric field. Work [3] also contained equations relating the J_a and J_c coefficients to the parameters of intra- and interstice interactions.

The next term in (7),

$$\Phi_{ME} = -[D_a(L_a^x E^x - L_a^y E^y) + d_a(l_a^x E^x - l_a^y E^y) + D_c(L_c^x E^x + L_c^y E^y) + d_c(l_c^x E^x + l_c^y E^y)]\frac{M_z^0}{M_0}, \tag{9}$$

describes the interactions of the antiferromagnetism vectors of electrons ($\mathbf{L}_\xi$) and nuclei ($\mathbf{l}_\xi$) with electric field **E**. Here, D_a, D_c, d_a, and d_c are the corresponding constants for these interactions. (Recall that $M_z^0/M_0 = 1$, and this multiplier is only retained in (9) as a test for invariance.) As distinct from (1), calculations by (9) were simplified by using invariant combinations in which polarization vector **P** components were replaced by electric field **E** components. The vectors $\mathbf{l}_\xi$ are related to the nuclear magnetizations of the sublattices, $\mathbf{m}_{\xi 1}$ and $\mathbf{m}_{\xi 2}$, by equations similar to (5).

The last term in (7) describes hyperfine interactions of the magnetizations $\mathbf{m}_{\xi 1}$ and $\mathbf{m}_{\xi 2}$ with $\mathbf{M}_\xi$ and $\mathbf{L}_\xi$,

$$\Phi_{HF} = -\frac{1}{2}\sum_\xi A_\xi(\mathbf{M}_\xi \mathbf{m}_\xi + \mathbf{L}_\xi \mathbf{l}_\xi). \tag{10}$$

The longitudinal components $M_{\xi 1}^z = M_{\xi 2}^z = \pm M_{\xi 0}$ determine the equilibrium values of hyperfine fields at the nuclei [5, 6],

$$H_{n\xi}^z = \pm A_\xi M_{\xi 0}, \tag{11}$$

and the eigenfrequencies of vector $\mathbf{m}_{\xi 1}$ and $\mathbf{m}_{\xi 2}$ oscillations (NMR frequencies),

$$\omega_{n\xi} = \gamma_n \left| H_{n\xi}^z \right| = \gamma_n A_\xi M_{\xi 0}, \tag{12}$$

where $\gamma_n = 1.05$ MHz /kOe is the gyromagnetic ratio for ^{55}Mn nuclei [5]. In (11), the minus sign refers to the site $\xi = a$, and the plus sign, to the site $\xi = c$. In review [7], the following values were given for these frequencies: $\omega_{na} = 126.26$ MHz and $\omega_{nc} = 143.7$ MHz. Their difference is much larger than the width of NMR lines, which allows us to study the resonance behavior of nuclear spins in the two sites independently. Below, we give the results obtained for the site $\xi = c$, for which the Φ_{ME} term [see (9)] is symmetric with respect to the replacement of *x* by *y*. For the site $\xi = a$, this term is antisymmetric with respect to this permutation. For this reason, the NMER signals from the nuclei in sites $\xi = a$ should be described by the formulas given below with E^y replaced by $-E^y$. For brevity, the ξ index will be omitted. With these replacements, the analysis given below is valid for a two-sublattice ferromagnet.

3. EFFECTIVE MAGNETIC FIELDS AT NUCLEI GENERATED BY ALTERNATING ELECTRIC FIELD $\mathbf{E}(t)$

In this section, we give the results obtained by calculating the effective fields

$$\mathbf{h}_\nu = -\frac{\partial \Phi}{\partial \mathbf{m}_\nu}, \quad \nu = 1, 2 \tag{13}$$

that act on the nuclear magnetic moments $\mathbf{m}_1$ and $\mathbf{m}_2$ when electric field $\mathbf{E}(t)$ alternating at a frequency of $\omega_{n\xi}$ [see (12), $\xi = c$] is switched on. It follows from (13) that the longitudinal components h_ν^z coincide with the equilibrium hyperfine field value H_n^z [see (11)], which determines the ω_n frequency [see (12)] and does not respond to field $\mathbf{E}(t)$. The transverse components h_ν^α ($\alpha = x, y$) include two terms, namely,

$$h_\nu^\alpha = h_{\nu l}^\alpha + h_{\nu L}^\alpha, \tag{14}$$

where

$$h_{\nu l}^\alpha = (-1)^{\nu - 1} d E^\alpha \tag{15}$$

is the field caused by the direct interaction between the magnetic moment $\mathbf{m}_\nu$ and field $\mathbf{E}(t)$, and

$$h_{\nu L}^\alpha = (-1)^{\nu - 1} \frac{1}{2} A L^\alpha, \tag{16}$$

which is the variable hyperfine field component caused by vector $\mathbf{L}$ oscillations under the action of $\mathbf{E}(t)$. As both $\omega_{n\xi}$ frequencies [see (12)] are much lower than the eigenfrequencies of vector $\mathbf{L}$ oscillations, the response of the $\mathbf{L}$ vector to field $\mathbf{E}(t)$ can be found by minimizing thermodynamic potential (7),

$$\frac{\partial \Phi}{\partial L^\alpha} = J L^\alpha - D E^\alpha(t) - \frac{1}{2} A l^\alpha = 0. \tag{17}$$

Substituting the solution to (17)

$$L^\alpha(t) = \frac{D}{J} E^\alpha(t) + \frac{A}{2J} l^\alpha(t), \tag{18}$$

into (14)–(16) and ignoring effects second-order in A yields

$$h_\nu^\alpha(t) = h_{\nu l}^\alpha + h_{\nu L}^\alpha = (-1)^{\nu - 1} \left(d + \frac{A}{2J} D \right) E^\alpha(t). \tag{19}$$

This equation describes two channels of NMER signal resonance excitation. Each of them is characterized by its own magnetoelectric constant. Currently, only experimental data on constant D are available. They were obtained from static magnetoelectric effect measurements and are summarized in [8]. These data were used in [4] to estimate the excitation level of nuclear spins under NMER conditions. For Cr_2O_3, this excitation can attain the same values as under usual NMR conditions in fields $\mathbf{E}(t)$ with an amplitude of 10^5 V/cm. In $TbPO_4$, the D constant is approximately 100 times larger than in Cr_2O_3. At such D values, the specified excitation level of nuclear spins should be attained at an amplitude of 10^3 V/cm. As far as the d constant is concerned [see (15)], estimating it requires the use of experimental NMER data.

4. ELECTRIC POLARIZABILITY CAUSED BY NUCLEAR SPIN OSCILLATIONS

As follows from (19), fields h_1^α and h_2^α differ only in sign ($h_1^\alpha = -h_2^\alpha$). As a result, there is a phase shift of π between the corresponding oscillations of the $\mathbf{m}_1$ and $\mathbf{m}_2$ vectors. The total magnetization is therefore zero,

$$\mathbf{m}(t) = \mathbf{m}_1(t) + \mathbf{m}_2(t) = 0.$$

For this reason, the usual NMR techniques are of no use for recording NMER signals. Nevertheless, we would be able to detect a NMER signal in heat loss measurements from the Q factor value. In the problem under consideration, the nuclear antiferromagnetism vector

$$\mathbf{l}(t) = \mathbf{m}_1(t) - \mathbf{m}_2(t) \tag{20}$$

becomes nonzero. As follows from (9), electric polarization $\mathbf{P}(t)$ oscillations are related to this vector. It is these oscillations that can play the role of NMER signals.

We calculated $\mathbf{P}(t)$ from the equation

$$P^\alpha(t) = -\frac{\partial \Phi}{\partial E^\alpha} = D L^\alpha(t) + d l^\alpha(t). \tag{21}$$

Substituting (18) for $L^\alpha(t)$ in (21) and retaining only the terms with $l^\alpha(t)$ that have resonance singularities at the NMER frequency yields

$$P^\alpha(t) = P_l^\alpha(t) + P_L^\alpha(t) = \left(d + \frac{A}{2J} D \right) l^\alpha(t). \tag{22}$$

Like (19), (22) determines two excitation channels of $\mathbf{P}(t)$ oscillations. The term with d describes the direct excitation of polarization $\mathbf{P}(t)$ by $\mathbf{l}(t)$ vector oscillations. The second term (with D) is responsible for indirect excitation via the hyperfine interaction between the $\mathbf{L}$ and $\mathbf{l}$ vectors. The $\mathbf{P}_l(t)/\mathbf{P}_L(t)$ ratio, as expected, coincides exactly with $\mathbf{h}_l(t)/\mathbf{h}_L(t)$,

$$\frac{\mathbf{P}_l(t)}{\mathbf{P}_L(t)} = \frac{\mathbf{h}_l(t)}{\mathbf{h}_L(t)} = 2\frac{dJ}{DA}. \tag{23}$$

Note that, when (21) and (22) are used to determine P_a^y, the replacements $L^y \longrightarrow -L^y$ and $l^y \longrightarrow -l^y$ should be made.

To finally determine the resonance response to field $\mathbf{E}(t)$ in the form of effective polarization (22), we must solve the equations for $\mathbf{m}_1$ and $\mathbf{m}_2$ (or $\mathbf{m}$ and $\mathbf{l}$) in constant [Eq. (11)] and alternating [Eqs. (14)–(16)] fields.

In the linear approximation under stationary excitation conditions, we obtain l^x and l^y in the form

$$l^x = 2\chi_n(\omega)\left(d + \frac{A}{2J}D\right)\left(E^x + i\frac{\omega}{\omega_n}E^y\right), \quad (24)$$

$$l^y = 2\chi_n(\omega)\left(d + \frac{A}{2J}D\right)\left(E^y - i\frac{\omega}{\omega_n}E^x\right), \quad (25)$$

where

$$\chi_n(\omega) = \chi_n(0)\frac{\omega_n^2}{\omega_n^2 - \omega^2} \quad (26)$$

is the NMR susceptibility. To determine l_a^y, both the $E^y \longrightarrow -E^y$ and $\omega_n \longrightarrow -\omega_n$ replacements should be made in (25); the latter is related to the negative sign of the hyperfine field H_{na}^z [see (11)].

Substituting (24) and (25) into (22) yields P_a^λ and P_c^λ in the form

$$P^x = 2\chi_n(\omega)\left(d + \frac{A}{2J}D\right)^2\left(E^x + i\frac{\omega}{\omega_n}E^y\right), \quad (27)$$

$$P^y = 2\chi_n(\omega)\left(d + \frac{A}{2J}D\right)^2\left(E^y - i\frac{\omega}{\omega_n}E^x\right). \quad (28)$$

As has been mentioned above, it makes little sense to quantitatively discuss these results in the absence of experimental NMER data, because we do not know the d values.

Equation (22) can also be applied to a pulsed nonlinear regime. For this purpose, we must substitute the l^α values found in the corresponding nonlinear approximation into it.

5. CONCLUSIONS

The new physical phenomenon, namely, the excitation of NMR by electric field $\mathbf{E}(t)$ (NMER), was predicted by one of these authors [9] for the examples of, first, a hypothetical two-sublattice ferromagnet with sublattices related to each other by a center of symmetry and, secondly, an antiferromagnet without a center of symmetry. More recently, this phenomenon was considered for several other antiferromagnets, including the four-sublattice antiferromagnets Cr_2O_3, α-Fe_2O_3 [4], and trirutiles Fe_2TeO_6 [10], which exist in reality. As follows from this work, the Mn_2Sb compound can be classified with two-sublattice ferromagnets from the point of view of NMER. This compound is also unusual because magnetoelectric effects should exist in it, but, as has been mentioned in the Introduction, they cannot be studied by the usual magnetostatic methods. One of the techniques for studying these effects [through exciting electrically active magnons by field $\mathbf{E}(t)$] was considered in [3]. This technique requires using frequencies on the order of 10^2–10^4 GHz. Studies of the properties of magnetic materials in fields $\mathbf{E}(t)$ at the lower boundary of this frequency range are already under way, although for other manganese compounds RMn_2O_5, where R = Eu, Er, and Gd [11]. A signal in field $\mathbf{E}(t)$ was detected close to the temperature of the structural phase transition, which is evidence of the important role played by magnetoelastic interactions in the formation of this signal.

The NMER technique can be used to study magnetoelectric effects at lower frequencies. For Mn_2Sb, these frequencies are 126.3 and 143.7 MHz [7]. We hope that the present publication will stimulate interest in experimental NMER studies.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 02-02-16440) and the Presidium of the Ural Division of the Russian Academy of Sciences.

REFERENCES

1. E. A. Turov, A. V. Kolchanov, V. V. Men'shenin, I. F. Mirsaev, and V. V. Nikolaev, *Symmetry and Physical Properties of Antiferromagnetics* (Fizmatlit, Moscow, 2001).
2. D. N. Astrov, Zh. Éksp. Teor. Fiz. **38**, 984 (1960) [Sov. Phys. JETP **11**, 708 (1960)].
3. I. F. Mirsaev and E. A. Turov, Zh. Éksp. Teor. Fiz. **121**, 419 (2002) [JETP **94**, 356 (2002)].
4. M. I. Kurkin, V. V. Leskovets, V. V. Nikolaev, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **45**, 653 (2003) [Phys. Solid State **45**, 685 (2003)].
5. E. A. Turov and M. P. Petrov, *Nuclear Magnetic Resonance in Ferro- and Antiferromagnets* (Nauka, Moscow, 1969; Wiley, New York, 1972).
6. M. I. Kurkin and E. A. Turov, *NMR in Magnetically Ordered Materials and Its Application* (Nauka, Moscow, 1990).
7. A. M. Portis and R. H. Lindquist, in *Magnetism*, Ed. by G. T. Rado and H. Suhl (Academic, New York, 1965), Vol. 2A, p. 357.
8. T. H. O'Dell and E. A. D. White, Philos. Mag. **22**, 649 (1970).
9. E. A. Turov, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 92 (2001) [JETP Lett. **73**, 87 (2001)].
10. I. F. Mirsaev and E. A. Turov, Zh. Éksp. Teor. Fiz. **124**, 338 (2003) [JETP **97**, 305 (2003)].
11. E. I. Golovenchits and V. A. Sanina, in *Abstracts of 33rd Meeting on Low-Temperature Physics* (Inst. Fiz. Metal., Ural. Otd. Ross. Akad. Nauk., Ekaterinburg, 2003), p. L200.

Translated by A. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 1006–1014.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1149–1158.

SOLIDS
Structure

A Multisublattice Magnetic Phase Induced by External Field in a Singlet Magnet

V. M. Kalita[a] and V. M. Loktev[b]

[a]*Institute of Physics, National Academy of Sciences of Ukraine, Kiev, 03028 Ukraine*
[b]*Institute for Theoretical Physics, National Academy of Sciences of Ukraine, Kiev, 03143 Ukraine*
e-mail: vloktev@bitp.kiev.ua

Received June 23, 2003

Abstract—The theory of magnetization in a longitudinal magnetic field is developed for an easy-plane multisublattice antiferromagnet with a singlet ground state and a strong single-ion anisotropy exceeding the magnitude of exchange interaction. The magnetic-field-induced phase transition from the singlet (magnetically disordered) state to a multisublattice antiferromagnetic state represents a displacive magnetic phase transition. At $T = 0$, this transition proceeds continuously and belongs to second-order phase transitions, while at $T \neq 0$, the behavior changes to jumplike and the process becomes the first-order phase transition. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The investigation of multisublattice (three-, four-sublattice, etc.) antiferromagnets possessing easy-plane anisotropy, despite considerable achievements in this direction of research, still receive much attention. Of special interest among these systems are hexagonal antiferromagnets of the ABX_3 type (where A is an alkali metal ion, B is a transition metal ion, and X is a halogen), in which the spins of magnetic B^{2+} ions form, on the one hand, antiferromagnetic (AFM) chains along the C_3 axis and, on the other hand, triangle structures in the basal plane (see reviews [1–3]). The phase transitions observed in such antiferromagnets exposed to a magnetic field are sometimes difficult to identify both with respect to the order (first versus second) and the type (order–disorder, etc.).

In these hexagonal antiferromagnets, the magnetic anisotropy always has a single-ion nature (owing to the orbital contribution) and may vary in extent. However, of most interest (especially from the standpoint of basic knowledge) is the case when the anisotropy is comparable to (or even stronger than) the Heisenberg exchange interactions [4]. Such a relation between the parameters of various spin interactions takes place, for example, in $CsFeBr_3$ and $CsFeCl_3$ antiferromagnets where (for Fe^{2+} ions with a pseudospin of $S = 1$) the constant of single-ion anisotropy is $D \approx 20$–30 K and the constant of exchange interaction is $J_{ch} \approx 3$–5 K for a pair of nearest ions (in the chains) belonging to adjacent planes and $J_{pl} \approx 0.3$–0.4 K for the same ions in the basal planes, and it is essential that both exchange interactions exhibit the AFM character [1, 5–7].

For such values of the energy parameters, whereby $D \gg J_{ch} + J_{pl}$, all ions in the magnet occur in the same singlet spin state and no multisublattice magnetic structure is formed in the crystal (in contrast to what would be required by the exchange interactions if there were no single-ion anisotropy) [1]. From the physical standpoint, this corresponds to a situation when the lowest state among three possible single-ion spin states with the spin projections $S_z = \pm 1, 0$ onto the C_3 axis is that with the zero projection. This ordered state (in fact, exhibiting no magnetic order) possesses the van Vleck character of susceptibility whose difference from that of usual paramagnets is especially pronounced for $T < D$.

When such an antiferromagnet is exposed to an external magnetic field **H** directed along the C_3 axis, the initial order of levels in the magnetic ions is gradually altered and, as the field strength increases, another state of the aforementioned triplet—that with nonzero spin projection—may become the lowest (ground). This magnetization proceeds in a self-consistent manner because of the interplay of different interactions: the exchange interaction and **H** favor the spontaneous polarization of ions, whereas the single-ion anisotropy, on the contrary, hinders this process [8]. Besides, the single-ion anisotropy tends to orient the ground-state magnetic moment of the ion perpendicularly to the hard axis C_3 and, hence, perpendicularly to **H**. In turn, the field forms the average spin projection onto **H**, thus producing canting of the average spin. As a result, the field $\mathbf{H} \parallel C_3$ will induce a phase transition from the singlet state to a phase (previously called the oblique phase [8]) with a complex magnetic order, in which ions are spin-polarized and the spins are canted toward the field.

As was noted above, the spin projection in the singlet state is zero (spin polarization in the ground state is absent), all ions are identical, and the AFM exchange is

not manifested. In the oblique phase, spontaneous polarization of the ground state during the AFM exchange is accompanied by the formation of a multisublattice structure with different spatial orientations. This field-induced transition from the van Vleck paramagnetism to antiferromagnetism can be (and has to be) classified as a displacive phase transition. The possibility of such magnetic phase transitions in the magnets with a large single-ion anisotropy was recently considered in [9, 10] (it should be noted that no any displacement of ions actually takes place and the term "displacive" only refers to the type of transition).

It can be suggested that the singlet state and a multisublattice structure, as well as a displacive magnetic phase transitions between these states are observed in $CsFeBr_3$ antiferromagnet. However, there are contradictory considerations [11–14] concerning the order of the phase transition between these states. Moreover, the aforementioned questions concerning magnetic polarization in the ground stage and its predominating role in determining the type of the phase transition were even not formulated. A description of the magnetization of this antiferromagnet [8] also bypassed this important question.

An analysis of the phase transitions between the singlet state and the multisublattice oblique phase at $T = 0$ showed that this process is continuous and belongs to second-order phase transitions [5, 8]. It should be emphasized that, in the longitudinal field (parallel to the C_3 axis) at $T = 0$, the singlet state has no magnetization at all and (see above) the system occurs in a state with the spin projection equal to zero. However, as soon as $T \neq 0$ (even when $D \gg T$), the situation is significantly complicated: by virtue of the paramagnetic process, the external field produces magnetization of the system, whereby the magnetic moments of ions in all sublattices are equal and oriented along the hard axis. The susceptibility of this state is much smaller as compared to that of the usual antiferromagnets because the population of levels in ions with nonzero spin projections at $D \gg T$ is exponentially small.

Despite a small value of the magnetization, the continuous character of transition from the field-magnetized singlet state to the oblique phase can no longer be retained because of different contributions to the origin of the new order (without and with self-consistency in the first and second case, respectively). If the continuity were retained, a procedure of spin polarization of the ground state would be absolutely different from that described above and (as will be shown below) it would imply that the magnetic moments arising with this polarization continuously turn out of the hard axis in the plane and then turn back to this axis in a stronger field. Obviously, such a "nonmonotonic" orientation is physically impermissible because the field oriented along the hard axis and not reducing the symmetry of the system in the course of magnetization only decreases the role of single-ion anisotropy. It is this circumstance that allows the AFM exchange to be manifested by forming a multisublattice AFM state oriented in the basal plane, the spins of which are rotated by the same field toward the hard axis. It should be emphasized that spontaneous deviation of the magnetic moments away from the hard axis would imply that, in the very beginning of spin flop, the exchange fields acting upon the field-induced average magnetizations (also small because of a low population of levels with nonzero spin projections) exceed both the anisotropy and the external field, which is impossible under real physical conditions.

Although the experiment of Tanaka *et al.* [5] does not allow an unambiguous conclusion to be made about the type and order of the phase transition in the ternary halide $CsFeBr_3$, we will proceed with consistent investigation of the model multisublattice spin system possessing a large single-ion anisotropy (exceeding exchange fields) in the longitudinal magnetic field. Below we will theoretically describe such a system assuming that its parameters and the relations between these characteristics correspond to those for $CsFeBr_3$.

2. MODEL HAMILTONIAN AND ITS EIGENSTATES

For simplicity, we restrict the consideration to bilinear isotropic exchange interactions, single-ion anisotropy, and Zeeman contribution. In this case, the simplest model Hamiltonian of a system with a structure analogous to that of $CsFeBr_3$ can be written in the following form:

$$H = \frac{1}{2}\sum_{\mathbf{n},\alpha,\mathbf{m},\beta} J_{\alpha\beta}\mathbf{S}_{\mathbf{n}\alpha}\cdot\mathbf{S}_{\mathbf{m}\beta} + D\sum_{\mathbf{n},\alpha}(S^Z_{\mathbf{n}\alpha})^2 - h\sum_{\mathbf{n},\alpha}S^Z_{\mathbf{n}\alpha}, \quad (1)$$

where the subscripts α and β indicate magnetic sublattices ($\alpha \neq \beta$), the total number of which in the system under consideration is six; **n** and **m** are the vectors setting spin positions in the sublattices; $D > 0$ is the constant of single-ion anisotropy (easy-plane structure); $h = \mu_\beta g H_z$ is the magnetic field in energy units; and H_z is the applied magnetic field (oriented along the crystallographic axis C_3 ($C_3 \parallel z$) and perpendicular to the easy axis. For the given field orientation the spins of various sublattices will be equally canted toward the field vector.

In $CsFeBr_3$ (and crystals of this family), the exchange interaction is spatially anisotropic (i.e., it depends on the mutual arrangement of spins in the lattice). The exchange constant J_{pl} in the basal plane differs from the value J_{ch} in the direction of hard axis (i.e., in the direction of chains). With allowance of the struc-

tural features, the exchange constant J_{ch} tends to establish the antiparallel orientation of spins in the neighboring planes, while J_{pl} orients spins closest to the easy axis at an angle of $2\pi/3$. As a result, the given hexagonal antiferromagnet with a finite value of average spin per lattice site acquires a multisublattice structure with the total number of magnetic sublattices equal to six.

An analysis of the possible quantum eigenstates of Hamiltonian (1) will be performed in the self-consistent field approximation, whereby the interspin fluctuations are ignored and the average of the product of spin operators of various sites is replaced by the product of average values for these sites. In this case, the ground state energy E_{gr} per cell (for spins belonging to different sublattices, three in one plane and three in the adjacent plane) is

$$E_{gr} = \frac{1}{2}\sum_{\alpha,\beta} J_{\alpha\beta} z_{\alpha\beta} \mathbf{s}_\alpha \cdot \mathbf{s}_\beta + D\sum_\alpha Q_\alpha - h\sum_\alpha s_\alpha^z, \quad (2)$$

where $\mathbf{s}_\alpha$ are the average vectors of ion spins in the ground state; $z_{\alpha\beta}$ is the number of nearest neighbors (three for spins in the same plane, two for spins in adjacent planes); we also introduce quantum-mechanical average values of the squares of z-projections of spin operators (called components of the quadrupole spin moment Q_α).

Let us introduce coordinate systems (ξ_α, η_α, ζ_α) for the spins of each sublattice, so that the average spin in the α sublattice is oriented along ζ_α, while ξ_α axis lies in the $z\zeta_\alpha$ plane. Then, the wave function of the ground spin state of the α sublattice has the following form [4]:

$$\psi_\alpha^{(0)} = \cos\phi_\alpha|1\rangle + \sin\phi_\alpha|-1\rangle, \quad (3)$$

where $|\pm1\rangle$, $|0\rangle$ are the eigenfunctions of the operator $S_{\mathbf{n}\alpha}^\zeta$. Using (3), we calculate the average values of the spin and the quadrupole moment components,

$$s_0 = \cos 2\phi, \quad Q_0^{\zeta\zeta} = 1, \quad Q_0^{\xi\xi} = \frac{1}{2}(1 + \sin 2\phi), \quad (4)$$

where subscript "0" indicates averaging over function (3). In expression (4), we omit the sublattice indexes because the reduced values are independent of α for the given field direction.

Using the values (4), the energy (2) can be represented as

$$E_{gr} = 9J_{pl}\cos^2 2\phi(3\cos^2\theta - 1) + 6J_{ch}\cos^2 2\phi(3\cos^2\theta - 1) + 6D\left[\cos^2\theta + \frac{\sin^2\theta}{2}(1 + \sin 2\phi)\right] - 6h\cos 2\phi\cos\theta, \quad (5)$$

where θ is the angle between spins of each sublattice and the magnetic field, equal to the angle between the crystallographic axis and ζ_α axes. Note that the ground state energy of a crystal was also determined and minimized in [5]. However, in contrast to (3), the selected wave functions were expressed in the general form in the crystallographic coordinate system (rather than in the intrinsic coordinate systems of sublattices) and contained the parameters of rotation of vectors $|\pm1\rangle$ and $|0\rangle$ in the Hilbert space, which significantly complicated the interpretation of a relationship between the observable and unknown (variational) parameters.

Our approach is more illustrative: spin configurations, determining the ground state in various possible phases and their mutual transformations in the magnetic field, are found from the equations derived by minimization of the energy (5) with respect to ϕ and θ:

$$\cos 2\phi\sin 2\phi[6J_{pl}(3\cos^2\theta - 1) + 4J_{ch}(2\cos^2\theta - 1)] - D\sin^2\theta\cos 2\phi - 2h\sin 2\phi\cos\theta = 0, \quad (6)$$

$$\cos\theta\sin\theta[(9J_{pl} + 4J_{ch})\cos^2 2\phi + D(1 - \sin 2\phi)] - h\cos 2\phi\sin\theta = 0. \quad (7)$$

These equations can be reduced to the equations for determining the ground state, which were obtained by Ostrovskii and Loktev [4] using a self-consistent procedure and special conditions of orthogonality for the vectors of various states. Leading eventually to the same solutions, Eqs. (6) and (7) are preferred because they additionally determine the conditions of phase stability.

Let us analyze the solutions of this system of equations. The first solution corresponds to a collinear ferromagnetic state and is realized under the conditions $\sin\theta = 0$ and $\sin 2\phi = 0$. In this case, the projections of spins of sublattices are aligned in the field and their values are maximum and equal $s_0 = S = 1$. The second solution corresponds to a three-sublattice AFM state with the Loktev structure [15] symmetric in the plane, which is valid for $\cos\theta = 0$ and $\sin 2\phi = -D/(6J_{pl} + 4J_{ch})$. This solution is possible only in the absence of magnetic field ($h = 0$). In the corresponding 120° spin structure, the values of spin projections are equal and smaller than the limit $S = 1$: $s_0 = \sqrt{1 - D^2/(6J_{pl} + 4J_{ch})^2}$.

The third solution refers to the oblique phase ($h \neq 0$), in which we also have $s_0(h) < 1$. Equation (7) directly indicates that, in contrast to the quasi-classical solution, the angle between the spin of sublattice and the direction of field **H** is a nonlinear function of the field strength:

$$\cos\theta = \frac{h\cos 2\phi}{(9J_{pl} + 4J_{ch})\cos^2 2\phi + D(1 - \sin 2\phi)}. \quad (8)$$

As the field strength grows, the spins are more canted toward the field direction and, hence, $s_0(h)$ increases so that always $s_0 \le s_0(h) \le 1$.

A field corresponding to the transition from the oblique phase to the state upon complete spin-flop, whereby all spins are parallel (oriented perpendicularly to the easy axis), can be determined by substituting $\cos\theta = 1$ and $\cos 2\phi = 1$ into formula (8) for the canting angle. As a result, we obtain the relation $h_{\text{flop}} = 9J_{\text{pl}} + 4J_{\text{ch}} + D$, which agrees with the expression for the flop field obtained in [5] and, moreover, coincides with the formula obtained in the quasi-classical approximation for $s_0 = 1$. As can be seen, the value of the spin-flop field is additive with respect to the anisotropy and exchange, although the physical mechanisms of their action are absolutely different.

Finally, one more (fourth) solution is possible as determined by the conditions $\cos\theta = 0$ and $\cos 2\phi = 0$. This solution can be realized both for nonzero magnetic fields ($h \neq 0$) and for $h = 0$. The reduction of spin projections in this state is maximum—up to $s_0 = 0$, in which case $\sin 2\phi = -1$. This corresponds to the van Vleck singlet (paramagnetic) state without magnetic order [16], which is also called a quadrupole spin state [17–19] because $Q^{\xi\xi} - Q^{\eta\eta} \neq 0$ (this difference has a limiting value).

It was demonstrated [9] that the phase transition from the singlet state to the oblique phase in a ferromagnet can be described using methods of the Landau theory of phase transitions. We suggest that the role of the order parameter can be performed by the spin projection of the ground level of an ion. Below we implement this approach taking into account a multisublattice character of the antiferromagnet under consideration.

3. PHASE TRANSITIONS FROM A SINGLET GROUND STATE TO THE OBLIQUE PHASE IN THE LONGITUDINAL MAGNETIC FIELD: $T = 0$

The phase transition from singlet to AFM state at $T = 0$ proceeds via magnetic polarization of ions in the ground state. This transition to a magnetically ordered state should be classified as a displacive phase transition, with the microscopic values of spin projections being equal to their macroscopic average values. In this case, the formation of a magnetic structure even at $T = 0$ can be described in terms of the phenomenological theory of phase transitions. Indeed, obtaining an expression for $\cos\theta$ from Eq. (7) and taking into account that $s_0 = \cos 2\phi$, the expression for the ground state energy at small values of s_0 ($s_0 \ll 1$) and h ($h \neq 0$) can be presented in the form of a series analogous to the Landau potential:

$$E_{gr} = \frac{3}{2D}[D^2 - D(6J_{\text{pl}} + 4J_{\text{ch}}) - h^2]s_0^2 + \frac{3}{8D}\left[D^2 - \frac{D - 2(9J_{\text{pl}} + 4J_{\text{ch}})}{D}h^2\right]s_0^4 + \frac{3}{16D} \times \left[D^2 - \frac{D^2 - 2(9J_{\text{pl}} + 4J_{\text{ch}})(9J_{\text{pl}} + 4J_{\text{ch}} - D)}{D^2}h^2\right]s_0^6. \quad (9)$$

As can be seen, expansion (9) does not contain the first powers of s_0 and h (second-order terms include their products). In addition, relation (9) indicates that (similar to the case of ferromagnets) the field oriented along the axis decreases the single-ion anisotropy. However, in contrast to the case of a ferromagnet (where this decrease is independent of the isotropic ferromagnetic exchange), the process in an antiferromagnet has a more complex character: the external field favoring the formation of a one-sublattice ferromagnetic phase simultaneously counteracts both the single-ion anisotropy and the AFM exchange.

The coefficient at s_0^2 in expansion (9) vanishes and then changes sign in the field $h = h_{QP}$ at which the system exhibits a transition from a singlet ground state to the oblique phase. This critical field is determined by the formula

$$h_{QP} = \sqrt{D^2 - D(6J_{\text{pl}} + 4J_{\text{ch}})}. \quad (10)$$

Note that, in the approximation of $s_0 \ll 1$, we have $h_{QP} \ll D$. This implies that, for $h = h_{QP}$, the coefficients at higher power of s_0 are positive and their values vary only slightly with the field h. Taking this into account, we infer from relation (9) that the magnetic-field-induced phase transition between the nonmagnetic (singlet) state with $s_0 = 0$ and a multisublattice spin-polarized state with $s_0 \neq 0$ is a second-order phase transition during which the magnetization varies in a continuous manner. Thus, restricting the expansion (9) for E_{gr} to the fourth-order terms, we conclude that ions in the ground state at $h < h_{QP}$ are not polarized (not "magnetized") and $s_0 = 0$. At the same time, for $h - h_{QP} > 0$, the average spin projection onto the quantization axis for each sublattice is

$$s_0(h) = 2\frac{h_{QP}^{1/2}}{D}(h - h_{QP})^{1/2}, \quad (11)$$

and the spin canting toward the hard axis in the magnetic field is described by the relation

$$\cos\theta = \frac{h_{QP}^{3/2}}{D^2}(h - h_{QP})^{1/2}. \quad (12)$$

Using relations (11) and (12), we can determine magnetization of the crystal along the hard axis as

$$m^z = 2\frac{h_{QP}^2}{D^3}(h - h_{QP}). \quad (13)$$

According to the above results, the longitudinal magnetization $\chi_\parallel = \partial m^z/\partial h$ in the singlet state is zero, while in the oblique phase it acquires a constant value

$$\chi_\parallel(T = 0, h > h_{QP}) = 2\frac{h_{QP}^2}{D^3}. \quad (14)$$

This is an unusual result, since both spin modulus and the canting angle in the course of spin flop are functions of the field.

The aforementioned oblique phase has essentially the structure of a multisublattice antiferromagnet, in which the spins of ions in various sublattices are canted toward the magnetic field. In usual antiferromagnets, the phase transition to such a canted multisublattice state is considered as a transition of the order–disorder type (such transitions are also called orientational phase transitions [20]). In contrast, the above phase transition from the singlet state without initial magnetic order to the oblique phase with a field-induced magnetic order, while also being a transition of the order–disorder type, represents a displacive magnetic phase transition.

4. THE FREE ENERGY OF A SINGLET GROUND STATE ANTIFERROMAGNET IN A LONGITUDINAL MAGNETIC FIELD

The description of a field-induced phase transition from the singlet ground state to AFM state at nonzero temperatures is complicated because of the need for taking into account the populations of all states of ions now occurring not only on the singlet ground level without magnetization, but on some other levels as well. At $T \neq 0$, the system features both the process of spin polarization considered above and the paramagnetic process.

For determining the ion states, previously we used a single-particle Hamiltonian [4]

$$H_0 = -h_{\text{exch}}^{(\alpha)}\mathbf{S}_{\mathbf{n}\alpha} + D(S_{\mathbf{n}\alpha}^z)^2 - hS_{\mathbf{n}\alpha}^z, \quad (15)$$

where $h_{\text{exch}}^{(\alpha)}$ is the average (exchange) field acting on the spin $\mathbf{S}_{\mathbf{n}\alpha}$ of the nth ion in the α sublattice. The eigenstates of Hamiltonian H_0 will be determined by introducing (similar to the case of $T = 0$) the intrinsic coordinate systems. At $T = 0$, the quantization axes were directed along the spin projections in the ground state of ions; now it is more convenient to orient these axes along the thermodynamic average spin vectors of each sublattice. It can be shown [4] that, for such coordinate systems (even making nonzero angles with the longitudinal field), the eigenstates of ions in antiferromagnets with easy-plane anisotropy are described by the wave functions

$$\psi_\alpha^{(0)} = \cos\phi_\alpha|1\rangle + \sin\phi_\alpha|{-1}\rangle, \quad \psi_\alpha^{(1)} = |0\rangle,$$
$$\psi_\alpha^{(2)} = -\sin\phi_\alpha|1\rangle + \cos\phi_\alpha|{-1}\rangle. \quad (16)$$

Using relations (16), we readily determine the partial spin projections onto the quantization axes in each state: $s_\alpha^{(0)} = -s_\alpha^{(2)} = \cos 2\phi_\alpha$, $s_\alpha^{(2)} = 0$; the corresponding partial averages of the operator $(S_\alpha^\zeta)^2$ are constant and equal to 1, 0, and 1, and the averages of the operator $(S_\alpha^\xi)^2$ are $(1/2)(1 + \sin 2\phi_\alpha)$, 1, and $(1/2)(1 - \sin 2\phi_\alpha)$.

Let us write down the thermodynamic averages for the magnetization modulus $m_\alpha = |\mathbf{m}_\alpha|$ and the operator $(S_\alpha^z)^2$. Note that, despite different orientations of the magnetization vectors $\mathbf{m}_\alpha$ in various sublattices, the values of these observables in various sublattices are the same:

$$m_\alpha = \Delta p_\alpha \cos 2\phi_\alpha, \quad (17)$$

$$\overline{(S_\alpha^z)^2} = p_\alpha \cos^2\theta_\alpha + \left(1 - \frac{p_\alpha}{2} + \frac{\Delta p_\alpha}{2}\sin 2\phi_\alpha\right)\sin^2\theta_\alpha, \quad (18)$$

where $\Delta p_\alpha = p_\alpha^{(2)} - p_\alpha^{(2)}$, $p_\alpha = p_\alpha^{(0)} + p_\alpha^{(2)}$, and $p_\alpha^{(0)}$ and $p_\alpha^{(2)}$ are the probabilities of the ground state and the second excited state (16), respectively.

According to the definition, the free energy $F = E - TS_{en}$, where E is the internal energy and S_{en} is the entropy. In the self- consistent field method, the entropy is configurational. For a multisublattice spin system, it is defined as sum $S_{en} = \sum_\alpha S_{en}^{(\alpha)}$ of the entropies of sublattices.

The internal energy of system (1) per particle can be presented in the following form:

$$E = \frac{1}{2}\sum_{\alpha,\beta} J_{\alpha\beta} z_{\alpha\beta} \mathbf{m}_\alpha \mathbf{m}_\beta + D\sum_\alpha \overline{(S_\alpha^z)^2} - h\sum_\alpha m_\alpha^z. \quad (19)$$

The entropy of the α sublattice is given by the standard expression:

$$S_{en}^{(\alpha)} = -\sum_{j = 0, 1, 2} p_\alpha^{(j)} \ln p_\alpha^{(j)}, \quad (20)$$

where $p_\alpha^{(j)}$ are the probabilities of single-ion states (16) satisfying the obvious condition $\sum_{j = 0, 1, 2} p_\alpha^{(j)} = 1$. In terms of relations (17) and (18), the entropy can be expressed as

$$S_{en}^{(\alpha)} = -\frac{p_\alpha + \Delta p_\alpha}{2}\ln\frac{p_\alpha + \Delta p_\alpha}{2} - \frac{p_\alpha - \Delta p_\alpha}{2}\ln\frac{p_\alpha - \Delta p_\alpha}{2} - (1 - p_\alpha)\ln(1 - p_\alpha). \quad (21)$$

Now we can obtain the final expression for the free energy of an antiferromagnet with $S = 1$ and an easy-

plane single-ion anisotropy in a longitudinal magnetic field:

$$\begin{aligned} F &= 9J_{pl}(\Delta p)^2(3\cos^2\theta - 1)\cos^2 2\phi \\ &+ 6J_{ch}(\Delta p)^2(2\cos^2\theta - 1)\cos^2 2\phi \\ &+ 6D\left(p\cos^2\theta + \left(1 - \frac{p}{2} + \frac{\Delta p}{2}\sin 2\phi\right)\sin^2\theta\right) \\ &- 6h\Delta p\cos 2\phi\cos\theta + 6T\left\{\frac{p+\Delta p}{2}\ln\frac{p+\Delta p}{2}\right. \\ &\left. + \frac{p-\Delta p}{2}\ln\frac{p-\Delta p}{2} + (1-p)\ln(1-p)\right\}, \end{aligned} \quad (22)$$

where the sublattice indexes (being equivalent) are omitted. This expression for the free energy is principally different from its traditional representation, including that used in [16, 17, 19]. Using Eq. (22), it is possible to follow the process of polarization of single-ion states (see below).

5. PHASE STATES OF A SINGLET GROUND STATE ANTIFERROMAGNET IN A LONGITUDINAL MAGNETIC FIELD AT $T \neq 0$

The equilibrium states of a singlet ground state antiferromagnet correspond to a minimum of the free energy (22). Here, the variation parameters are both ϕ, θ, and p, Δp. Differentiating Eq. (22) with respect to these variables and equating the derivatives to zero, we obtain the corresponding equations of state

$$\begin{aligned} \frac{\partial F}{\partial \phi} &= 6\{-\cos 2\phi\sin 2\phi[6J_{pl}(3\cos^2\theta - 1) \\ &+ 4J_{ch}(2\cos^2\theta - 1)](\Delta p)^2 \\ &+ D\Delta\sin^2\theta\cos 2\phi + 2h\Delta p\sin 2\phi\cos\theta\} = 0, \end{aligned} \quad (23)$$

$$\begin{aligned} \frac{\partial F}{\partial \theta} &= 6\left\{-\cos\theta\left[(9J_{pl} + 4J_{ch})\cos^2 2\phi(\Delta p)^2\right.\right. \\ &\left.\left. + 2D\left(\frac{3}{2}p - 1 - \frac{\Delta p}{2}\sin 2\phi\right)\right] + h\Delta p\cos 2\phi\right\}\sin\theta = 0, \end{aligned} \quad (24)$$

$$\begin{aligned} \frac{\partial F}{\partial \Delta p} &= 3\left\{2\Delta p[3J_{pl}(3\cos^2\theta - 1) + 2J_{ch}(2\cos^2\theta - 1)]\right. \\ &\left. + D\sin^2\theta\sin 2\phi - 2h\cos 2\phi\cos\theta + T\ln\frac{p+\Delta p}{p-\Delta p}\right\} = 0, \end{aligned} \quad (25)$$

$$\begin{aligned} \frac{\partial F}{\partial p} &= 6\left[D\left(\cos^2\theta - \frac{\sin^2\theta}{2}\right)\right. \\ &\left. + \frac{T}{2}\ln\frac{p^2 - (\Delta p)^2}{4(1-p)^2}\right] = 0. \end{aligned} \quad (26)$$

Let us consider the solutions of this system of equations corresponding to the singlet state of the crystal and the oblique phase. For the singlet solution, we have $\cos\theta = 1$ and $\cos 2\phi = 1$, while p and Δp are determined from the equations

$$2\Delta p(6J_{pl} + 2J_{ch}) - 2h + T\ln\frac{p+\Delta p}{p-\Delta p} = 0, \quad (27)$$

$$D + \frac{T}{2}\ln\frac{p^2 - (\Delta p)^2}{4(1-p)^2} = 0; \quad (28)$$

the quantization axis for this solution coincides with the hard axis and is directed along the field. The ground state of the ion is a singlet with the ground state function $\psi_\alpha^{(0)} = |0\rangle$ (see (16)). The two other (in this case, excited) states have limiting values of spins onto the quantization axis ± 1. Despite the fact that the spin projections in the ground state are zero, the thermodynamic average of the system magnetization is nonzero and equals $m = \Delta p$ (per magnetic ion). The nonzero magnetization appears only due to a difference in the populations of excited states (having limiting values of spin projections).

Differentiating Eqs. (27) and (28) with respect to h, we obtain expressions for the longitudinal magnetic susceptibility of the singlet state,

$$\chi_\parallel(T, h) = \frac{1}{6J_{pl} + 2J_{ch} + \dfrac{T}{p - (\Delta p)^2}}. \quad (29)$$

In the region of $D \ll T$, the temperature dependence of the magnetic susceptibility for $h \longrightarrow 0$ will be the same as that for usual antiferromagnets in the paramagnetic state. Indeed, substituting $\Delta p = 0$ and $p = 1/3$ into Eq. (29), we obtain

$$\chi_\parallel(T, 0) = \frac{1}{3(T - T_\theta)}, \quad (30)$$

where $T_\theta = -(2J_{pl} + (2/3)J_{ch})$ is the paramagnetic temperature.

A significant difference of the magnetic susceptibility in the singlet state of the crystal (essentially, in the van Vleck paramagnetic state) from the susceptibility of the paramagnetic phase of a usual antiferromagnet arise at low temperatures such that $D/T \gg 1$. Under this condition, the populations of excited states are much lower as compared to that of the nonmagnetic ground state, while a weak field is incapable of producing significant differences in the populations of two excited ion states, so that we can assume $\Delta p \longrightarrow 0$. Thus, the

temperature dependence of the magnetic susceptibility at $D/T \gg 1$ and $h \longrightarrow 0$ has the following form:

$$\chi_{\parallel}(T, 0) = \frac{2e^{-D/T}}{T}. \quad (31)$$

According to this, the longitudinal magnetic susceptibility of the crystal in the singlet ground state for $T \longrightarrow 0$ also tends to zero.

It is also interesting to study the field dependence of the magnetic susceptibility in the other limiting case, whereby $h \ll D$ and $h \gg T$. This implies that the temperature is sufficiently low, while the field is significant and exceeds the temperature (in the energy units adopted). In order to justify introduction of the latter condition, we note that the singlet state in the compound $CsFeBr_3$ is observed at temperatures from 1 to 3 K and the fields from 3 to 10 T. In this case, in contrast to that considered above, the magnetic field sharply changes the populations of excited levels. Owing to this sufficiently large magnetic field, we can realize the limiting situation in which the probability of the ion state with the spin projection along the field is much greater than the probability of a state with the opposite spin projection. As a result, for the equal populations of these levels produced by thermal excitations (magnetic disorder), the sufficiently large magnetic field ($h \gg T$) produces actually the ideal order for these two levels. Under these conditions, the magnetic susceptibility of the singlet state is

$$\chi_{\parallel}(T, h) = \frac{1}{2}\chi_{\parallel}(T, 0)e^{h/T}. \quad (32)$$

Apparently, this nonlinear field dependence described by the exponential law is observed in $CsFeBr_3$ in large magnetic fields [5].

The second solution of Eqs. (23)–(26) corresponding to the oblique phase exists in the fields $h > h_{QP}$. Let us analyze this solution assuming that the average magnetization of sublattices is $m \ll 1$, that is, in the limit of $h \longrightarrow h_{QP}$. This means that the exchange field in Hamiltonian (15) is small as compared to both h and D. The polarization in the ground state of the ion also depends on the temperature T but, in the case under consideration, the magnitude of polarization is much lower than the limiting value of unity. Under these conditions, we may calculate the populations of levels in the oblique phase with neglect of the nonlinearity caused by the exchange interaction. Then, using Eqs. (25) and (26), we obtain the following expressions for the populations determined entirely by the D/T ratio:

$$p = \frac{2 + \Delta p}{3}, \quad \Delta p = \frac{1 - e^{-D/T}}{1 + 2e^{-D/T}}. \quad (33)$$

The magnitude of polarization $s = \cos 2\phi$ in the ground state of ions in the oblique phase and the orientation of quantization axes are determined from Eqs. (23) and (24). Taking into account formula (33), these equations can be written as

$$\cos 2\phi \sin 2\phi [6J_{pl}(3\cos^2\theta - 1) + 4J_{ch}(2\cos^2\theta - 1)](\Delta p(T))^2 - D\Delta p(T)\sin^2\theta \cos - 2\phi 2h\Delta p(T)\sin 2\phi \cos\theta = 0, \quad (34)$$

$$\cos\theta \sin\theta [(9J_{pl} + 4J_{ch})(\Delta p(T))^2 \cos^2 2\phi + D\Delta p(T)(1 - \sin 2\phi)] - h\Delta p(T)\cos 2\phi \sin\theta = 0. \quad (35)$$

As can be seen, Eqs. (34) and (35) are fully equivalent to Eqs. (6) and (7) but contain the temperature-dependent coefficients. These quantities are conveniently designated as $J_{pl}(T) = J_{pl}(\Delta p(T))^2$, $J_{ch}(T) = J_{ch}(\Delta p(T))^2$, and $D(T) = D\Delta p(T)$.

The part of the free energy dependent on the magnetic field and describing the polarization can be written as

$$\Delta F = \frac{3}{2}\left[D(T) - (6J_{pl}(T) + 4J_{ch}(T)) - \frac{h^2\Delta p(T)}{D}\right]s^2 + \frac{3}{8}D(T)s^4. \quad (36)$$

Equating to zero the coefficient at s^2 in this relation, we determine the magnitude of the critical field:

$$h_{QP}(T) = D\sqrt{\left(1 - \frac{6J_{pl} + 4J_{ch}}{D}\right)\Delta p(T)}. \quad (37)$$

Minimizing the functional (36), we determine the magnitude of polarization in the ground state:

$$s(T, h) = 2\frac{\sqrt{h_{QP}(T)}\sqrt{h - h_{QP}(T)}}{D}. \quad (38)$$

Finally, using Eq. (35) we determine the spin canting angle in the sublattices for $h \longrightarrow 0$:

$$\cos\theta = \frac{h_{QP}^{3/2}(T)}{D^2}\sqrt{h - h_{QP}(T)}. \quad (39)$$

Using the above expressions, we determine the crystal magnetization for the given solution as the sum of magnetization vectors of all sublattices. The magnetization is directed along the hard axis and has the value (per magnetic ion)

$$m^z = 2\frac{h_{QP}^2(T)[h - h_{QP}(T)]}{D^3}\Delta p(T). \quad (40)$$

According to expression (40), the magnetic susceptibility of the system corresponding to the solution for the

oblique phase is independent of the field:

$$\chi_{QP} = \chi_{\parallel}(T, h > h_{QP}) = 2\frac{h_{QP}^2(T)}{D^3}\Delta p(T). \quad (41)$$

This behavior of the magnetic susceptibility was experimentally observed by Haseda *et al.* [21] during measurements of longitudinal (along the C_3 axis) magnetization of $CsFeCl_3$ and $RbFeCl_3$ at various temperatures. In the interval of magnetic fields 4–6 T $\le H \le$ 10–11 T at $T \le 2.5$ K, the susceptibility of these antiferromagnets has proved to be constant, while outside this interval it exhibited a sharp drop—exactly as expected for both small and large field outside the region of existence of the oblique phase.

It should be noted that, despite the fact that the susceptibility in the singlet state is exponentially small (see relations (31) and (32), it can still be on the same order of magnitude as the susceptibility (41) of the oblique phase because $h_{QP}/D \ll 1$ and $D \gg T$.

6. PHASE TRANSITIONS IN A LONGITUDINAL MAGNETIC FIELD AT $T \neq 0$.

The above analysis showed that the system of equations (23)–(26) has two solutions at $T \neq 0$ for the fields $h \ll D$. The first solution corresponds to a one-sublattice crystal phase with a nonmagnetic singlet ground state and the two other states having limiting values of the spin projections onto the quantization axis, the latter axis being directed along the field.

In the second solution, the ion spin projections are not equal to their limiting values and the ground state is polarized. This corresponds to a multisublattice AFM state; for $h \longrightarrow h_{QP}$, the quantization axes are oriented perpendicularly to the field. Obviously, the passage from the first solution to the second involves a jumplike rotation of the quantization axes. In other words, the transition from the singlet state to the oblique phase at $T \neq 0$ is the first-order phase transition, in contrast to the second-order phase transition at $T = 0$.

The magnetic field for this phase transition can be determined from the condition of equal free energies for the two solutions. For $h \ll D$, the free energy of the singlet phase can be written as

$$F_{SP} = -T\ln(1 + 2e^{-D/T}\cosh h/T). \quad (42)$$

If the field is such that $h \gg T$ but smaller than the constant of anisotropy, relation (42) can be rewritten as

$$F_{SP} = -T^2\chi_{\parallel}(T, h). \quad (43)$$

The free energy for the second solution equals the sum of a single-ion component, which is dependent on D/T but independent of h, and the contribution (36) describing the magnetic-field-induced polarization of ions in the ground state. Thus, the expression for the equilibrium free energy dependent only on h for the oblique phase at $s \ll 1$ is as follows:

$$F_{QP} = -T\ln(1 + 2e^{-D/T}) + \Delta F(h). \quad (44)$$

Taking into account expressions (37)–(41) and assuming $T/D \ll 1$, the expression for F_{QP} can be transformed to

$$F_{QP} = -T^2\chi_{\parallel}(T, 0) - \frac{1}{2}\chi_{QP}(T)[h - h_{QP}(T)]^2. \quad (45)$$

For determining the critical field for the phase transition from the singlet state to the oblique phase, we equate the expressions for free energies of these phase and arrive at the transcendental equation

$$T^2[\chi_{\parallel}(T, h_C) - \chi_{\parallel}(T, 0)] = \frac{1}{2}\chi_{QP}(T)[h_C - h_{QP}(T)]^2, \quad (46)$$

where h_C is the critical field for the phase transition.

As can be seen from Eq, (46), the value of h_C is somewhat greater than h_{QP}. This result confirms the above conclusion that the phase transition from the singlet phase to the oblique phase at $T \neq 0$ is the first- order transition and a jump in magnetization has to take place at $h = h_C$. It is interesting to note that, as the temperature T decreases, the values of h_C and h_{QP} approach each other and $h_C \longrightarrow h_{QP}$ when $T \longrightarrow 0$.

7. CONCLUSIONS

It was demonstrated that the magnetic phase transitions in singlet ground state magnets can be described using an approach based on the Landau theory of phase transitions. The role of the order parameter is performed by the spin polarization of single-particle states of paramagnetic ions. We have shown that the ternary magnetic halogenides ABX_3 with easy-plane single-ion magnetic anisotropy feature a phase transition from the singlet (one- sublattice) state to an AFM phase induced by the longitudinal magnetic field. This transition represents a displacive magnetic phase transition. An important special feature of the system under consideration is that the phase transition is continuous at $T = 0$ and exhibits a jumplike character at $T \neq 0$. This change in the phase transition character from the second to first order is related to the magnetic-field-induced paramagnetic process, which cannot be ignored in the system studied.

Another, also very important peculiarity of the system under consideration is manifested in the course of magnetization at low temperatures. First, the longitudinal component of the magnetic susceptibility in the singlet phase (i.e., in the initial stage of magnetization) is strongly nonlinear, exhibiting exponential growth with increasing external field h and decreasing with the temperature T. Second, this magnetic susceptibility component ceases to depend on h upon transition to the AFM

state at $h > h_C$. This behavior is rather unusual. Indeed, the magnetic susceptibility of usual antiferromagnets (obeying the quasi-classical description) in the spin-flop phase is constant. However, the spin-flop phase in such a system is only the result of a change in the AFM order caused by the external field, while the moduli of spins of sublattices in the spin-flop phase remain essentially constant in the course of canting toward the field. In a singlet ground state magnet at $h < h_C$, the AFM state (including multisublattice) is absent. It is only at $h = h_C$ that sublattices begin to form and acquire nonzero magnetic moments increasing with h, whereby the field also induces canting of the spins of sublattices toward the hard axis. It is unusual that the field dependence of the average spins and canting angles are such that the magnetic susceptibility of the whole crystal in the oblique phase is analogous to that for the quasi-classical antiferromagnets.

The results of our investigation qualitatively well agree with the data of recent experimental investigation of the static magnetic properties of $CsFeBr_3$ [5]. This antiferromagnet exhibits a clear displacive magnetic phase transition from the singlet state to the oblique phase. From additional experimental data on the heat capacity of $CsFeBr_3$, this phase transition at $T \neq 0$ is of the first order, although the crystal magnetization proceeds smoothly and exhibits no jump. As the temperature is decreased to approach $T = 0$, $CsFeBr_3$ exhibits a clear tendency to change the transition order from first to second [5], but this behavior requires separate experimental investigation.

The field dependence of the longitudinal magnetization of $CsFeBr_3$ is well consistent with the theory proposed above. The magnetic susceptibility of $CsFeBr_3$ exhibits a nonlinear (exponential) dependence on the applied field strength in the singlet state and remains constant at $h > h_C$, after the transition to the oblique phase. However, according to the available data [5], it is still difficult to provide for a direct description of the magnetization and susceptibility curves for a $CsFeBr_3$ crystal. It should also be noted that continuous variation of the magnetization of $CsFeBr_3$ during the first-order phase transition is not excluded, if we take into account the possibility of formation of an intermediate state simultaneously containing both phases [22]. However, consideration of the possible domain formation goes outside the scope of this study and is a subject for separate investigation.

ACKNOWLEDGMENTS

The authors are grateful to S.M. Ryabchenko for fruitful discussion and useful critical remarks.

This study was supported in part by the National Academy of Sciences of Ukraine, grant nos. 1.4.1.VTs/95 and 0102U002332.

REFERENCES

1. R. S. Gekht, Usp. Fiz. Nauk **159**, 261 (1989) [Sov. Phys. Usp. **32**, 871 (1989)].
2. M. F. Collins and O. A. Petrenko, Can. J. Phys. **75**, 605 (1997).
3. B. S. Dumesh, Usp. Fiz. Nauk **170**, 403 (2000) [Phys. Usp. **43**, 365 (2000)].
4. V. M. Loktev and V. S. Ostrovskiĭ, Fiz. Nizk. Temp. **20**, 983 (1994) [Low Temp. Phys. **20**, 775 (1994)].
5. Y. Tanaka, H. Tanaka, and T. Ono, cond-mat/0104287.
6. B. Dorner, D. Visser, U. Stiegenberger, *et al.*, Z. Phys. B **72**, 487 (1988).
7. A. Harrison and D. Visser, J. Phys.: Condens. Matter **4**, 6977 (1992).
8. V. M. Kalita, I. M. Ivanova, and V. M. Loktev, Fiz. Nizk. Temp. **28**, 667 (2002) [Low Temp. Phys. **28**, 475 (2002)].
9. V. M. Kalita and V. M. Loktev, Fiz. Nizk. Temp. **28**, 1244 (2002) [Low Temp. Phys. **28**, 883 (2002)].
10. V. M. Kalita and V. M. Loktev, Fiz. Tverd. Tela (St. Petersburg) **45**, 1450 (2003) [Phys. Solid State **45**, 1523 (2003)].
11. H. Kawamura, J. Phys. Soc. Jpn. **61**, 1299 (1992).
12. H. Kawamura, J. Phys.: Condens. Matter **10**, 4707 (1998).
13. H. Kadowaki, S. M. Shapiro, T. Inami, and Y. Ajiro, J. Phys. Soc. Jpn. **57**, 2640 (1988).
14. Y. Ajiro, T. Nakashima, Y. Uno, *et al.*, J. Phys. Soc. Jpn. **57**, 2648 (1988).
15. V. M. Loktev, Fiz. Nizk. Temp. **5**, 295 (1979) [Low Temp. Phys. **5**, 142 (1979)].
16. A. K. Zvezdin, V. M. Matveev, A. A. Mukhin, and A. I. Popov, *Rare-Earth Ions in Magnetic-Ordered Crystals* (Nauka, Moscow, 1985).
17. F. P. Onufrieva, Zh. Éksp. Teor. Fiz. **89**, 2270 (1985) [Sov. Phys. JETP **62**, 1311 (1985)].
18. Yu. N. Mitsaĭ, A. N. Maĭorova, and Yu. A. Fridman, Fiz. Tverd. Tela (St. Petersburg) **34**, 66 (1992) [Sov. Phys. Solid State **34**, 34 (1992)].
19. V. V. Val'kov and S. G. Ovchinnikov, *Quasi-Particles in Strongly Correlated Systems* (Sib. Otd. Ross. Akad. Nauk, Novosibirsk, 2001).
20. K. P. Belov, A. K. Zvezdin, A. M. Kadomtseva, and R. Z. Levitin, *Reorientational Transitions in Rare-Earth Magnets* (Nauka, Moscow, 1979).
21. T. Haseda, N. Wada, M. Hata, and K. Amaya, Physica B (Amsterdam) **108**, 841 (1991).
22. V. G. Bar'yakhtar, A. N. Bogdanov, and D. A. Yablonskiĭ, Usp. Fiz. Nauk **156**, 47 (1988) [Sov. Phys. Usp. **31**, 810 (1988)].

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 1015–1026.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1159–1171.

SOLIDS
Structure

New Micromagnetic States of Magnetically Soft Nanoparticles with a Nearly Cubic Shape

B. A. Ivanov and E. V. Tartakovskaya

Institute of Magnetism, National Academy of Sciences of Ukraine, Kiev, 03142 Ukraine
e-mail: bivanov@i.com.ua

Received November 12, 2003

Abstract—The ground state of nonellipsoidal particles can be inhomogeneous due to the effect of a demagnetizing field. The approach proposed here for studying such particles is based on the combination of symmetry analysis and perturbation theory. The general formulation of this approach, which makes it possible to analyze weakly inhomogeneous states for particles with a complex shape, is considered. The ground state of cubic particles of magnetically soft materials is calculated analytically, and the effect of small strains of cubic particles on the magnetization distribution in the particles is investigated. It is shown for the example of magnetically soft cubic particles that even a small deviation of the particle shape from symmetrical may result in the realization of a special magnetic state in such particles, in which the symmetry in the magnetization distribution is lower than the particle symmetry. A change in the parameters of a particle can substantially modify its magnetic properties and may even induce a phase transition to a state with a different symmetry. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Advanced in sputtering technologies and electron-beam lithography during the last decade have made it possible to prepare artificial magnetic particles (magnetic dots) of a nanometer size as well as their ordered system (superlattices [1, 2]. Magnetic dots may have various shapes and are mainly prepared from magnetically soft ferromagnets (e.g., Co, Fe, or permalloy) deposited on a nonmagnetic substrate. The superlattices of magnetic dots are important for practical applications, in particular, for designing new devices for high-density magnetic recording [3, 4], magnetic field sensors [5], and logical elements of computers [6].

Superlattices of magnetic dots are also interesting as a basically new object of the fundamental physics of magnetism. Indeed, from the standpoint of traditional physics of magnetism, magnetic dots with a size on the order of tens or hundreds of nanometers (which is smaller than or comparable to the one-domain size) are typical monodomain particles, which have been studied for more than 50 years. However, detailed analysis of these particles in recent years demonstrated a number of peculiarities that have not been discussed before. The main one is that various nonuniform distributions of magnetization can be observed in such particles. If the size of a particle becomes larger than a certain critical value, but is still smaller than the one-domain size, a vortex state can be realized in the particle [7]. For smaller nonellipsoidal particles, the ground state a specific weakly inhomogeneous magnetic configuration determined by the magnetic dipole interaction (see [8–12]). In [8–10], a method was proposed for determining the micromamagnetic pattern of such a distribution from the known average direction of the magnetic moment. Such a simplification is possible for magnetically rigid particles, for which the average direction of magnetization is determined by magnetic anisotropy, and for an elongated particle (e.g., having the shape of a cylinder or a parallelepiped), in which the mean magnetic moment direction is determined by a uniform demagnetizing field (anisotropy of shape). However, magnetically soft particles of a symmetric shape, viz., flat squares (with a thickness smaller than the square side) or cubic particles, were found to be more convenient for application and preparation. Anisotropy in such particles is negligibly small, while a uniform demagnetizing field possesses a high symmetry and does not determine unambiguously the average direction of the magnetic moment. In this case, both the average direction of magnetization and nonuniform deviations from it are determined by a nonuniform demagnetizing field and the problem must be solved self-consistently. As a result, the ground state of such quasi-monodomain particles turned out to be complicated and was treated in a series of publications [13–16] devoted to only quadratic permalloy and supermalloy particles. The authors of these publications made use of the fruitful idea that the average direction of the magnetic moment in symmetric particles is determined by the weak nonuniform component of the demagnetizing field. It was shown experimentally and with the help of numerical simulation for square particles that, depending to the ratio of the square side to the particle thickness, the particle can be in one of two micromagnetic states, viz., the flower state, in which the average direction of the magnetic moment is parallel to the side of the square, and the leaf

state, in which the average direction of the magnetic moment is parallel to the diagonal of the square. The authors of [13–16] also carried out numerical variational analysis of the corresponding micromagnetic states choosing some simple test functions. However, the approach developed by these authors did not permit a generalization to the case of particles with a more complex shape (e.g., cubic or rectangular). The calculations were performed without taking into account the natural boundary conditions for magnetization (see [17, 18]). For this reason, the ground state of a small particle has not been determined unambiguously even for particles of a simple shape. The important (in our opinion) question concerning the variation of the ground state upon small but experimentally feasible deviations of particles from a precisely preset shape (i.e., a weak lowering of symmetry of the uniform demagnetizing field) has not been discussed either.

The construction of an analytic theory for micromagnetic states began in [10], where the micromagnetic structure was calculated on the basis of perturbation theory in small deviations relative to the known average direction of the magnetic moment in a thin cylinder. The small parameter of the theory was the squared ratio of the particle size to the exchange length of the material, which corresponds to the ratio of the magnetic dipole energy to the exchange energy. Consequently, the criterion of smallness for this parameter corresponds to the fulfillment of the standard criterion of one-domain nature of a nanoparticle. In our previous publications [19, 20], we generalized this theory and applied it for calculating the average direction of the magnetic moment and corresponding micromagnetic states in flat symmetric particles. In this theory, the calculations of uniform and nonuniform demagnetizing fields are treated, respectively, as the zeroth and first approximations in perturbation theory. For square particles, the results of our calculations are in qualitative agreement with the experimental data [14–16]. For flat rectangular particles, we indicate the existence of a new micromagnetic state, which was called the intermediate state. The direction of the magnetic moment $\langle \mathbf{M}_0 \rangle \equiv \mathbf{M}_0$ averaged over the particle volume is not associated with any symmetry axes of the particle for such a state. In the range of the intermediate state, the average moment forms an angle with the longer side of the rectangle, the magnitude of this angle depending on the ratio of the uniform and nonuniform components of the demagnetizing field, while the micromagnetic distribution contains terms with symmetries corresponding to the flower state as well as the leaf state. The results of our calculations were completely confirmed by numerical simulation [20].

This study is devoted to analytic calculation of the ground state of particles with a higher symmetry as compared to flat particles, i.e., particles of cubic shape. The approach used here is based on the combination of the symmetry analysis and calculations in perturbation theory. We also studied the effect of small deformations of cubic particles on the magnetization distribution in these particles. In a square particle, a uniform demagnetizing field "lays" the magnetic moment into the plane of the square, and calculation in the next order in perturbation theory is required only for determining the direction of the magnetic moment in this plane, while in a cubic particle, the direction of magnetic moment $\mathbf{M}_0$ in space is determined only by the nonuniform component of the demagnetizing field. Numerical simulation carried out in [8] indicates that a flower-type state is realized in a cubic particle; i.e., magnetic moment vector $\mathbf{M}_0$ is parallel to one of the edges of the cube.

In Section 2, we give a general formulation of the approach that makes it possible to analyze weakly inhomogeneous states for particles with a complex shape. The expansion of the system energy into a power series in components of vector $\mathbf{M}_0$, which naturally arises when perturbation theory is used, is verified by the symmetry analysis of the effective energy as a function of $\mathbf{M}_0$. In Section 3, the case of cubic particles is briefly considered. Section 4 is devoted to the determination of the magnetic state, in particular, the average direction of the magnetic moment in particles of a nonrectangular shape close to the cubic shape. It is shown that such particles can exhibit a special form of the intermediate state, in which the symmetry of magnetization distribution is lower than the particle symmetry. It should be noted that the application of the standard method of computer simulation presuming the discretization of the problem by dividing the volume of a particle into small domains with a shape repeating the particle shape [21] is quite difficult in the case of nonrectangular particles. For this reason, the calculation based on the proposed approach is an important tool for their theoretical investigation. The concluding part of this paper contains a discussion of the results and analysis of possible generalizations of the algorithm developed here (e.g., its applicability for describing the magnetization reversal of particles in a weakly inhomogeneous state).

2. FORMULATION OF THE PROBLEM AND GENERAL RELATIONS

The energy functional for a particle in an isotropic ferromagnet taking into account the exchange and magnetic dipole interactions can be written in the form

$$W = \int_V d\mathbf{r}\left\{\frac{l_e^2}{2}[(\nabla \mathbf{M})^2] - \frac{1}{2}\mathbf{M}\cdot\mathbf{H}_m\right\}, \qquad (1)$$

where the demagnetizing field is defined in the usual way as

$$\mathbf{H}_m = -\frac{\partial}{\partial \mathbf{r}}\int d\mathbf{r}'\left[\mathbf{M}(\mathbf{r}')\frac{\partial}{\partial \mathbf{r}'}\right]\frac{1}{|\mathbf{r}-\mathbf{r}'|},$$

$l_e = \sqrt{A/M_0^2}$ is the exchange length, M_0 is the saturation magnetization, A is the nonuniform exchange constant,

and integration is carried out over the particle volume. The magnetization distribution is described by the equation

$$(l_e^2\nabla^2\mathbf{M} + \mathbf{H}_m)\times\mathbf{M} = 0, \quad (2)$$

with the boundary conditions [17, 18]

$$(\partial\mathbf{M}/\partial x_i)n_i = 0, \quad (3)$$

where n is the vector of the normal to the particle surface. If the characteristic size l of the particle is smaller than the exchange length, we have

$$l_e^2|\nabla\mathbf{M}|^2 \propto (l_e/l)^2 M_0^2 \gg MH_m \propto 4\pi M_0^2$$

and the energy of the magnetic dipole interaction is smaller than the exchange energy. For $(l/l_e) \longrightarrow 0$, the magnetization distribution becomes uniform and we can assume that the nonuniform magnetization part $\delta\mathbf{M}$ is small ($|\delta\mathbf{M}| \ll M_0$) for $(l/l_e) \ll 1$. In this case, we can construct a recurrent procedure stemming from the fact that the magnetization is uniform in the zeroth approximation in (l/l_e). In this case, in the first approximation in $(l/l_e)^2$, the demagnetizing field $\mathbf{H}_m = \mathbf{H}_m^{(0)}$ is calculated for uniform magnetization $\mathbf{M}_0$. Field $\mathbf{H}_m^{(0)}$ can be written in terms of the tensor of effective demagnetizing factors $N_{ik}(\mathbf{r})$, $H_{m,i}^{(0)} = N_{ik}(\mathbf{r})M_{0,k}$, and magnetic field $\mathbf{H}_m^{(0)}$ is a function of coordinates only by virtue of the coordinate dependence of $N_{ik}(\mathbf{r})$.

The presence of a nonuniform field induces small nonuniform deviations of $\mathbf{M}$ from $\mathbf{M}_0$, $\mathbf{M} = \mathbf{M}_0 + \delta\mathbf{M}$, with an amplitude proportional to $(l/l_e)^2$. It is these deviations and the demagnetizing field associated with these deviations that determine the type of the ground state of a small nonellipsoidal particle. We will describe the direction of $\mathbf{M}_0$ by angular variables, which can be conveniently chosen in the form

$$\mathbf{M}_0 = M_0[\mathbf{e}_z\sin\theta + \cos\theta(\mathbf{e}_x\cos\varphi + \mathbf{e}_y\sin\varphi)].$$

To calculate $\delta\mathbf{M}$, it is convenient to pass to a system of coordinates in which the unit vector $\mathbf{e}_3$ is directed along the average magnetization vector, $M_0\mathbf{e}_3 = \mathbf{M}_0$, the remaining unit vectors being defined by the formulas $M_0\mathbf{e}_1 = \partial\mathbf{M}_0/\partial\theta$ and $M_0\mathbf{e}_2 = (\partial\mathbf{M}_0/\partial\varphi)/\sin\theta$. In this case, vector equation (2) in the linear approximation in $\delta\mathbf{M} = m_1\mathbf{e}_1 + m_2\mathbf{e}_2$ can easily be written in the form of two independent equations

$$\begin{aligned} l_e^2\nabla^2 m_1 + \frac{\partial}{\partial\theta}(M_{0,i}N_{ik}M_{0,k}) &= 0, \\ l_e^2\nabla^2 m_2 + \frac{\partial}{\sin\theta\partial\varphi}(M_{0,i}N_{ik}M_{0,k}) &= 0. \end{aligned} \quad (4)$$

In this approximation, the particle energy, taking into account Eqs. (4), can be written in the form of the functional quadratic in m_1 and m_2,

$$W^{(2)} = -\int_V d\mathbf{r}\frac{l_e^2}{2}\{(\nabla m_1)^2 + (\nabla m_2)^2\}. \quad (5)$$

To solve Eqs. (4) for the given shape of particles, we write the components of tensor $N_{ik}(\mathbf{r})$ in the form of the expansion in the complete orthonormal set of functions $f_\alpha(\mathbf{r})$, which are solutions to the D'Alembert equation and satisfy boundary conditions (3) on the particle surface:

$$\begin{aligned} &\nabla^2 f_\alpha(\mathbf{r}) + l_e^2\lambda_\alpha f_\alpha(\mathbf{r}) = 0, \\ &\int_V d\mathbf{r}\, f_\alpha(\mathbf{r}) f_\beta(\mathbf{r}) = \delta_{\alpha\beta}, \quad \frac{\partial f_\alpha(\mathbf{r})}{\partial r_i}n_i = 0. \end{aligned} \quad (6)$$

Writing

$$N_{ik}(\mathbf{r}) = \sum_\alpha \nu_{ik}^{(\alpha)} f_\alpha(\mathbf{r}),$$

we can easily obtain explicit expressions for m_1 and m_2 in terms of coefficients $\nu_{ik}^{(\alpha)}$, e.g.,

$$m_1 = \frac{1}{2}\sum_\alpha f_\alpha(\mathbf{r})\frac{l^2}{l_e^2\lambda_\alpha}\frac{\partial}{\partial\theta}(M_{0,i}\nu_{ik}^{(\alpha)}M_{0,k}). \quad (7)$$

Further, we can write the expression for energy $W^{(2)}$ in the universal form

$$W^{(2)} = -\frac{1}{2}\left(\frac{l}{l_e}\right)^2\Lambda_{ik,lm}(u_{ik}u_{lm} + v_{ik}v_{lm}), \quad (8)$$

where tensor $\Lambda_{ik,lm} = \Lambda_{ki,lm} = \Lambda_{lm,ik}$ is defined only by the particle shape,

$$\Lambda_{ik,lm} = \sum_\alpha \frac{1}{\lambda_\alpha}\nu_{ik}^{(\alpha)}\nu_{lm}^{(\alpha)}, \quad (9)$$

while the symmetric rank-two tensors u_{ik} and v_{ik} are defined as

$$\begin{aligned} u_{ik} &= \frac{1}{2}\frac{\partial}{\partial\theta}(M_{0,i}M_{0,k}), \\ v_{ik} &= \frac{1}{2\sin\theta}\frac{\partial}{\partial\varphi}(M_{0,i}M_{0,k}). \end{aligned} \quad (10)$$

The important properties of tensor $\Delta_{ik,lm}$ (in particular, the number of its independent components) can be determined without calculations from symmetry considerations only. In particular, for an important case of particles, which can densely fill the 3D space, we can use the formal similarity of each term in formula (8) with the expression for the elastic energy of a deformed crystal (see [22]). In this case, the structure of tensor

$\Lambda_{ik,\,lm}$ is the same as for the elastic moduli tensor for the crystal with the shape of the unit cell coinciding with the shape of the particle. However, the number of independent components $\Lambda_{ik,\,lm}$ is actually smaller than that of the elastic moduli tensor for the crystal since the tensor components u_{ik} and v_{ik} are connected via the obvious relations

$$u_{ii} = v_{ii} = 0, \quad u_{ik}u_{ik} = v_{ik}v_{ik} = 1, \tag{11}$$

following from the condition $\mathbf{M}^2 = \mathbf{M}_0^2 = \text{const}$. In fact, the form of energy (8) for the known coefficients $\Lambda_{ik,\,lm}$ can be simplified further, noting that the combinations $u_{ik}u_{lm} + v_{ik}v_{lm}$ (but not $u_{ik}u_{lm}$ and $v_{ik}v_{lm}$ separately) appearing in the expression for energy are proportional to the corresponding invariants composed of the magnetization components. By way of example, we consider the formulas

$$u_{xx}^2 + v_{xx}^2 = M_x^2M_0^2 - M_x^4,$$

$$u_{xx}u_{yy} + v_{xx}v_{yy} = -M_x^2M_y^2,$$

$$u_{xy}^2 + v_{xy}^2 = \frac{1}{4}(M_x^2 + M_y^2)M_0^2 - M_x^2M_y^2;$$

the remaining combinations of this type can easily be restored by transposition of indices.

Thus, in the first nontrivial approximation in small parameter $(l/l_e)^2$, the calculation of the particle energy has been reduced to the following two problems: (i) writing of the solution to the Klein-Gordon equation for the scalar function $f_\alpha(\mathbf{r})$ in the bulk of the particle and (ii) the calculation of the sums in expression (9) for $\Lambda_{ik,\,lm}$. For many cases, the solution of this problem is known; for example, for particles of cylindrical shape, $f_\alpha(\mathbf{r})$ can be written in the familiar form in terms of Bessel's functions. For the above-mentioned particles densely filling the space, both problems can easily be solved in general form. If the filling of the space by translation of a particle copy results in a Bravais lattice with translation vectors $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$, the corresponding complete set is determined by the inverse lattice vectors $\mathbf{g} = m\mathbf{b}_1 + n\mathbf{b}_2 + k\mathbf{b}_3$, $(\mathbf{b}_i, \mathbf{a}_k) = \delta_{ik}$. To construct this set, we must just select the functions satisfying boundary conditions (3) from the general set of the form $\exp(i\pi\mathbf{g}\cdot\mathbf{r})$ (here, we have factor π instead of the standard factor 2π since we are interested in the solutions for which the particle size is equal to half-period). In this case, α is a set of integers n, m, and k, and the eigenvalues are $\lambda_\alpha \equiv \lambda_{n,\,m,\,k} = \pi^2(n^2 + m^2 + k^2)$. For such a set, the Fourier components of tensor $N_{ik}(\mathbf{r})$ can be determined analytically, and the sums in relation (9) converge rapidly and can easily be calculated numerically.

3. CUBIC PARTICLE

We will apply the theory developed above to the case of a cubic particle with a side length of $2l$. We assume that the origin of the coordinates coincides with the center of the cube; i.e., in the previous formulas, we have

$$\int_V d\mathbf{r} = \int_{-l}^{l} dx \int_{-l}^{l} dy \int_{-l}^{l} dz.$$

In this case, $\mathbf{g} = (m\mathbf{e}_x + n\mathbf{e}_y + k\mathbf{e}_z)/2l$ and the boundary conditions are satisfied for real-valued solutions containing $\cos[p\pi\xi/l]$ and $\sin[\pi(p + 1/2)\xi/l]$, where $\xi = x$, y, or z and $p = m$, n, or k. It is sufficient for our purpose to calculate two independent (diagonal and nondiagonal) Fourier components of tensor $N_{ik}(\mathbf{r})$, e.g.,

$$N_{xx} = \sum_{m,\,n,\,k} v_{xx}^{(m,\,n,\,k)} \cos\left(\frac{m\pi x}{l}\right)\cos\left(\frac{n\pi y}{l}\right)\cos\left(\frac{k\pi z}{l}\right),$$

$$N_{xy} = \sum_{m,\,n,\,k} v_{xy}^{(m,\,n,\,k)} \sin\left(\frac{(2m+1)\pi x}{2l}\right) \times \sin\left(\frac{(2n+1)\pi y}{2l}\right)\cos\left(\frac{k\pi z}{l}\right), \tag{12}$$

where

$$v_{xx}^{(m,\,n,\,k)} = \frac{1}{l^3}\int_V d\mathbf{r}\cos\left(\frac{m\pi x}{l}\right)\cos\left(\frac{n\pi y}{l}\right) \times \cos\left(\frac{k\pi z}{l}\right)\frac{\partial}{\partial x}\int_V d\mathbf{r}'\frac{\partial}{\partial x'}\frac{1}{|\mathbf{r}-\mathbf{r}'|},$$

$$v_{xy}^{(m,\,n,\,k)} = \frac{1}{l^3}\int_V d\mathbf{r}\sin\left(\frac{(2m+1)\pi x}{2l}\right) \times \sin\left(\frac{(2n+1)\pi y}{2l}\right)\cos\left(\frac{k\pi z}{l}\right)\frac{\partial}{\partial y}\int_V d\mathbf{r}'\frac{\partial}{\partial x'}\frac{1}{|\mathbf{r}-\mathbf{r}'|}.$$

the remaining components can be obtained from relations (12) using the obvious cyclic permutation of variables (x, y, z) and integers m, n, and k. The integrals in the formulas for $v^{(m,\,n,\,k)}$ contain only integrable singularities and these coefficients can easily be determined numerically.

In the case of the cubic symmetry, tensor $\Lambda_{ik,\,lm}$ in formula (8) contains three independent invariants $\Lambda_{xx,\,xx}$, $\Lambda_{xx,\,yy}$, and $\Lambda_{xy,\,xy}$. However, in view of relations (11), only one of the three invariants quadratic in u_{ik} and v_{ik} is independent. As a result, the energy can be written in the form

$$W_{\text{nonuniform}} = -C\frac{Vl^2}{l_e^2M_0^2}(u_{xy}^2 + u_{xz}^2 + u_{yz}^2 + v_{xy}^2 + v_{xz}^2 + v_{yz}^2).$$

Calculation of constant C gives $C = 0.61$. Passing to the

magnetization components, we find that the expression in the brackets is equal to $(M_x^2M_y^2 + M_x^2M_z^2 + M_y^2M_z^2)$; i.e., it coincides with the standard expression for the cubic anisotropy energy. This leads to the required expression for the magnetic energy in the angular variables:

$$W_{\text{nonuniform}}(\varphi_0, \theta_0) = C\frac{M_0^2Vl^2}{2l_e^2}(\sin^2 2\varphi\cos^2\theta + 4\sin^2\theta)\cos^2\theta. \quad (13)$$

Since $C > 0$, function $W_{\text{nonuniform}}(\varphi, \theta)$ assumes the minimal value (equal to zero) at six points, namely, for $\mathbf{M}_0 = \pm M_0\mathbf{e}_x$, $\mathbf{M}_0 = \pm M_0\mathbf{e}_y$, and $\mathbf{M}_0 = \pm M_0\mathbf{e}_z$. These three cases are obviously identical since they indicate the average orientation of the magnetic moment along one of the cube edges. The analytic formulas for the corresponding micromagnetic state become especially visual for $\theta = 0$ and $\varphi = 0$; i.e., the average vector $\mathbf{M}_0 = M_0(1, 0, 0)$ is parallel to the x axis. Since the sums in relations (12) converge well [12, 20], it is sufficient for a qualitative analysis to retain only the first terms in these sums. Then the components of the weakly nonuniform part of the magnetic moment in the linear approximation in $\delta\mathbf{M}$ that we are dealing with exhibit the coordinate dependence standard for a flower-type state [19],

$$\partial\mathbf{M} = M_0\frac{1}{\pi^2} \times \frac{l^2}{l_e^2}\nu_{xy}^{(0,0,0)}\left(0, \sin\frac{\pi x}{2l}\sin\frac{\pi y}{2l}, \sin\frac{\pi x}{2l}\sin\frac{\pi z}{2l}\right), \quad (14)$$

where the numerical coefficient $\nu_{xy}^{(0,0,0)} \approx -4.31$. As usual, the maximal deviation of the total magnetic moment $\mathbf{M}(\mathbf{r}) = \mathbf{M}_0 + \delta\mathbf{M}$ from the average direction is observed at the vertices of a particle, while $\mathbf{M}(\mathbf{r})$ at the center coincides with $\mathbf{M}_0$ (see Fig. 2a in [8]). The micromagnetic state for the magnetic moment orientation along some other edge of the cube can be obtained by rotating $\mathbf{M}_0$ through $\pi/2$ and by cyclic permutation of x, y, and z in (14).

It should be noted that, although we calculated in this section the particle energy for an arbitrary orientation of vector $\mathbf{M}_0$, the calculation could be significantly simplified in fact, which is important when higher orders in perturbation theory have to be employed. Indeed, the form of the energy as an invariant of the cubic symmetry group is beyond any doubt, and the value of coefficient C can be determined by calculating energy for two symmetric direction of $\mathbf{M}_0$.

4. CALCULATION OF THE AVERAGE DIRECTION OF MAGNETIC MOMENT IN PARTICLES WITH A NEARLY CUBIC SHAPE

In this section, we will consider some important examples of distortions in the cubic shape of a particle, assuming that the deformation can be reduced to extension or compression of the particle along one of the symmetry axes of the cube. Obviously, all such deformations will induce a nonzero anisotropy of the shape associated with the uniform part of the demagnetizing field. The symmetry of this anisotropy is lower than the cubic symmetry and the resultant direction of magnetization is determined by the competition of these two contributions.

It is quite easy to predict the outcome of a tetragonal or rhombic deformation, i.e., a change in the particle size along directions parallel to one of the edges of the cube. The lowering of the demagnetizing field symmetry leads to the magnetic moment orientation along the longest edge; in other words, the flower-type state, which was considered in detail in the previous publications [12] for the elongated parallelepiped obtained in this way, "survives" in the system. For a considerable compression of the cube (with the ratio of edges exceeding 2.6–2.8), a transition to the leaf state takes place; however, in this study we confine our analysis to the case of small deformations of cubic particles.

A more important question, which has never been discussed earlier, concerns the change in the direction of the magnetic moment upon a change in the angles determining the shape of the particle. Since a cube is not a mechanically rigid figure, spontaneous distortions of the shape of this type are quite feasible during the preparation of nanoparticles. In addition such distortions may take place for particles implanted in the bulk of an amorphous nonmagnetic matrix during its elastic deformation.

We will consider here two basically different cases of lowering of cubic symmetry (see Fig. 1). In the first case, the cube is stretched along the diagonal of one of the faces (e.g., in the (110) direction), being transformed into a rectangular prism with a rhombus as the base (Fig. 1a). Another example of a deformed particle is an equilateral rhombohedron (Fig. 1b), which can be obtained by extending (compressing) the cube along the principal diagonal (111). In both cases, we assume that the length of the edge remains unchanged and the change in the shape of the particle is described by only one parameter. It is convenient to choose for such a parameter the angle α between the particle edge and diagonal (110) in the xy plane. The deformation is small if the value of $\delta\alpha = \alpha - \pi/4$ is small.

To predict the results in specific cases, we will first qualitatively consider the physical features of the problem. A lowering of the particle symmetry induces a uniform demagnetizing field orienting the magnetic moment along a certain direction. In the case depicted in Fig. 1a, this direction (easy anisotropy axis of the

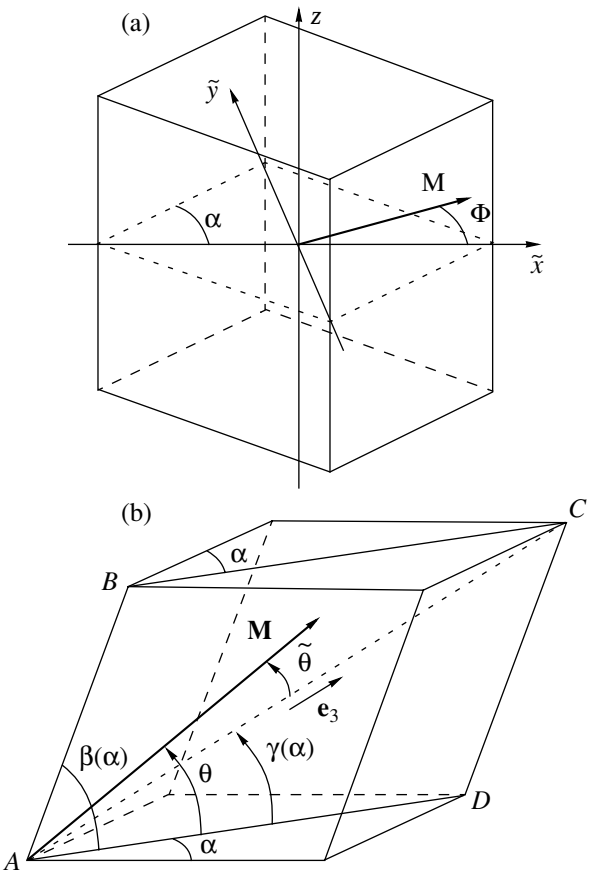


Fig. 1. The shape of particles obtained by deforming the cube along the symmetry directions: (a) prismatic deformation, extension along the diagonal of the face; (b) rhombohedral deformation, extension along the principal diagonal of the cube.

particle shape) is oriented along the longer diagonal (110) of the face. In the case of rhombohedral distortions (Fig. 1b), the selected axis is oriented along the principal diagonal (111) of the rhomboid. This axis serves as the easy axis during extension of the rhomboid and the hard axis during its compression. In both cases, the corresponding energy of the uniform demagnetizing field is proportional to $\delta\alpha$.

According to the results of calculations carried out in the previous section, a nonuniform demagnetizing field in a cubic particle tends to orient its magnetic moment along an edge of the cube. The energy associated with this field contains an additional small parameter $(l/l_e)^2$. If the particle shape is nearly cubic (i.e., $\delta\alpha$ is on the order of $(l/l_e)^2$), the contributions from the uniform and small nonuniform components of the magnetic fields can be comparable in magnitude. The competition of these contributions creates conditions for the emergence of an intermediate state, which was considered in [20] for flat particles. It was demonstrated by comparing the numerical data and the results of calculation in perturbation theory that the results of analysis are also valid qualitatively for $l \sim l_e$ ($l_e \approx 18$ nm for permalloy). Consequently, although we will use in subsequent calculations the formal inequality $l \ll l_e$ required for the applicability of perturbation theory, the results can be used for larger particles as well.

Another important circumstance considerably simplifying calculations is also worth noting. It will be proved below that the range of realization of the intermediate state is small ($\delta\alpha \sim 2°$–$3°$) even for $l \sim l_e$. Consequently, we can use the inequality $\delta\alpha \ll 1$. In the case of small deformations we are interested in, in the main approximation in small parameters $\delta\alpha$ and $(l/l_e)^2$, it is sufficient to use for the contribution of the nonuniform demagnetizing field formula (13) obtained for a cubic particle. The anisotropic contribution of the nonuniform field taking into account weak distortions of the shape gives the result of the next order of smallness in parameters $\delta\alpha$ and $(l/l_e)^2$ and does not lead to any significant effects; for this reason, this contribution will be disregarded here.

Prismatic Deformation of a Cube

Let us now analyze the particle shown in Fig. 1a. First, we must find the magnetic dipole energy of uniform fields. The calculations are simple when the tensor of demagnetizing coefficients contains only diagonal components. This is observed when the coordinate axes coincide with the symmetry axes of the system. In the given case, it is sufficient to leave the z axis unchanged and turn the x and y axes so that they coincide with the perpendicular diagonals of the rhomb. We define this rotation as shown in Fig. 1a, where the $\tilde{x}$ axis coincides with the longer diagonal. The contribution of uniform demagnetizing fields to the magnetic energy is a quadratic form in the magnetization components. To within an insignificant constant chosen so that the energy vanishes for a cubic particle, this energy can be written in the form

$$W_{\text{uniform}}(\varphi, \theta) = \frac{VM_0^2}{2} \times [(I_y - I_x)\cos^2\theta\sin^2\varphi + (I_z - I_x)\sin^2\theta]. \quad (15)$$

Here, parameters I_i are proportional to the diagonal components N_{ii} of the tensor of demagnetizing coefficient, integrated over the particle volume:

$$I_i = -\frac{4\pi}{V}\iiint N_{ii}(x, y, z)dV.$$

The corresponding integrals have only integrable singularities at the corners of the particle [23] and can easily be evaluated. To determine the average direction of

the magnetic moment in the given system, we must find the minimum of the total magnetic energy

$$W(\varphi, \theta) = W_{\text{uniform}}(\varphi, \theta) + W_{\text{nonuniform}}(\varphi, \theta),$$

where function $W_{\text{nonuniform}}(\varphi, \theta)$ can be calculated using formula (13) derived for a cubic particle if we replace $\varphi \longrightarrow \tilde{\varphi} = \varphi + \pi/4$ in accordance with the rotation of the coordinate system. If angle α becomes smaller than its value $\pi/4$ for the cube, parameters I_i must satisfy the relation $I_x < I_z < I_y$.

Minimization of the total energy of the particle over angle θ_0 shows that only two types of solutions with $\theta_0 = 0$ and $\pi/2$ can exist. The latter solution corresponds to states with $M_0 = \pm M_0 \mathbf{e}_z$ and their energy is $(M_0^2 V/2)(I_z - I_x)$. The states with $\theta_0 = 0$ correspond to the magnetic moment orientation in the *xy* plane,

$$\mathbf{M}_0 = M_0(\mathbf{e}_x \cos\varphi_0 + \mathbf{e}_y \sin\varphi_0),$$

where the value of angle φ_0 is determined by the extremum of the function of one variable $W(\varphi_0, \theta_0 = 0)$.

If we disregard the weakly nonuniform field (which can be done for very small particles), $W_{\text{nonuniform}} \ll W_{\text{uniform}}(\varphi_0, \theta_0)$ and the energy of the demagnetizing field is minimal for the average direction of the magnetic moment parallel to the $\tilde{x}$ axis (i.e., for $\varphi_0 = 0$). As the nonuniform energy (i.e., the particle size) increases, the energy minimum at point $\varphi_0 = 0$ vanishes and a new minimum appears for $\varphi_0 = \Phi \neq \pi k/2$, where $\cos 2\Phi = (I_y - I_x) l_e^2 / 4Cl^2$. In this state, the direction of $\mathbf{M}_0$ does not coincide with any symmetric directions of the body; in analogy with the previous publications, we refer to this state as the intermediate state. This state is stable in the entire range $0 < \cos 2\Phi < 1$ of its existence; it is realized for particles with the length of the edge larger than a certain critical value $2l_{\text{crit}}$, where $l_{\text{crit}} = l_e \sqrt{(I_{yy} - I_{xx})/4C}$.

The value of l_{crit} is determined by the exchange length of the material and by the strain; its value is small for $\alpha \longrightarrow 45°$, but rapidly increases with increasing strain. Thus, the possibility that the intermediate state is realized in a right prism of a preset shape depends on its size and the exchange length of the prism material. We carried out calculations in perturbation theory for permalloy particles. Figure 2 shows the results for the rectangular prism depicted in Fig. 1a. Three different values of α depend to three curves describing the dependence of angle Φ on the particle size, It should be noted that the value of l_c is small for small strains; it is smaller than the exchange length, which literally corresponds to the applicability of perturbation theory.

Thus, in a weakly deformed cubic particle, the direction of average magnetic moment $\mathbf{M}_0$ does not coincide with any symmetry axis of the particle. However, as the strain increases (or, which is the same, the linear size of the particle increases), the uniform magnetic field rotates vector $\mathbf{M}_0$ towards the symmetry axis (the longer diagonal of the base) and the magnetic moment remains parallel to this symmetry direction. The reorientation of the average magnetic moment upon a change in the particle parameters can be interpreted as a certain reorientation phase transition. It can easily be seen that this phase transition is a second-order transition, for which the order parameter is angle Φ. Indeed, states with $\Phi \neq 0$ possess a lower symmetry that the state with $\Phi = 0$ and the dependence $\Phi(l)$ in the vicinity of the phase transition has the typical root dependence

$$\Phi = \begin{cases} 0, & l < l_{\text{crit}}, \\ \sqrt{l/l_{\text{crit}} - 1}, & l \geq l_{\text{crit}}. \end{cases} \tag{16}$$

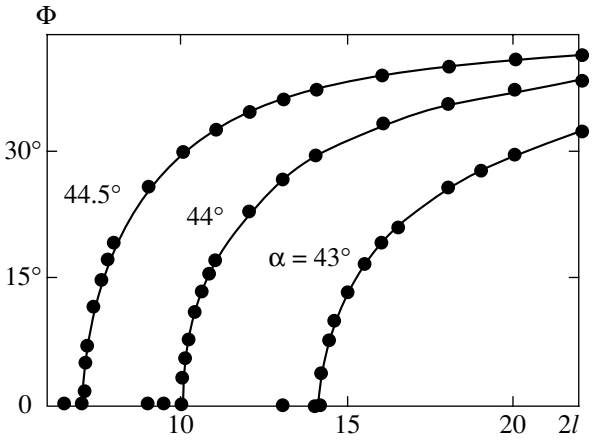


Fig. 2. Dependence of angle Φ describing the deviation of the average magnetization from the symmetry direction on the particle size *l* for three values of prismatic deformation of the cube. Here and in the following figures, the numerical values were obtained for permalloy.

This behavior is clearly shown in Fig. 2, where curves $\Phi = \Phi(l)$ in the vicinity of the value of l_{crit} corresponding to each angle α become nearly vertical (i.e., these curves have an infinitely large derivative). The pattern of phase transitions can be visualized more clearly via the dependence of angle Φ on angle α for a fixed size of the particle, which is depicted in Fig. 3. Here, angles $\alpha < 45°$ describe extension and angles $\alpha > 45°$ describe the compression of the particle along the direction (110). The emergence of two second-order phase transitions near which the magnetization is reoriented to symmetric states $\Phi \longrightarrow 0$ (extension) and $\Phi \longrightarrow \pi/2$ (compression) can be explained by the fact that the extension along one of the diagonals of a face in the given geometry is equivalent to compression along the other diagonal of the same face.

Thus, when the cube is deformed along a diagonal of its face, a symmetric state of the flower type, which is typical of an ideal cube, is never realized. Instead of

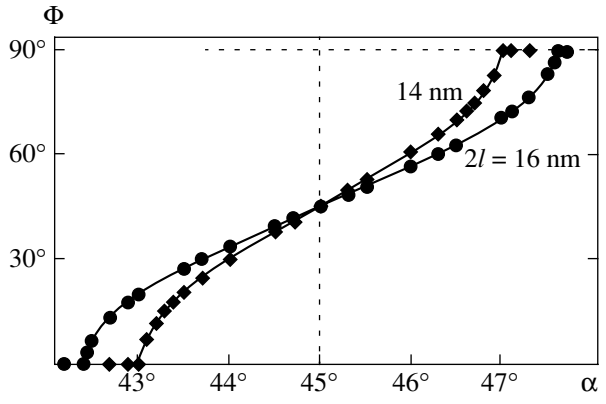


Fig. 3. Dependence of angle Φ on the value of angle α describing the prismatic strain of the particle for two values of particle size *l*.

this state, two types of states can be observed: a more symmetric state with $\Phi = 0$, resembling the leaf-type state observed earlier for flat square particles, appears for considerable strains, while for small strains, a less symmetric intermediate state is formed, in which the average direction of the magnetization is not connected with any symmetry axis of the particle.

Rhombohedral Deformation

Let us now consider the extension or compression of a cube along its principal diagonal (Fig. 1b corresponds to extension). The rhomboid has a uniaxial anisotropy with a third-order principal axis parallel to the spatial diagonal of the cube. As before, we assume that the parameter of the problem is angle α between the diagonal and the side of the rhomb at the base of the rhomboid. In the system with the $\mathbf{e}_3$ axis directed along the principal axis of the rhomboid (segment *AC* in the figure), the energy of uniform demagnetizing fields depends only on angle $\tilde{\theta}$ between the $\mathbf{e}_3$ axis and the magnetic moment. This energy can be written in the simple form

$$W_{\text{uniform}}(\theta_0) = \frac{1}{2}M_0^2 VB(\alpha)\sin^2\tilde{\theta}, \qquad (17)$$

where function $B(\alpha)$ is positive (negative) in the case of the extension (compression) of the cube. In the same system of coordinates, the energy of weakly nonuniform fields has the form

$$W_{\text{nonuniform}}(\tilde{\varphi}, \tilde{\theta}) = C\frac{M_0^2 Vl^2}{2l_e^2} \times\left(2\sin^2\tilde{\theta} + \frac{7\cos^4\tilde{\theta}}{3} + \frac{4\sqrt{2}\cos 3\tilde{\varphi}\cos\tilde{\theta}\sin^3\tilde{\theta}}{3}\right), \qquad (18)$$

which is typical of the rhombohedral symmetry with the C_3 axis parallel to the $\mathbf{e}_3$ axis. It was mentioned above that the form of energy (18) and value of constant *C* can be taken the same as for the undeformed cube. The first two terms in the parentheses describe purely uniaxial anisotropy, while the last term describes anisotropy in the basal plane. The energy extremum in the value of angle φ is determined by the relation $\sin 3\tilde{\varphi} = 0$; i.e., $\cos 3\tilde{\varphi} = \pm 1$. Further, the energy minimum corresponds to the value of $\cos\tilde{\theta}\cos 3\tilde{\varphi} = -1$; i.e., for each value of the projection of the magnetization on the C_3 axis, triple degeneracy of the minima of the total energy $W(\tilde{\varphi}, \tilde{\theta}) = W_{\text{uniform}}(\tilde{\theta}) + W_{\text{nonuniform}}(\tilde{\varphi}, \tilde{\theta})$ arises. Without any loss of generality of the problem, we can analyze the behavior of the magnetic moment lying in one of the three selected planes, choosing, for example, angle $\tilde{\varphi} = 0$; the remaining equivalent states can be obtained by rotations about the $\mathbf{e}_3$ axis through an angle of 120° and the changes in the direction of magnetization vector $\mathbf{M}_0$.

Thus, we consider the rotation of the magnetic moment in the *ABCD* plane passing through the principal diagonal and the corresponding diagonals of the faces (see Fig. 1b). For such a rotation, the energy of the magnet is defined by function $W(\tilde{\theta}) \equiv W(\tilde{\varphi}, \tilde{\theta})\big|_{\tilde{\varphi}=0}$. It should be noted above all that the equation $\partial W(\tilde{\theta})/\partial\tilde{\theta} = 0$ always has the solution $\tilde{\theta} = 0$ describing the orientation of the average magnetization along the C_3 axis. This solution corresponds to a state of the leaf type, which is most symmetric, but which is not realized in a cubic particle. The type of this extremum is determined by the sign of the strain. It should be noted that, in the case of a rhombohedral deformation, the extension and compression of the cube along the principal diagonal are not physically equivalent effects as in the case of extension and compression along the diagonals of the faces. Let us first describe the case of extension.

It was noted above that a state of the flower type is stable in an undeformed cube, while solution $\tilde{\theta} = 0$ (magnetization distribution of the leaf type) corresponds to the energy maximum. In the case of extension, the value of constant $B(\alpha) > 0$; the C_3 axis becomes the easy axis of induced anisotropy (17), and the highly symmetric state with $\tilde{\theta} = 0$ and a magnetization parallel to this axis may become stable. A simple analysis of energy $W(\tilde{\theta})$ taking into account formulas (17) and (18) shows that this state is stable to small deviations from the C_3 axis when the condition $B(\alpha) > 8Cl^2/3l_e^2$ is satisfied; i.e., for the preset strain amplitude α, this state is stable for quite small sizes of the particle, $l < l_c^{(1)}(\alpha) = l_e\sqrt{3B(\alpha)/8C}$. As the particle size increases above the critical value $l_c^{(1)}(\alpha)$, the action

of nonuniform demagnetizing fields violates the stability of this state and leads to the formation of a new state, in which the average magnetic moment forms a certain angle $\tilde{\theta} \neq 0$ with the principal axis of the rhomboid. It should be recalled that in fact there are six such states, three of which are obtained by rotations about the C_3 axis through an angle multiple to 120°; the remaining states can be obtained via the substitution $\mathbf{M}_0 \longrightarrow -\mathbf{M}_0$. The rotation about the C_3 axis does not change the magnetization in state $\tilde{\theta} = 0$; as a result, the symmetry of a state with $\tilde{\theta} \neq 0$ is lower than the symmetry of a state with $\tilde{\theta} = 0$. Consequently, the states with $\tilde{\theta} \neq 0$ are typical examples of the intermediate state. The search for solutions to the equation $\partial W(\tilde{\theta})/\partial\tilde{\theta} = 0$ is a cumbersome problem. To visualize the result, it is convenient to analyze these states qualitatively, returning to the coordinate system with the *x*, *y*, and *z* axes directed along the edges of an undeformed cube and to polar angle θ. Angles $\tilde{\theta}$ and θ are connected via the relation $\theta = \tilde{\theta} + \gamma(\alpha)$, where $\cos\gamma = 1/\sqrt{3}\cos\alpha\sqrt{4\cos^2\alpha - 1}$ (see Fig. 1b).

In terms of angle θ, under the condition $\tilde{\varphi} = 0$, the total energy of the particle $W = W_{\text{uniform}} + W_{\text{nonuniform}}$ can be written in the form

$$\frac{W(\theta)}{VM_0^2} = -\frac{\sqrt{2}B}{6}\sin 2\theta - \left(\frac{B}{12} - \frac{Cl^2}{4l_e^2}\right)\cos 2\theta + \frac{3Cl^2}{8l_e^2}\sin^2 2\theta. \quad (19)$$

After the substitution $\theta \longrightarrow \Theta/2$, this expression coincides with the formula for the energy of a uniaxial ferromagnet,

$$W_{ua}(\Theta) = -H_z\cos\Theta - H_x\sin\Theta + \frac{K}{2}\sin^2\Theta,$$

in an external magnetic field oriented arbitrarily relative to the easy axis (*z* axis); the states formed in this case are well known (see [24]). In such a system, in the range of values $H_z^{2/3} + H_x^{2/3} \leq K^{2/3}$, there exists a region of metastable states, while two phases with different magnetizations differing in the sign of the *z* projections coexist for $H_z = 0$ and $H_x < K$. The equality sign in this relation ($H_z^{2/3} + H_x^{2/3} = K^{2/3}$) determines the point of stability loss for one of the phases. After substitution of the corresponding relations from (19) into this condition, we obtain the instability condition for a phase with $\tilde{\theta} = 0$ determined above, while the second root of this equation leads to the stability loss condition for a phase with $\tilde{\theta} = 0$ in the form $B(\alpha) = 3.0445Cl^2/l_e^2$. Thus, the asymmetric phase is stable for $l > l_e^{(2)}(\alpha) = 0.573l_e\sqrt{B(\alpha)/C}$. The condition $B(\alpha) = 3Cl^2/l_e^2$ for vanishing of the coefficient of $\cos 2\theta$ in formula (19) gives the point $l = l_t(\alpha)$, $l_t(\alpha) = l_e\sqrt{B(\alpha)/3C}$ of the coexistence of a phase with $\tilde{\theta} = 0$ ($\theta = \gamma(\alpha)$, $\gamma(\alpha) \longrightarrow 35.3°$ as $\alpha \longrightarrow 45°$) and an asymmetric phase for which $\theta \approx \pi/2 - \gamma(\alpha)$ at this point ($\theta \longrightarrow 54.7°$ as $\alpha \longrightarrow 45°$). Thus, a symmetric phase with $\tilde{\theta} = 0$ realizes the absolute minimum of the particle energy for quite small particle size, $l < l_t(\alpha)$, while the asymmetric case is favorable for $l > l_t(\alpha)$. It should be noted that the values of $l_c^{(2)}$, the above value of the particle size for which the phase with $\tilde{\theta} = 0$ loses stability, $l_c^{(1)} \approx 0.612l_e\sqrt{B(\alpha)/C}$, and the size $l_t(\alpha) \approx 0.577l_e\sqrt{B(\alpha)/C}$ for which the phases coexist satisfy the inequality $l_c^{(2)} < l_t(\alpha) < l_c^{(1)}$.

Thus, the general pattern of states for a rhombohedral particle with a preset strain (preset value of α) can be described as follows. For small particles with $l < l_c^{(2)}$, only a symmetric distribution of the leaf type can be realized. For $l > l_c^{(2)}$, a less symmetric intermediate phase, which was initially metastable, appears. In a narrow region $l_c^{(2)}(\alpha) < l < l_c^{(1)}(\alpha)$ (this region is on the order of 1 nm for permalloy particles with an exchange length on the order of 18 nm), the particle has two stable states, one of which is the ground state and the other is metastable. The transition between these states occurs at a certain value of $l = l_t(\alpha)$; this transition corresponds to an abrupt change in the direction of magnetization. When the length of the edge of the particle becomes larger than the second critical value $2l_c^{(1)}(\alpha)$, the state with the average direction of the magnetic moment along the principal diagonal loses its stability and only one stable state (namely, the intermediate state) remains in the system. The value of θ increases with the size of the rhomboid, asymptotically tending to the value $\pi/2 - \beta(\alpha)$, where $\beta(\alpha)$ is the angle between the diagonal of a face and the opposite face of the particle; $\cos\beta = \cos 2\alpha/\cos\alpha$ (see Fig. 1b). The dependence of angle $\tilde{\theta}$ on *l* for the intermediate state in permalloy particles in the range of sizes $l > l_c^{(2)}(\alpha)$ for which this state is stable is depicted in Fig. 4 for three different values of α.

Thus, as in the case of prismatic deformation, for a strong extension of the particle along the principal diagonal, the uniform magnetic field turns the magnetic moment towards the strain axis, which can again be interpreted as a reorientation phase transition upon a change in the parameters of the particle. However, the

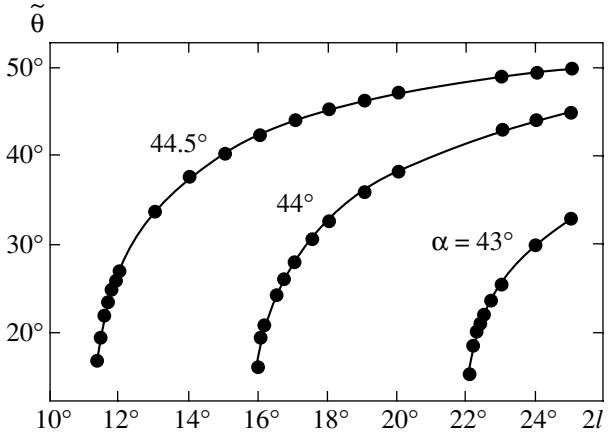


Fig. 4. Dependence of angle $\tilde{\theta}$ between the magnetic moment and the principal diagonal of the particle in the shape of an elongated rhomboid on the particle size l for the intermediate state in the region where this state is stable.

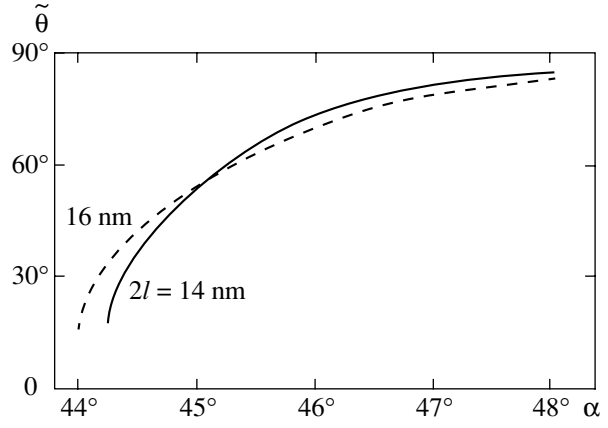


Fig. 5. Dependence of angle $\tilde{\theta}$ between the magnetic moment of the particle and the principal diagonal of the rhomboid obtained by extension ($\alpha < 45°$) or compression ($\alpha > 45°$) of the cube along its principal diagonal on angle α for two fixed values of the particle edge.

situation for this deformation differs from that considered above. Although the reorientation of the average magnetic moment occurs between the states with a higher and a lower symmetry, this phase transition is a typical first-order transition, for which the phase coexistence region is $l_c^{(2)}(\alpha) < l < l_c^{(1)}(\alpha)$.

In the case of the compression of the cube along its principal diagonal, induced anisotropy is an easy-plane anisotropy and a symmetric phase with the magnetization along the principal diagonal is never realized. In this case, there exists only the intermediate state in which the average magnetic moment $\mathbf{M}_0$ for a small particle size asymptotically tends to the position in the plane perpendicular to the principal diagonal, $\tilde{\theta} \longrightarrow \pi/2$, or $\theta \longrightarrow \pi/2 + \gamma(\alpha)$. For large values of l (or for small distortions of the cubic shape) the direction of $\mathbf{M}_0$ approaches the face of the particle, which is opposite to the diagonal of the *AD* face in Fig. 1b; i.e., $\theta \longrightarrow \beta(\alpha)$, or $\tilde{\theta} \longrightarrow \beta(\alpha) - \gamma(\alpha)$. Consequently, in this range of the parameters, the average magnetic moment is rotated monotonically and the reorientation transitions do not take place.

To visualize the pattern of the variation of the average value of the particle magnetization for various strains, we will trace the motion of vector $\mathbf{M}_0$ after the first-order phase transition for $l = l_t(\alpha)$, which exists in the case of extension of the cube ($\alpha < 45°$) upon an increase in the value of α to $\alpha = 45°$ and further upon the subsequent compression of the cube for $\alpha > 45°$. The dependence of angle $\tilde{\theta}$ on angle α in this case is a monotonic function of α, assuming the value $\tilde{\theta} = \pi/2 - \gamma(45°) \approx 54.775°$ for the undeformed cube ($\alpha = 45°$); at this point, the dependences $\tilde{\theta}(\alpha)$ for all values of the particle size intersect. The dependences $\tilde{\theta}(\alpha)$ obtained numerically for two fixed lengths of the edge for extension and compression of the cube ($\alpha < 45°$ and $\alpha > 45°$, respectively) along the principal axis are shown in Fig. 5.

5. CONCLUDING REMARKS AND DISCUSSION

Using cubic magnetically soft particles as an example, we have proved that even an insignificant deviation of the particle shape from symmetry might cause a significant change in the magnetic properties of the particle even up to a phase transition to another ground state with a lower symmetry. It was shown that particles having the same shape and made of the same material might have different ground states depending on their size. These conclusions are especially important for arrays of regular particles, which can be prepared most easily by the methods of electron-beam lithography. Such physical objects are investigated in physical experiments most frequently and can be used in practice [1–6]. The existence of intermediate magnetic states described in this paper can be regarded as an important factor affecting the properties of the array as a whole. In all probability, the stability of information recording on an individual particle and the presence of such states in the range of their existence will be an unfavorable factor for information recording devices operating on the arrays of magnetic nanoparticles. On the other hand, the existence of such states is interesting from the standpoint of the fundamental physics of magnetism. Particles in the intermediate state must exhibit all features known for bulk magnets in the vicinity of reorientation phase transitions. In particular, in the vicinity of a transition to the intermediate state, anomalies in the static susceptibility, softening of intrinsic vibrational modes, and other effects must be observed.

These peculiarities of the physical properties of particles in the intermediate magnetic state and their ordered arrays will probably be used in functional electronic devices.

Let us now consider the applicability of the method presented in this paper for describing particles with a size of up to several tens of nanometers (we will consider the estimates obtained for permalloy, which is commonly used in applications), which exhibit the quasi-homogeneous states considered in this paper. (For normal permalloy particles with a size exceeding 100 nm, an eddy state with a closed magnetic flux is realized [7]). It should be emphasized that the proposed method is based not only on perturbation theory, but is a self-consistent approach to studying the properties of small magnetic particles. The calculations carried out in the framework of perturbation theory can be treated as the first step in this direction. In fact, we can demonstrate as a results of this calculation that the static state can be described on the basis of the effective energy written in the form of a combination of powers of the average magnetic moment components, which is invariant to the symmetry group of the particle. A qualitatively identical form of energy is obviously applicable for $l \geq l_e$ as well. To draw conclusions about the type of ground state (in particular, the presence or absence of phase transitions between the symmetry-determined and intermediate phases and the type of these transitions), often it is important to know only the sign of the coefficients of these invariants. It is not surprising that the results of analysis based on this method are in accordance with the experimental data (e.g., the conclusion that the ground state of a cubic particle corresponds to a flower-type state) as well as the result obtained with the help of computer simulation [20].

The above analysis makes it possible to study particles with a complex shape, in particular, particles for which discretization of the problem by dividing the particle volume into small domains whose shape repeats the shape of the particle [21] is too complicated. It is important to note that the symmetry approach often leads to certain conclusions about the nature of states prior to routine but cumbersome calculations of the corresponding contributions to the energy in perturbation theory. By way of example, let us consider a cylindrical particle with a symmetry axis coinciding with the z axis. This particle obviously exhibits anisotropy of shape of the BM_z^2 type with coefficient B depending on the height-to-diameter ratio as well as the contribution of a nonuniform demagnetizing field. For the latter contribution, the terms in the expression for energy to within terms on the order of $(l/l_e)^2$ obviously contain the sum of two independent invariants $C_1 M_z^4 + C_2(M_x^2 + M_y^2)M_z^2$, which is sufficient for describing reorientation of the magnetic moment from the symmetry axis to the plane upon a change in the particle shape (i.e., coefficient B). Another example—the orientation of magnetization in the basal plane for a particle in the form of a regular hexagonal prism—can be completely described only by calculating the sixth-order terms in the magnetic moment components, which requires the inclusion of terms on the order of $(l/l_e)^4$. However, this cumbersome calculation can be simplified by the fact that it is possible to determine from symmetry considerations the form of the important term $C(l^4/M_0^4 l_e^4)[(M_x + iM_y)^6 + (M_x - iM_y)^6]$. To find this term (in fact, to calculate coefficient C), it is sufficient to consider the problem only for two directions of $\mathbf{M}_0$ in the basal plane.

Such an approach will also be useful in other problems; in particular, for describing the behavior of the average magnetic moment of nonellipsoidal particles in a strong external magnetic field $\mathbf{H}_0$. This also applies to the problem of reversal of magnetization of nonellipsoidal particles, which is important for applications. It should be noted that the inclusion of a weak field (in particular, the study of the field-induced spin reorientation of states with close vectors $\mathbf{M}_0$ in the vicinity of transitions to the intermediate state) does not present any difficulty and can easily be carried out on the basis of the effective energy of the type (13), (15), or (18), supplementing it with the Zeeman term in the form $-\mathbf{H}_0 \cdot \mathbf{M}_0$. However, the presence of a strong external field $|\mathbf{H}| \sim 4\pi M_0$ requires a certain modification of the calculation. In particular, the structure of the operators in Eqs. (4) changes in this case ($-l_e^2 \nabla^2 m_{1,2} \longrightarrow -l_e^2 \nabla^2 m_{1,2} + h_{1,2}$, where $h_{1,2} = H_{1,2}/M_0$ and H_1 and H_2 are the corresponding components of the external field). However, such a modification can easily be carried out. In particular, the eigenfunctions of the problem remain unchanged and the nature of variation of eigenvalues of the problem is clear (the eigenvalues increase with the field). We can easily predict the change in the results as well; if the contribution of the uniform demagnetizing field depends only on the particle geometry, the contribution of the next orders of perturbation theory, which is inversely proportional to eigenvalues λ_α, effectively decreases.

ACKNOWLEDGMENTS

We are grateful to J. Tucker and N. Usov for fruitful discussions.

REFERENCES

1. B. Hillebrands, C. Mathieu, C. Hartmann, *et al.*, J. Magn. Magn. Mater. **175**, 10 (1997).
2. E. F. Wassermann, M. Thielen, S. Kirsch, *et al.*, J. Appl. Phys. **83**, 1753 (1998).
3. V. Novosad, Y. Otani, A. Ohsawa, *et al.*, J. Appl. Phys. **87**, 6400 (2000).

4. S. O. Demokritov, B. Hillebrands, and A. N. Slavin, Phys. Rep. **348**, 441 (2001).
5. R. P. Cowburn, D. K. Koltsov, A. O. Adeyeye, and M. E. Welland, J. Appl. Phys. **87**, 7082 (2000).
6. R. P. Cowburn and M. E. Welland, Science **287**, 1466 (2000).
7. N. A. Usov, J. Magn. Magn. Mater. **203**, 277 (1999).
8. M. E. Schabes and H. N. Bertram, J. Appl. Phys. **64**, 1347 (1988).
9. N. A. Usov, J. Magn. Magn. Mater. **125**, L7 (1993).
10. N. A. Usov and S. E. Peschany, J. Magn. Magn. Mater. **130**, 275 (1994).
11. N. A. Usov and S. E. Peschany, J. Magn. Magn. Mater. **174**, 247 (1997).
12. N. A. Usov and S. E. Peschany, J. Magn. Magn. Mater. **135**, 111 (1994).
13. R. P. Cowburn, A. Ercole, S. I. Gray, and J. A. C. Bland, J. Appl. Phys. **81**, 6879 (1997).
14. R. P. Cowburn, A. O. Adeyeye, and M. E. Welland, Phys. Rev. Lett. **81**, 5414 (1998).
15. R. P. Cowburn and M. E. Welland, Appl. Phys. Lett. **72**, 2041 (1998).
16. R. P. Cowburn and M. E. Welland, Phys. Rev. **58**, 9217 (1998).
17. W. F. Brown, Jr., *Micromagnetics* (Interscience, New York, 1963; Nauka, Moscow, 1979).
18. A. I. Akhiezer, V. G. Bar'yakhtar, and S. V. Peletminskiĭ, *Spin Waves* (Nauka, Moscow, 1967; North-Holland, Amsterdam, 1968).
19. E. V. Tartakovskaya and B. A. Ivanov, Mater. Sci. Forum **373–376**, 213 (2001).
20. E. V. Tartakovskaya, B. A. Ivanov, and J. W. Tucker, J. Appl. Phys. **89**, 8348 (2001).
21. H. Fukushima, Y. Nakatani, and N. Hayashi, IEEE Trans. Magn. **34**, 193 (1998).
22. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity*, 3rd ed. (Nauka, Moscow, 1965; Pergamon, New York, 1986).
23. A. Thiaville, D. Tomas, and J. Miltat, Phys. Status Solidi A **170**, 125 (1998).
24. L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Gostekhizdat, Moscow, 1957; Pergamon Press, Oxford, 1960).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 1027–1036.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1172–1183.

SOLIDS
Electronic Properties

Evolution of the Optical and Magnetooptical Properties of Amorphous Metal–Insulator Nanocomposites

E. A. Gan'shina[a,*], M. V. Vashuk[a], A. N. Vinogradov[a], A. B. Granovsky[a], V. S. Gushchin[a], P. N. Shcherbak[a], Yu E. Kalinin[b], A. V. Sitnikov[b], Chong-Oh Kim[c], and Cheol Gi Kim[c,]**

[a]*Moscow State University, Moscow, 119992 Russia*
[b]*Voronezh State Technical University, Voronezh, 394026 Russia*
[c]*ReCAMM, Chungnam National University, Taejon 305-764, Korea*
**e-mail: eagan@magn.ru*
***e-mail: cgkim@cnu.ac.kr*
Received October 13, 2003

Abstract—Magnetic, optical, and magnetooptical (MO) properties of $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ and $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ granular nanocomposites of the amorphous ferromagnetic metal–insulator type were studied in a broad range of the magnetic component concentrations x. The MO response of nanocomposites increases in the vicinity of the percolation transition. Using the experimentally determined values of optical constants and the equatorial Kerr effect, the diagonal and nondiagonal components of the permittivity tensor of nanocomposites were calculated for the first time. The nondiagonal components of this tensor are nonlinear functions of x, the most pronounced variations being observed near the percolation threshold. Experimental data on the MO effect and the permittivity tensor were theoretically modeled within the framework of the effective medium approximation and the Maxwell–Garnett approximation. The most adequate description was obtained with the symmetrized Maxwell–Garnett approximation, which provides for a good (semiquantitative) agreement between theory and experiment under certain assumptions about the microstructure of nanocomposites. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The considerable interest in modern composite materials is related to a variety of unusual and useful properties, which make these materials highly promising both for practical applications and for basic research. The magnetic granular alloys can be divided into two types: ferromagnetic metal–nonmagnetic metal and ferromagnetic metal–insulator. The latter alloys are also frequently referred to as composites. Granular alloys of the former type exhibit the phenomenon of giant magnetoresistance [1], while the latter systems exhibit tunneling magnetoresistance and the giant Hall effect [2]. The magnetooptical (MO) properties of granular alloys attract attention because these materials can serve as magnetoactive media for perpendicular magnetic recording and in contactless MO magnetic field sensors and temperature sensors. Recently, it was suggested to use granular alloys as MO elements in magnetophotonic crystals. All these applications stimulate the search for optimum compositions ensuring increased MO response.

Specific features inherent in nanodimensional objects must be manifested not only in the magnetotransport properties, but in the linear and nonlinear MO effects as well, although the nature of these manifestations in nanostructural materials is still incompletely clear. Such effects must be sensitive to variations in the magnetic and electron structures, scattering mechanisms, characteristic dimensions, and the shape and topology of nanodimensional inhomogeneities. This was evidenced by the results of recent investigations of the MO spectra of granular alloys and nanocomposites [3–11], which revealed numerous peculiarities in the linear and nonlinear MO Kerr effect and some other properties.

For example, investigations of granular systems showed evidence of enhanced MO response [3, 4], which could be due to variation of both the MO parameters and the optical characteristics. In particular, investigation of the MO spectra of a multilayer $[Co(x)/SiO_2(y)]_n$ system in the region of 3 eV [3] showed a severalfold increase in the MO response amplitude as compared to that for a homogeneous cobalt film of the same thickness. An analogous resonance increase in the MO effect was observed for the MnAs/GaAs multilayers [4]. An analysis of the spectral dependence of the permittivity tensor components of this system showed that the main contribution to the enhanced MO response was due to an increase in the MO activity in a narrow spectral interval. Investigations of the granular $Co/Al_2/O_3$ [5] and $CoFeZr/SiO_n$ [6] systems of the metal–insulator type revealed a new photo-

refractive effect in the IR wavelength range, the magnitude of which was ten times greater than that of the traditional even and odd MO effects.

The MO spectra of granular alloys were described using several calculation methods based on the effective medium approximation (EMA) [12–17]. Systems of the metal–metal type have been successfully described in the Bruggeman approximation [13, 14]. However, this approach is frequently inapplicable to systems of the metal–insulator type featuring a strong dependence of their properties on the microstructure. This requires using a symmetrized Maxwell–Garnett (SMG) approximation [15, 16] taking into account finer details of the system topology.

Despite a large number of publications devoted to the MO properties of magnetic nanocomposites, no systematic investigations of such systems were undertaken that would allow constructing the effective permittivity tensor. Analysis of this tensor can provide more complete information about the observed phenomena. On the other hand, the MO properties of granular alloys based on amorphous components remain almost unstudied. In the nearest future, such materials will probably combine all the unique properties of magnetic composites and those of the amorphous substances.

This paper reports on the results of investigation of the magnetic, optical, and MO properties of $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ and $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ granular nanocomposites of the amorphous ferromagnetic metal–insulator type. Based on these experimental data, the permittivity tensor of nanocomposites of this type was calculated for the first time and theoretically modeled within the framework of the Bruggeman approximation and the Maxwell–Garnett approximation. Special attention was devoted to finding correlations between the microstructure of a nanocomposite and its MO response. The extremely strong dependence of the MO effects on various structural parameters (in particular, on the shape, size, and the distribution of granules) of nanocomposites is an important advantage of the MO methods of investigation [3, 4, 7–10].

2. SAMPLE PREPARATION AND EXPERIMENTAL METHODS

The samples of $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ and $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ granular nanocomposites were prepared by ion beam sputter deposition in an argon atmosphere [18]. This process was carried out in a vacuum chamber equipped with three ion beam sources and alloy or composite targets. The alloy targets ($Co_{45}Fe_{45}Zr_{10}$ and $Co_{41}Fe_{39}B_{20}$) were prepared by induction melting of the alloy components in vacuum. The components were high-purity cobalt (99.98% Co), carbonyl iron, boron, and zirconium taken in a ratio corresponding to the desired composition. A composite targets comprised an alloy target with single crystal quartz plates fastened on its surface. The thickness of sputter-deposited film samples was varied within 0.15–6.5 μm. The average granule size monotonically decreased with increasing insulator content: from 5–7 nm in the composite with a high metal content (x = 56%) to 2–4 nm in the composite with x = 30% (here and below the magnetic (metal) phase concentration x is given in at.%). The structure of nanocomposites was studied by transmission microscopy, and their compositions were checked by electron probe X-ray microanalysis.

The measurements of the electric resistivity as a function of the composition showed the presence of a percolation threshold at $x_{per} \approx 43\%$. The results of magnetoresistance measurements revealed the tunneling magnetoresistance effect, whereby the magnetoresistance exhibits a sharply pronounced maximum near the percolation threshold and a decrease to zero behind this point. For the $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ system, room-temperature tunneling magnetoresistance reached 4% in a magnetic field of 11 kOe; the magnitude of this effect in the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system was somewhat lower [18].

The MO properties of nanocomposites were studied by measuring the MO Kerr effect in the equatorial (transverse) geometry, which consists in a relative change of the intensity of the p-wave of a linearly polarized light upon reflection from samples in magnetized and nonmagnetized states:

$$\delta = \frac{I(H) - I(0)}{I(0)},$$

where $I(H)$ and $I(0)$ are the reflected light intensities measured with and without applied magnetic field.

The spectral and field dependences of the transverse Kerr effect (TKE) were measured using an automated MO spectrometer in the range of incident photon energies from 0.5 to 4 eV. The measurements were performed for several angles of incidence of the primary light beam. The amplitude of the applied alternating magnetic field reached up to 3.5 kOe. The response signal was detected using a dynamic technique, which allowed measuring the relative changes of the reflected light intensity as small as 10^{-5} with an error not exceeding 5%. The static near-surface hysteresis loops were measured using the MO magnetometer in a meridional geometry at E = 1.98 eV and the angle of light incidence $\varphi = 70°$. The optical constants n and k were determined ellipsometrically [19] in the photon energy range from 0.5 to 3.4 eV. All measurements were performed at room temperature.

3. EXPERIMENTAL RESULTS

Figure 1a shows the spectral dependences of the TKE in $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ nanocomposites. For this material, as well as in other granular alloys, the MO spectra of nanocomposites significantly differ from the

spectrum of the corresponding homogeneous amorphous alloy $Co_{41}Fe_{39}B_{20}$ (x = 100%). Indeed, the MO response of $Co_{41}Fe_{39}B_{20}$ monotonically decreases to nearly zero with decreasing incident photon energy, whereas the TKE spectrum of $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ nanocomposite measured at $\varphi = 70°$ exhibits a change in the sign of the response signal and has a large negative maximum in the region of E = 1.3 eV.

Analogous behavior was observed in the spectra of samples of the $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ system (see Fig. 1b), the only difference being that the maximum in the signal response at $\varphi = 70°$ shifts to E = 1.0 eV. Figure 1b well illustrates variation of the TKE spectrum (typical of both systems studied) depending on the light incidence angle φ. A decrease in this angle to 60° sharply changes the shape of spectrum, which becomes closer to that for a homogeneous amorphous alloy of the corresponding composition.

The dependence of the TKE on the concentration of magnetic component in the nanocomposite is manifested for light incidence angles close to 70°, which is related to their proximity to the Brewster angle for nanocomposites with $x < x_{per}$. Figure 2 shows the concentration dependence of the TKE for nanocomposites of the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system. As can be seen from these data, the TKE signal in the "red" spectral range exhibits a nonmonotonic dependence on the ferromagnetic component concentration, featuring a break at an x value close to x_{per}. In addition, the absolute value of TKE at 1.3 eV in $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ with x = 43% amounts to 10^{-2}, which is more than three times the analogous value for the corresponding homogeneous alloy $Co_{41}Fe_{39}B_{20}$ (x = 100%) for the same photon energy. Thus, the granular system near the percolation threshold exhibits a significant increase in the TKE magnitude in a narrow photon energy range. This increase takes place despite an almost one-half lower content of the ferromagnetic component in the composite as compared to that in the homogeneous alloy. An analogous result is observed for the $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ system, where the maximum effect also takes place in the vicinity of x_{per}.

Figure 3 shows the hysteresis loops observed for samples of the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system with various concentrations x of the magnetic phase. For $x < x_{per}$, the dependence of the TKE on the magnetic field strength exhibits no hysteresis and has a shape characteristic of superparamagnets. The measurements performed for various sample orientations showed that the sample material occurred in the isotropic state. However, in the region of x close to x_{per}, the magnetic properties of this system exhibit a significant change. A sample with x = 47% exhibited anisotropic ferromagnetic behavior of the field dependence of TKE, with a coercive force of $H_c \approx 10$ Oe and a saturation field of $H_s \approx 30$ Oe along the easy magnetization axis and $H_s \approx$ 100 Oe along the hard magnetization axis. As the mag-

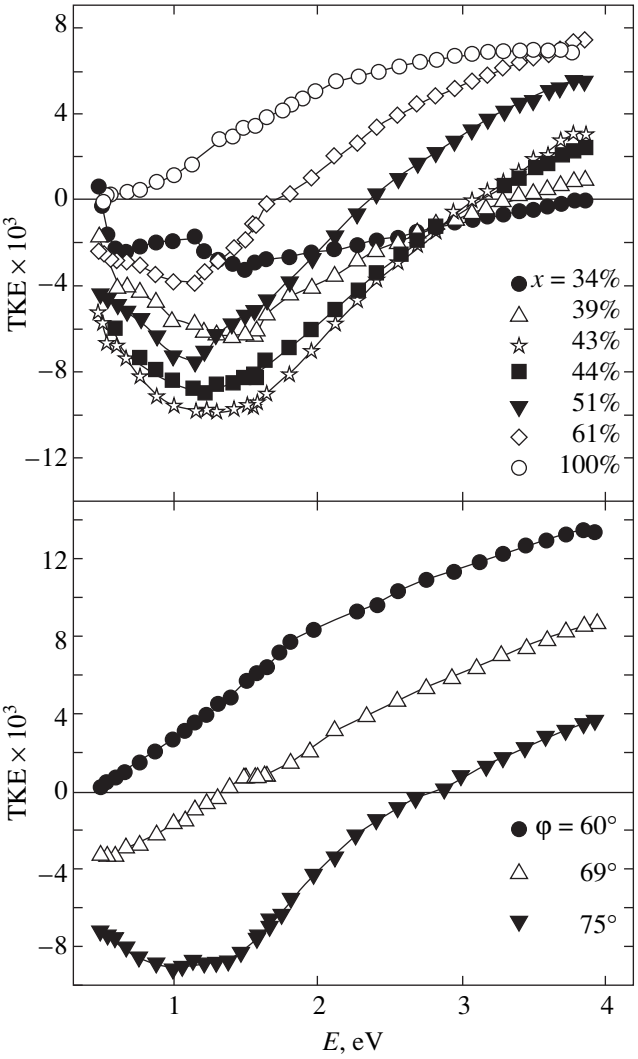


Fig. 1. The spectra of equatorial (transverse) Kerr effect (TKE) for samples of (a) the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system with various x measured for the light incidence angle $\varphi = 70°$ and (b) the $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ system with x = 47% and various φ values.

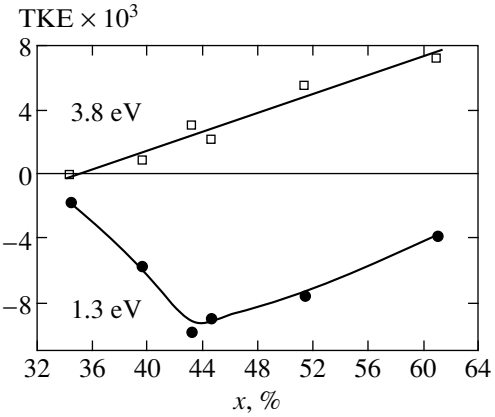


Fig. 2. Plots of the TKE magnitude versus concentration x for the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system, measured for $\varphi = 70°$ and the incident photon energies E = 1.3 and 3.8 eV.

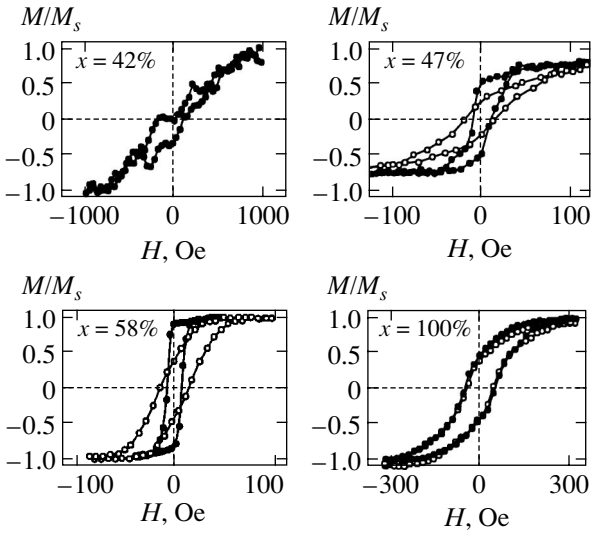


Fig. 3. Hysteresis loops measured along the (black symbols) easy and (open symbols) hard magnetization axis for samples of the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system with various concentrations x.

netic phase content grows further, both H_c and H_s increase, but the magnetic anisotropy of the system drops.

Thus, based on the analysis of the field dependence of TKE, we may conclude that the metal phase in nanocomposites with small x is concentrated in isolated granules of small size and the system occurs in an isotropic superparamagnetic state. As the concentration x increases, the metal granules grow along a certain direction occurring in the film plane and eventually forms an anisotropic ferromagnetic phase. The value of x corresponding to the appearance of the ferromagnetic phase correlates with the value of x_{per} determined from the concentration dependence of the electric conductivity. Further increase in the metal phase content and the growth of granules lead to a decrease in the magnetic anisotropy. The existence of a preferred direction for the growth of granules is probably related to some features in the technology of film deposition, primarily, to the application of a constant magnetic field during the ion beam sputtering of targets [18].

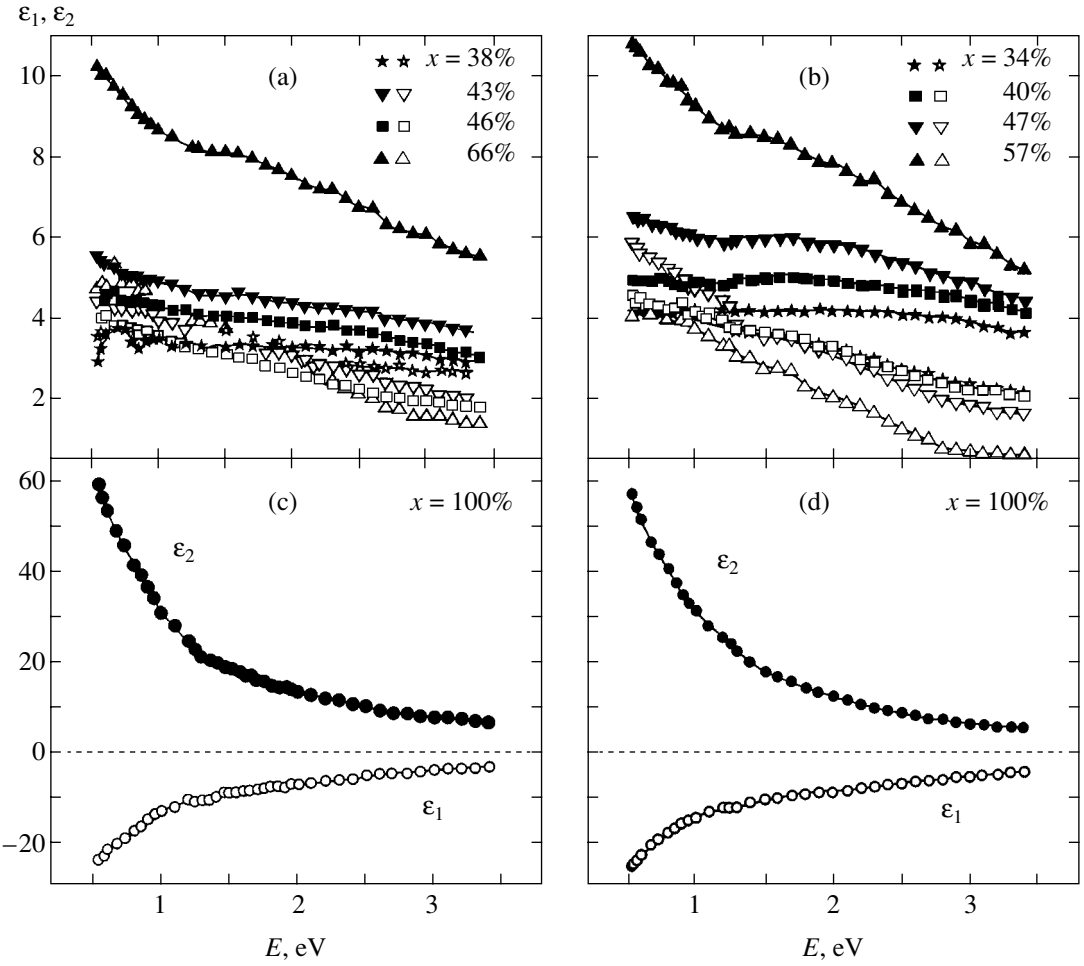


Fig. 4. Spectra of the (open symbols) real part ε_1 and (black symbols) imaginary part ε_2 of the diagonal component of the permittivity tensor for alloys of the (a, c) $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ and (b, d) $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ systems with various x.

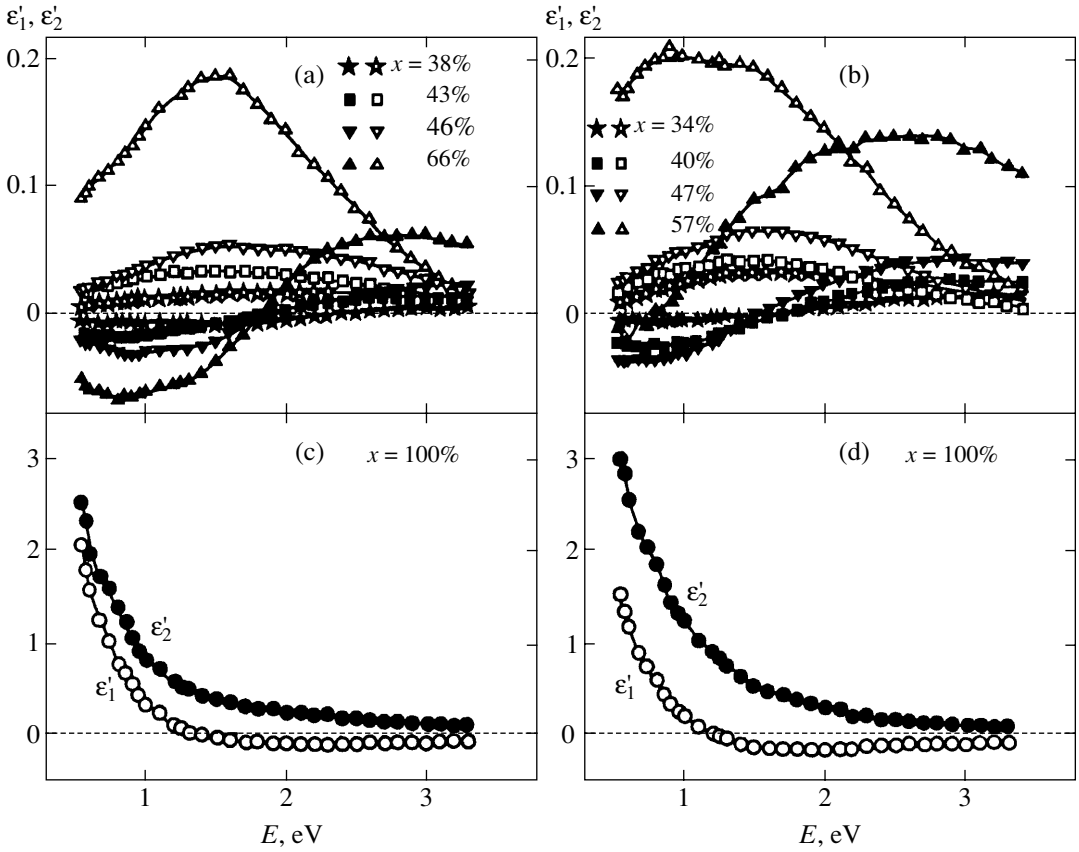


Fig. 5. Spectra of the (open symbols) real part ε_1' and (black symbols) imaginary part ε_2' of the nondiagonal component of the permittivity tensor for alloys of the (a, c) $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ and (b, d) $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ systems with various x.

Using the experimentally determined optical constants and MO data, it is possible to calculate the diagonal ($\varepsilon = \varepsilon_1 - i\varepsilon_2$) and nondiagonal ($\varepsilon' = \varepsilon_1' - i\varepsilon_2'$) components of the permittivity tensor of nanocomposites. In the case of a gyroelectric medium magnetized along the z axis, this tensor appears as

$$\hat{\varepsilon} = \begin{pmatrix} \varepsilon & i\varepsilon' & 0 \\ -i\varepsilon' & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{pmatrix}. \quad (1)$$

The corresponding TKE value is given by the formula [20]

$$\delta = (a\varepsilon_1' + b\varepsilon_2')\frac{\sin 2\varphi}{a^2 + b^2}, \quad (2)$$

where φ is the light incidence angle, $a = \varepsilon_2(2\varepsilon_1\cos^2\varphi - 1)$, $b = \cos^2\varphi(\varepsilon_2^2 - \varepsilon_1^2 + 1) + \varepsilon_1 - 1$, $\varepsilon_1 = n^2 - k^2$, $\varepsilon_2 = 2nk$, n is the refraction coefficient, and k is the absorption coefficient. Using the TKE values determined for two incidence angles and the known values on n and k, we can solve the last equation and determine the components of the permittivity tensor.

As might have been expected based on comparison of the MO spectra of composites and the corresponding homogeneous alloys, the components of the permittivity tensor of granular alloys are significantly different from the analogous values for the homogeneous samples. The optical spectra presented in Fig. 4 show an increase in the imaginary part ε_2 of the diagonal component of the permittivity tensor of nanocomposites with increasing x, this growth being most pronounced in the near-IR spectral range featuring a large contribution due to the conduction electrons. This behavior is typical of all granular nanocomposites. However, the values of ε_2 for both $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ and $(Co_{41}Fe_{39}Zr_{10})_x(SiO_2)_{100-x}$ systems remain significantly smaller as compared to those for the corresponding homogeneous amorphous alloy. In comparison to the case of homogeneous amorphous alloys, the real part ε_1 of the diagonal component of the permittivity tensor of nanocomposites changes sign and remains smaller by absolute value (in the near-IR range, by a factor of about 5).

Figure 5 shows the dispersion relations of the nondiagonal components ε_1' and ε_2' ($\varepsilon' = \varepsilon_1' - i\varepsilon_2'$) of the

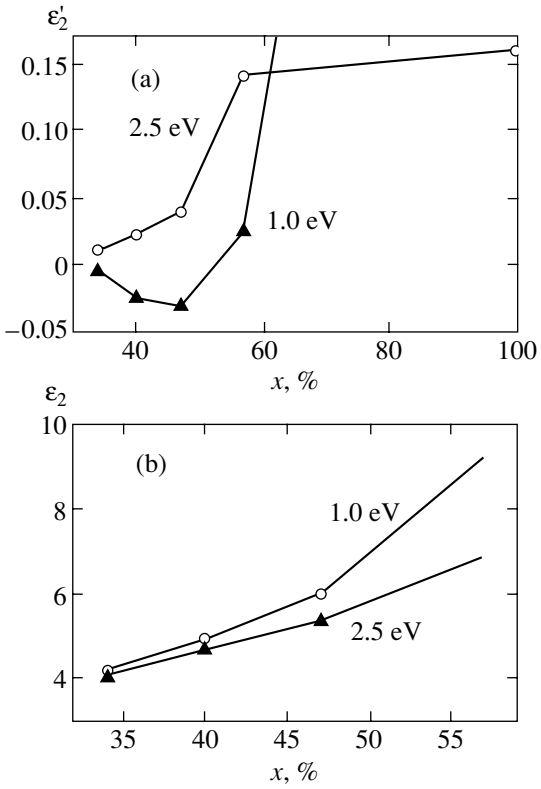


Fig. 6. Plots of the imaginary parts of the (a) nondiagonal ε_2' and (b) diagonal ε_2 components of the permittivity tensor versus concentration x for alloys of the $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ system measured for the photon energies E = 1.0 and 2.5 eV.

permittivity tensor of various nanocomposites. As can be seen, the real part ε_1' for the nanocomposites of both systems increases with decreasing photon energy; for $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$, this value exhibits a maximum at $E_{fet} \sim 1.5$ eV. In the same spectral region, the imaginary part of the nondiagonal component ε_2' of the permittivity tensor crosses the zero level. As the photon energy is further decreased, the ε_1' value exhibits a monotonic decrease, whereas ε_2' increases. This behavior of the nondiagonal components of the permittivity tensor is typical of all nanocomposites with the metal component concentration $x < x_{per}$.

For the samples of $(Co_{45}Fe_{45}Zr_{10})_{57}(SiO_2)_{43}$ (x = 57% > x_{per}), E_{fet} sharply changes to 1 eV. Figure 6 presents the concentration dependences of ε_2 and ε_2' measured for several selected energies of incident photons. As can be seen from these data, the variation of x influences the MO properties (ε_2') much more significantly than the optical properties (ε_2). The most noticeable feature in the behavior of ε_2' is observed for E = 1 eV at x = 44%, that is, near the percolation threshold. This behavior is probably related to a sharp change in microstructure of the granular alloy in the vicinity of the percolation transition which is most strongly manifested in the MO spectra in the near-IR spectral region and is expected to reveal the contribution of dimensional effects.

It should be noted that, in contrast to the TKE spectra, the spectra of the nondiagonal component of the permittivity tensor responsible for the MO activity show no increase in comparison to the corresponding homogeneous amorphous alloy. Moreover, the ε_2' value in the "red" spectral region for $Co_{45}Fe_{45}Zr_{10}$ is more than ten times that for the $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ nanocomposite with x = 43%, which showed the maximum effect in this very energy range.

4. DISCUSSION OF RESULTS

In order to interpret the experimental results, we have modeled the TKE curves within the framework of the Bruggeman approximation [13, 14] and the symmetrized Maxwell–Garnett approximation [15, 16].

According to the Bruggeman theory, the MO spectra can be calculated proceeding from the optical and MO parameters of the alloy components, the volume filling coefficient f for the magnetic particles, and their shape. The calculation algorithm is as follows [8]. In the linear approximation with respect to magnetization, the effective permittivity tensor of a granular ferromagnetic alloy can be written as

$$\hat{\varepsilon} = \begin{pmatrix} \varepsilon^{EMA} & i\varepsilon'^{EMA} & 0 \\ -i\varepsilon'^{EMA} & \varepsilon^{EMA} & 0 \\ 0 & 0 & \varepsilon^{EMA} \end{pmatrix}. \quad (3)$$

Under the assumption of weak fields and ellipsoidal particles, expressions for the diagonal and nondiagonal parts of the permittivity tensor within the framework of the Bruggeman approximation appear as

$$f(\varepsilon^1 - \varepsilon^{EMA})/(\varepsilon^{EMA} + (\varepsilon^1 - \varepsilon^{EMA})L_j) + (1-f)(\varepsilon^0 - \varepsilon^{EMA})/(\varepsilon^{EMA} + (\varepsilon^0 - \varepsilon^{EMA})L_j) = 0, \quad (4)$$

$$f(\varepsilon'^{EMA} - \varepsilon'^1)/[\varepsilon^1 + (\varepsilon^1 - \varepsilon^{EMA})L_j]^2 + (1-f)\varepsilon'^{EMA}/[\varepsilon^{EMA} + (\varepsilon^0 - \varepsilon^{EMA})L_j]^2 = 0. \quad (5)$$

Here, $\varepsilon^1 = \varepsilon_1^1 - i\varepsilon_2^1$ and $\varepsilon'^1 = \varepsilon_1'^1 - i\varepsilon_2'^1$ are the diagonal and nondiagonal parts of the permittivity tensor of the magnetic component, respectively, and $\varepsilon^0 = \varepsilon_1^0 - i\varepsilon_2^0$ is the permittivity of the nonmagnetic component. Solving these equations, one can readily calculate the TKE using relation (2).

The Bruggeman approximation does not take into account the existence of separate granules in the

medium, which hinders the description of alloys with microstructures possessing complicated topologies.

This disadvantage is eliminated in the Maxwell–Garnett approximation which, as well as its symmetrized variant [12, 17]], allows for the presence of separate particles. The symmetrized Maxwell–Garnett approximation, in contrast to the usual one, considers the medium as composed on the particles of two types. One type (A) comprises a component of material 1 inside material 2, whereas another type (B) represents component 2 inside material 1. In our case, material 1 is the metal phase and material 2 is the insulating matrix (SiO_2). The probability of encountering particles of each type in the alloy is described by the Sheng model [21],

$$P_A = u_1/(u_1 + u_2),$$

$$P_B = u_2/(u_1 + u_2),$$

where

$$u_1 = (1 - f^{1/3})^3,$$

$$u_2 = [1 - (1 - f)^{1/3}]^3.$$

The effective tensor of permittivity for such a microstructure is calculated in two stages. In the first step, the effective permittivity tensor is calculated for the particles of each type within the framework of the Maxwell–Garnett approximation. The first step, owing to the special features of this approximation, reflects the presence of separate, clearly defined granules. Then, the effective medium is composed of the particles of types A and B and the MO spectra are calculated using the algorithm based on the Bruggeman theory.

Advantages of the symmetrized Maxwell–Garnett approximation are (i) the possibility of calculating spectra for arbitrary values of the total content of the metal component with allowance for the microstructure and (ii) the accurate description of the percolation transition. These factors are very important for calculations of the properties of systems of the metal–insulator type.

Our attempt to describe the experimental results within the framework of the Bruggeman approximation did not give satisfactory results. The calculated curves did not fit the spectral dependences of TKE obtained for the samples studied (Fig. 7). This is related to the fact that the simple Bruggeman theory considers aggregate rather than granular medium and does not take into account the existence of separate particles in the nanocomposite. In order to allow for the influence of microstructure on the optical and MO properties of samples, we have modeled their behavior within the framework of the symmetrized Maxwell–Garnett approximation. Calculations of the permittivity tensor components of an effective medium in this approximation involve the volume filling coefficient f. This value differs from the total content x of the magnetic component because of the difference between the densities of the matrix and

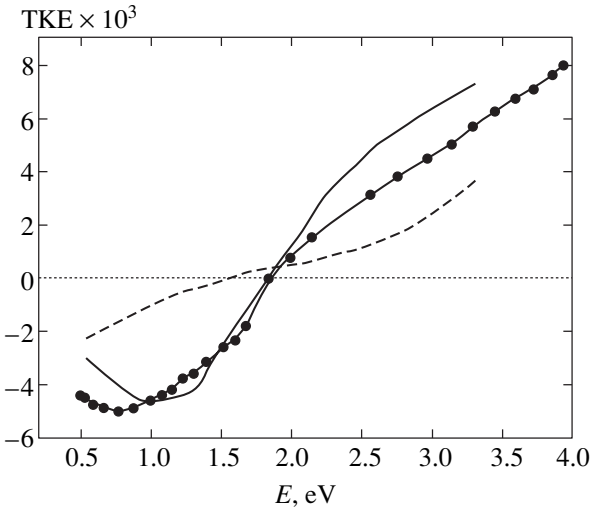


Fig. 7. The TKE spectra of a $(Co_{45}Fe_{45}Zr_{10})_{47}(SiO_2)_{53}$ nanocomposite for φ = 69°: (●) experiment; (dashed curve) Bruggeman effective medium approximation for x = 44% and L = 0.33; (solid curve) symmetrized Maxwell–Garnett approximation for x = 44%, L_A = 0.23, and L_B = 0.43.

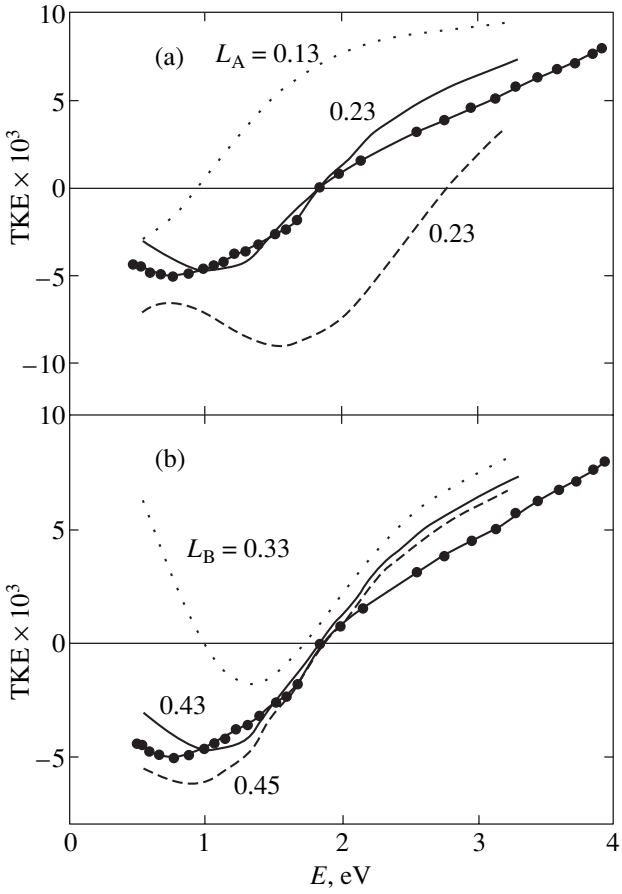


Fig. 8. Comparison of the experimental TKE spectra (black symbols) of a $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ sample with the results of modeling using the symmetrized Maxwell–Garnett approximation for various (a) L_A and (b) L_B form factors for x = 44%, L_B = 0.43 (a) and L_A = 0.23 (b). Black circles correspond to experiment.

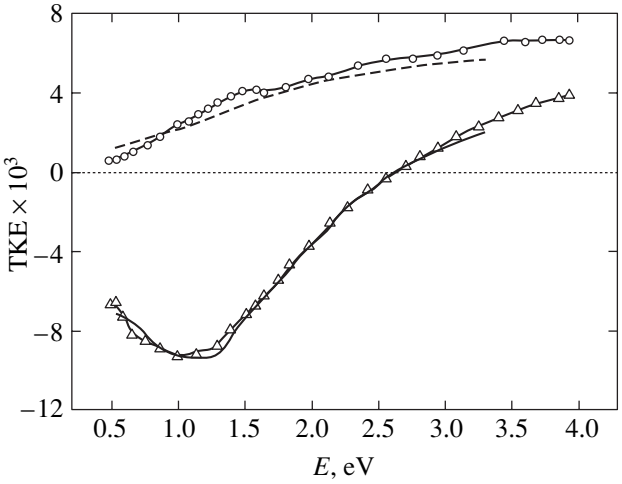


Fig. 9. Comparison of the experimental TKE spectra of $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ samples with small x and the results of modeling using the symmetrized Maxwell–Garnett approximation: (dashed line) $x = 35\%$, $L_A = 0.25$, $L_B = 0.43$, $\alpha = 0.25$; (○) experiment, $x = 34\%$, $\varphi = 60°$; (solid line) $x = 42\%$, $L_A = 0.18$, $L_B = 0.25$, $\alpha = 0.5$; (△) experiment, $x = 40\%$, $\varphi = 75°$.

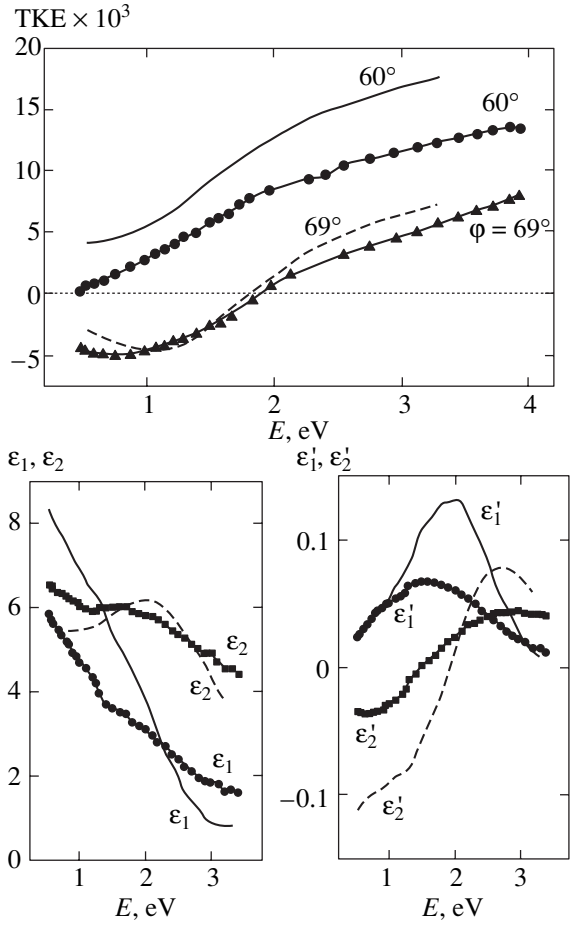


Fig. 10. Comparison of the experimental (black symbols) optical and TKE spectra of a $(Co_{45}Fe_{45}Zr_{10})_{47}(SiO_2)_{53}$ sample with the results of modeling (solid and dashed curves) using symmetrized Maxwell–Garnett approximation for $x = 44\%$, $L_A = 0.23$, and $L_B = 0.43$.

the dispersed component. We recalculated x into f using the formulas

$$f = \frac{1}{1 + V_b/V_a}, \tag{6}$$

$$\frac{V_b}{V_a} = \frac{1-x}{x}\beta, \tag{7}$$

where $\beta = \rho_a P_b/\rho_b P_a$; ρ_a, P_a and ρ_b, P_b are the density and molecular weight, respectively, of the corresponding component. For alloys of the $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ and $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ systems, we have $\beta = 1.12$ and 1.22, respectively.

Selection of the model parameters was first performed for the MO spectra measured at a light incidence angle of 70°, because the MO effects are more sensitive than optical spectra with respect to the form factor of particles [8].

Figures 8–11 present the results of calculations performed within the framework of the Maxwell–Garnett theory and compare the model curves to the experimental TKE curves. As can be seen, variation of the form factors (L_A and L_B) for the particles of both types significantly influences the shapes of the model spectra. For example (see Fig. 8a), a change in the form factor L_A only by 0.1 toward any side from the optimum value adequately describing the experimental data leads to a significant deviation of model curves from the measured spectrum in the entire spectral range. Analogous variations of the shape of type B particles (form factor L_B) also strongly change the shape of the model MO spectra, especially in the near-IR range (Fig. 8b). Thus, only the L_A and L_B values from a very narrow interval provide for a good (i.e., both qualitative and quantitative) agreement between theory and experiment. Therefore, based on the results of modeling, it is possible to judge on the shape of particles in a given composite. In the aforementioned example, a sample with the composition $(Co_{45}Fe_{45}Zr_{10})_{47}(SiO_2)_{53}$ ($x = 47\%$) comprises particles of type A with a form factor of 0.43 (almost spherical, slightly oblate in the field direction) and type B with $L_B = 0.23$ (close to ellipsoids elongated in the field direction).

It should also be noted that, for description of the TKE spectra of samples with a low content of the metal component (superparamagnets) within the Maxwell–Garnett approximation, an additional coefficient α was introduced in calculations of the nondiagonal components of permittivity. This parameter, called the undermagnetization factor, took into account that the samples with a low content of the magnetic phase are far from saturation. Using the results of modeling, it is possible to estimate the degree of undermagnetization during the TKE measurements. For example, in $(Co_{45}Fe_{45}Zr_{10})_{47}(SiO_2)_{53}$ samples with $x = 40\%$, the undermagnetization to saturation was 50% (see Fig. 9), while a sample with $x = 34\%$ had $\alpha = 0.25$.

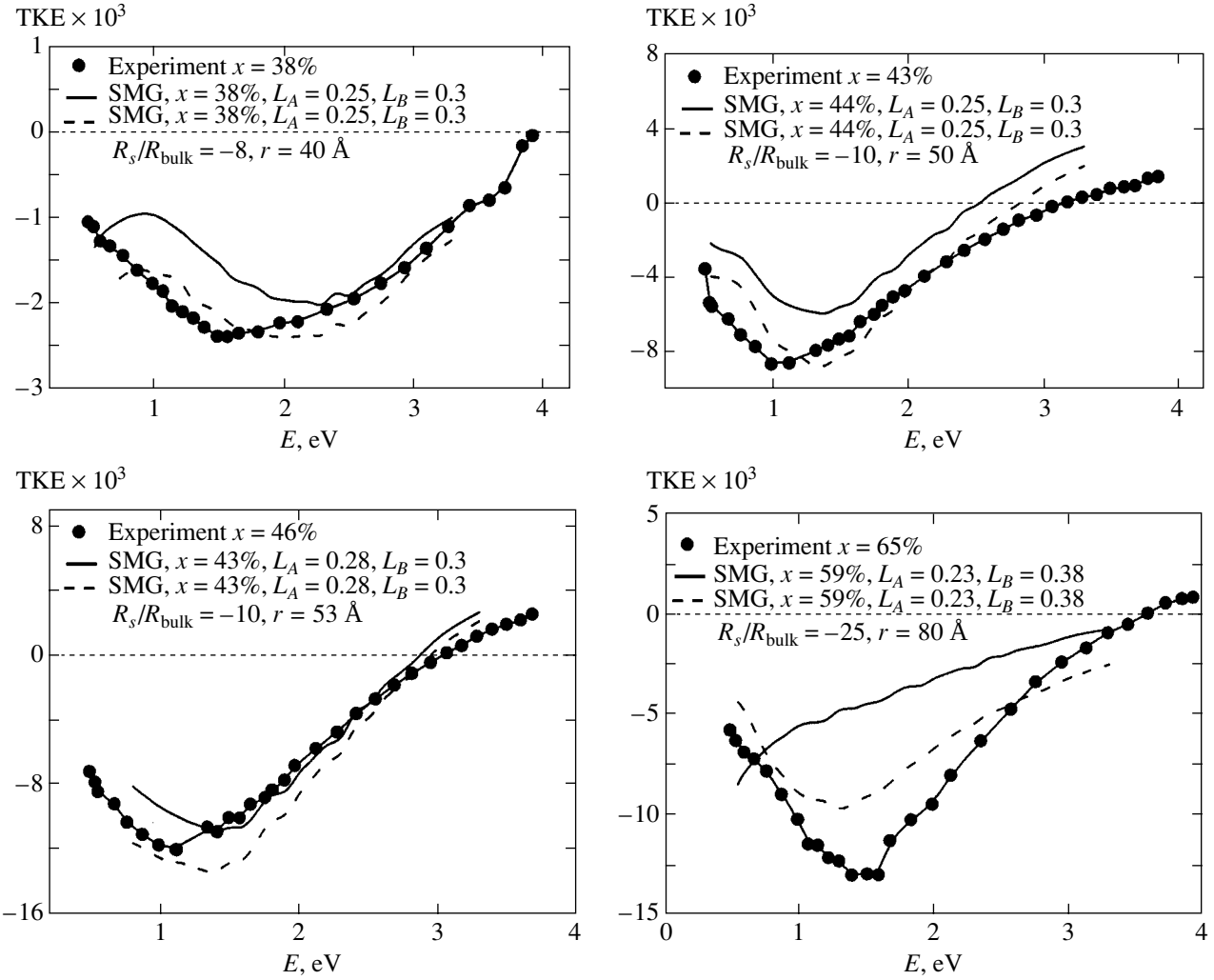


Fig. 11. Comparison of the experimental TKE spectra (black circles correspond to experiment) of $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ samples with various x and the results of modeling (solid and dashed curves) using the symmetrized Maxwell–Garnett approximation with allowance for the quasi-classical size effect.

A comparison of the results of modeling for the permittivity tensor and the TKE spectra (Fig. 10) shows that, even for a good quantitative coincidence of the measured TKE spectra with calculated curves, there is only a semiquantitative agreement between theory and experiment for both diagonal and nondiagonal components of the permittivity tensor. This discrepancy is probably related to the more complicated real microstructure of the nanocomposites studied and the presence of a certain distribution of the grains with respect to shape and size.

It should be noted that theoretical modeling for the samples of $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ nanocomposites was performed without taking into account the quasi-classical dimensional effect. However, allowance for this effect (reflecting a change in the mean free time of electron motion in the grains as compared to that in the bulk samples) improves the agreement between theory and experiment [22]. This is illustrated in Fig. 11 showing the model TKE spectra of samples of the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system. These plots present the results of calculations involving parameters of the quasi-classical dimensional effect such as the ratio of coefficients of the anomalous Hall effect for the surface of granules and a massive sample (R_s/R_{bulk}) and the size of metal inclusions (r). Making allowance for the quasi-classical dimensional effect is described in detail elsewhere [22]. As can be seen from Fig. 11, allowance for this effect changes the TKE spectrum even for the samples with a high metal content (i.e., with large granules) from that characteristic of a massive $Co_{41}Fe_{39}B_{20}$ sample toward the spectra observed for granular nanocomposites of the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system. Note also that the size of granules used in making allowance for the quasi-classical dimensional effect increases with decreasing insulator content, in good agreement with the structural data.

5. CONCLUSIONS

Based on the results of investigations, we may conclude that magnetic nanocomposites of the amorphous metal–insulator type near the percolation threshold exhibit a significant increase in the MO response in the near-IR spectral range. A change in the microstructure and topology of granular alloys influences their MO spectra to a greater extent than the optical characteristics.

The concentration dependences of the nondiagonal components of the permittivity tensor are nonlinear, the most pronounced variations being observed near the percolation threshold. This fact allows the x_{per} value to be determined by MO methods.

The observed increase in the MO response in the vicinity of the percolation threshold for the photon energies in the region of 1.3 eV for the $(Co_{41}Fe_{39}B_{20})_x(SiO_2)_{100-x}$ system and 1.0 eV for the $(Co_{45}Fe_{45}Zr_{10})_x(SiO_2)_{100-x}$ system is caused by a change in the optical and MO parameters related to variations in the topology and microstructure of the nanocomposites, rather than by an increase in the MO activity in the near-IR range.

The results of theoretical modeling of the spectra of TKE and the permittivity tensor components in the effective medium approximation showed that the best agreement between experimental and theoretical curves can be obtained using the symmetrized Maxwell–Garnett approximation, which allows variation of the microstructure in the granular alloys to be traced and the parameters of microstructure to be evaluated. Allowance for the quasi-classical dimensional effect also improves the quality of description of the experimental results.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 03-02-16127, 03-02-06301, and 03-02-17164) and the Korean Science and Engineering Foundation (grant no. 033-A39-002).

REFERENCES

1. H. Fujimori, S. Mitani, and S. Ohnuma, Mater. Sci. Eng. B **31**, 219 (1995).
2. A. Pakhomov, X. Yan, and Y. Xu, J. Appl. Phys. **79**, 6140 (1996).
3. E. Gan'shina, A. Granovsky, B. Dieny, *et al.*, Physica B (Amsterdam) **299**, 260 (2001).
4. H. Akinaga, M. Mizuguchi, T. Manado, *et al.*, J. Magn. Magn. Mater. **242–245**, 470 (2002).
5. I. V. Bykov, E. A. Gan'shina, A. B. Granovskiĭ, and V. S. Gushchin, Fiz. Tverd. Tela (St. Petersburg) **42**, 487 (2000) [Phys. Solid State **42**, 498 (2000)].
6. A. B. Granovskiĭ, I. V. Bykov, E. A. Gan'shina, *et al.*, Zh. Éksp. Teor. Fiz. **123**, 1256 (2003) [JETP **96**, 1104 (2003)].
7. E. Gan'shina, A. Granovsky, V. Gushin, *et al.*, Physica A (Amsterdam) **241**, 45 (1997).
8. E. Gan'shina, R. Kumaritova, A. Bogorodisky, *et al.*, J. Magn. Soc. Jpn. **23**, 379 (1999).
9. A. B. Granovsky, E. A. Gan'shina, A. N. Vinogradov, *et al.*, Phys. Met. Metallogr., Suppl. 1 **92**, S52 (2001).
10. V. G. Kravets, A. K. Petford-Long, and A. F. Kravets, J. Appl. Phys. **87**, 1762 (2000).
11. T. V. Murzina, E. A. Gan'shina, V. Guschin, *et al.*, Appl. Phys. Lett. **73**, 3769 (1998).
12. Yu. I. Petrov, *Physics of Small Particles* (Nauka, Moscow, 1982).
13. D. A. G. Bruggeman, Ann. Physik (Leipzig) **24**, 636 (1935).
14. T. K. Xia, P. M. Hui, and D. S. Stroud, Appl. Phys. **67**, 2736 (1989).
15. G. A. Niklasson and C. G. Granqvist, J. Appl. Phys. **55**, 3382 (1984).
16. A. Granovsky, M. Kuzmichov, and J. P. Clerc, J. Magn. Soc. Jpn. **23**, 382 (1999).
17. J. C. M. Garnett, Philos. Trans. Roy. Soc. London **203**, 385 (1904).
18. Yu. E. Kalinin, A. T. Ponomarenko, A. V. Sitnikov, and O. V. Stogneĭ, Fiz. Khim. Obrab. Mater. **5**, 14 (2001).
19. J. R. Beattie and G. K. T. Kohn, Philos. Mag. **460**, 222 (1955).
20. G. S. Krinchik, *Physics of Magnetic Phenomena* (Mosk. Gos. Univ., Moscow, 1985).
21. P. Sheng, Phys. Rev. Lett. **45**, 60 (1980).
22. A. B. Granovskiĭ, M. V. Kuz'michev, and A. N. Yurasov, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron. **6**, 67 (2000).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 1037–1039.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1184–1187.

SOLIDS
Electronic Properties

De Haas–van Alphen Oscillations in the Organic Quasi-Two-Dimensional Metal $(ET)_8[Hg_4Cl_{12}(C_6H_5Cl)_2]$

R. B. Lyubovskiĭ[a,b], S. I. Pesotskiĭ[a,b,c], V. I. Nizhankovskiĭ[b], W. Biberacher[c], and R. N. Lyubovskaya[a]

[a]*Institute of Problems of Chemical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
[b]*International Laboratory of High Magnetic Fields and Low Temperatures, 53-421, Wroc̆law, Poland*
[c]*Walter Meissner Institut, D-85748, Garching, Germany*
e-mail: pesot@icp.ac.ru
Received October 16, 2003

Abstract—The behavior of de Haas–van Alphen oscillations in the quasi-2D organic metal $(ET)_8[Hg_4Cl_{12}(C_6H_5Cl)_2]$ was studied in detail. The section of the Fermi surface of this metal is a two-dimensional network of magnetic breakdown orbits. Only two frequencies, which corresponded to allowed closed orbits, F_A and F_{MB}, were detected. This is in agreement with the earlier studies of Shubnikov–de Haas oscillations in this metal. The reason for the absence of other allowed frequencies remains unclear. The angular dependences of the amplitudes of F_A and F_{MB} oscillations contain a series of "spin zeros." An analysis of their positions led us to suggest that many-particle interactions were weakened in $(ET)_8[Hg_4Cl_{12}(C_6H_5Cl)_2]$. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The crystalline and band structures of the organic quasi-2D conductor $(ET)_8[Hg_4Cl_{12}(C_6H_5Cl)_2]$ [further, (Cl, Cl)], which has metallic properties in the temperature range 0.5–300 K, have been described in detail in [1, 2]. Band calculations show that the initial Fermi surface in the conducting plane of this metal consists of two intersecting open hole orbits characteristic of low-dimensional electronic systems. The hybridization of molecular orbitals that form the conduction band results in the formation of the final Fermi surface in the form of two closed orbits, electronic (*E*) and hole (*H*) (see Fig. 1), which have different shapes but enclose equal areas of about 13% of the area of the first Brillouin zone [2]. Even the earliest studies of Shubnikov–de Haas oscillations [3, 4] showed the presence of six rather than one, as expected, frequencies. Depending on the angle θ between the field direction and the normal to the conduction plane, all these frequencies are described by the law $F_i(\theta) = F_i(0)/\cos\theta$ characteristic of cylindrical Fermi surfaces, which are typical of quasi-2D electronic systems. Among the observed oscillations, those with a frequency $F_A(0) = 245$ T obviously predominate. This frequency corresponds to approximately 11% of the first Brillouin zone area. These oscillations were assigned to the motion of carriers in the classical closed orbits *E* and *H*. An analysis of the temperature and angular dependences of the amplitude of Shubnikov–de Haas oscillations of frequency F_A showed that the hole and electronic orbits were characterized by not only equal areas but also carriers of equal masses. The nature of the other frequencies remained unclear for several years, primarily because of the small oscillation amplitudes and related experimental problems.

Several years ago, high-quality (Cl, Cl) single crystals were synthesized and studied in quasi-pulsed fields of up to 35 T. The results [5, 6] not only included the frequencies reported earlier but even increased the spectrum of Shubnikov–de Haas oscillation frequen-

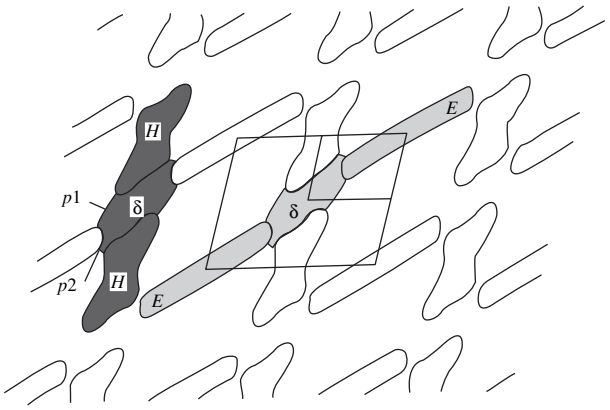


Fig. 1. Fermi surface and the first Brillouin zone in the conduction plane of the organic quasi-2D metal $(ET)_8[Hg_4Cl_{12}(C_6H_5Cl)_2]$.

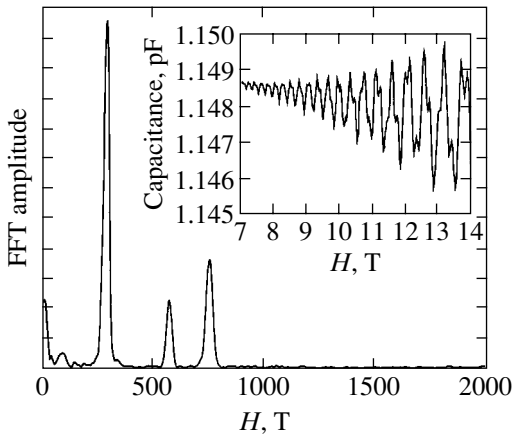


Fig. 2. Fast Fourier transform (FFT) of de Haas–van Alphen oscillations in the organic metal $(ET)_8[Hg_4Cl_{12}(C_6H_5Cl)_2]$ (the oscillations are shown in the inset). Temperature T = 0.45 K and polar angle $\theta = 32.5°$.

cies to fifteen at certain field directions. An analysis of these results led the authors of [6] to the following conclusions:

(1) Magnetic breakdown transitions $p1$ and $p2$ between the hole and electronic Fermi surface regions become possible in magnetic fields on the order of 10 T (Fig. 1). This results in the formation of a 2D network of magnetic breakdown orbits.

(2) Both closed magnetic breakdown orbits and open orbits, which are related to each other by magnetic breakdown transitions and correspond to the quantum interference effect [7], exist simultaneously for all frequencies other than the F_A frequency that corresponds to two initial closed orbits E and H.

(3) Two of the frequencies observed experimentally, $F_A(0) = 242$ T and $F_{MB}(0) = 633$ T, are only responsible for the motion of carriers in closed orbits [F_{MB} corresponds to two closed orbits enclosing equal areas, $H + \delta + H$ and $E + \delta + E$ (see Fig. 1), including four magnetic breakdown transitions each]. In any event, this motion significantly predominates in F_{MB} oscillations.

(4) The other frequencies make up two groups. The first group includes the frequencies (in particular, the frequency corresponding to 100% of the first Brillouin zone area) that largely originate from the quantum interference effect. The frequencies of the second group (in particular, the frequency corresponding to area δ between the initial Fermi surface sheets) have not been given satisfactory explanation. In principle, they can be described within the framework of both oscillating [8] and fixed [9] electrochemical potentials, but neither of these models has been substantiated experimentally.

This work presents the results of a detailed study of de Haas–van Alphen oscillations in (Cl, Cl). These results lend support to the versions suggested in [5, 6].

2. RESULTS AND DISCUSSION

The samples for measurements were high-quality (Cl, Cl) single crystals of weights up to 110 μg. The de Haas–van Alphen oscillations were studied by monitoring changes in the torque moment of a capacitive sensor [10]. A magnetic field of up to 14 T was generated by a superconducting magnet. The temperatures that we used were in the range 0.45–1.3 K.

De Haas–van Alphen oscillations were studied at angles $\theta = (-21°–80°)$. Their frequencies depended on θ according to the standard equation $F_i(\theta) = F_i(0)/\cos\theta$. At all magnetic field directions, the number of observed frequencies did not exceed three (see Fig. 2), namely, $F_A(0) \approx 242$ T, $2F_A$ (the second harmonic of the fundamental frequency), and $F_{MB}(0) = 630$ T. The fundamental frequency F_A corresponds to charge motions in the classical E and H orbits. Effective mass calculations for the fundamental frequency showed that equal masses, which amounted to $m_A = (1.18 \pm 0.05)m_0$ in the conduction plane (m_0 is the mass of the free electron), corresponded to both orbits. The F_{MB} frequency corresponds to the motion of carriers in two closed magnetic breakdown orbits with equal areas [$H + \delta + H$ and $E + \delta + E$ (see Fig. 1)]. The effective masses in these orbits were also equal, $m_{MB} = (2.28 \pm 0.05)m_0$ in the conduction plane. According to the effective mass concept [11], we can expect that the mass related to magnetic breakdown orbits should be twice the m_A mass for the main closed orbits. This mass ratio is satisfied to within measurement errors, $2m_A = 2.36m_0 \sim 2.28m_0 = m_{MB}$.

The angular dependences of the amplitude of de Haas–van Alphen oscillations are shown in Fig. 3. Figure 3a corresponds to oscillations at frequency F_A, and Fig. 3b, to oscillations at frequency F_{MB}. Both dependences contain minima. Curve *a* has two minima at 30.1° and 59.5°, and curve *b*, three minima at 20.6°, 47.4°, and 60.7°. These minima are called spin zeros and originate from the splitting of Landau levels in a magnetic field [12]. The condition of the existence of such zeros is

$$\cos(\pi pmg/2m_0) = 0,$$

where p is the harmonic number and g is the g-factor. The presence of two or more sequential spin zeros and the assumption that the effective mass depends on the angle by the $m(\theta) = m(0)/\cos\theta$ law typical of cylindrical Fermi surfaces allow us to unambiguously calculate the splitting factor

$$S = gm(0)/2m_0 = ((2n + 1)\cos\theta_n)/2,$$

where θ_n is the position of the nth spin zero and $n = 0, 1, 2, 3, \ldots$. The splitting factor for oscillations with the F_A frequency is $S_A = 1.29 \pm 0.04$; for the F_{MB} frequency, it is $S_{MB} = 2.36 \pm 0.04$. The ratio between the effective masses $m_A/m_{MB} = 0.52$ is approximately equal to the

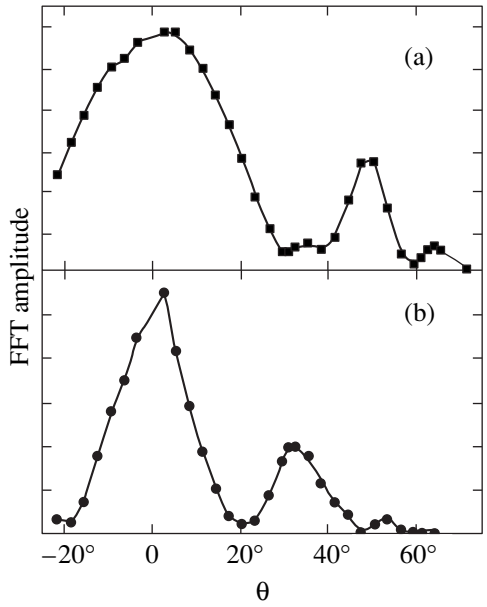


Fig. 3. Polar angle θ dependences of the amplitude of de Haas–van Alphen oscillations: (a) oscillations with frequency F_A and (b) oscillations with frequency F_{MB}; T = 0.45 K.

ratio between the splitting factors $S_A/S_{MB} = 0.54$, which leads us to suggest that the splittings of the Landau levels of the closed orbits under consideration are characterized by equal g-factors. This is a quite expected result, considering how these orbits are formed [2]. The insignificant difference of the splitting factors from the corresponding reduced masses (1.29 and 1.18 or 2.36 and 2.28) can be indirect evidence in favor of an insignificant contribution of many-particle interactions in (Cl, Cl), as has been noted earlier for some organic metals [13].

One more point should be mentioned. The effective mass for the F_A frequency, $m_A = 1.18m_0$, determined from de Haas–van Alphen oscillations closely agrees with that obtained in magnetoresistive measurements, $m_A = 1.17m_0$ [6]. At the same time, the effective masses for the magnetic breakdown orbits are noticeably different, $m_{MB} = 2.28m_0$ and $1.95m_0$ for de Haas–van Alphen and Shubnikov–de Haas [6] oscillations, respectively. This difference can be caused by the existence of an interferometer with a frequency equal to F_{MB} but a substantially lower mass in addition to the magnetic breakdown frequency F_{MB} [6]. This interferometer makes no contribution to de Haas–van Alphen oscillations, but its contribution to resistive oscillations can be noticeable and cause effective mass underestimation.

The question of the absence of other frequencies for which closed magnetic breakdown orbits exist in the observed spectrum of de Haas–van Alphen oscillations remains open. Jointly studying magnetoresistance and magnetization oscillations in higher magnetic fields would probably clarify the situation.

3. CONCLUSIONS

Two de Haas–van Alphen oscillation frequencies (not counting harmonics) are only observed for the organic metal $(ET)_8[Hg_4Cl_{12}(C_6H_5Cl)_2]$, at least, in fields of up to 14 T. One of these corresponds to two classical closed orbits enclosing equal areas, and the other, to two magnetic breakdown orbits also enclosing equal areas. The behavior of the amplitudes of these oscillations and the corresponding effective masses is in reasonable agreement with the suggestion that the Fermi surface of $(ET)_8[Hg_4Cl_{12}(C_6H_5Cl)_2]$ is formed through the hybridization of the orbitals that form the conduction band. The results of this work are in close agreement with those obtained in studying Shubnikov–de Haas oscillations in pulsed and stationary fields. The reason for the absence of other frequencies related to allowed closed magnetic breakdown orbits is unclear, and the problem requires additional study.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic research (project no. 03-02-16606) and State Contract no. 40.020.1.1.1166.

REFERENCES

1. R. N. Lyubovskaya, O. N. Dyachenko, V. V. Gritsenko, *et al.*, Synth. Met. **42**, 1907 (1991).
2. L. F. Vieros and E. Canadell, J. Phys. I France **4**, 939 (1994).
3. R. B. Lyubovskii, S. I. Pesotskii, A. Gilevskii, and R. N. Lyubovskaia, J. Phys. I France **6**, 1809 (1996).
4. R. B. Lyubovskii, S. I. Pesotskii, A. Gilevskii, and R. N. Lyubovskaya, Zh. Éksp. Teor. Fiz. **107**, 1698 (1995) [JETP **80**, 946 (1995)].
5. R. B. Lyubovskii, S. I. Pesotskii, C. Proust, *et al.*, Synth. Met. **113**, 227 (2000).
6. C. Proust, A. Audouard, L. Brossard, *et al.*, Phys. Rev. B **65**, 155106 (2002).
7. R. W. Stark and C. B. Friedberg, J. Low Temp. Phys. **1**, 111 (1974).
8. J. Y. Fortin and T. Ziman, Phys. Rev. Lett. **80**, 3117 (1998).
9. V. M. Gvozdikov, Yu. V. Pershin, E. Steep, *et al.*, Phys. Rev. B **65**, 165102 (2002).
10. P. Christ, W. Biberacher, H. Muller, and K. Andres, Solid State Commun. **91**, 451 (1994).
11. L. M. Falicov and H. Stachowiak, Phys. Rev. **147**, 505 (1966).
12. D. Shoenberg, *Magnetic Oscillations in Metals* (Cambridge Univ. Press, Cambridge, 1984; Mir, Moscow, 1986).
13. S. I. Pesotskiĭ, R. B. Lyubovskiĭ, W. Biberacher, *et al.*, Zh. Éksp. Teor. Fiz. **121**, 504 (2002) [JETP **94**, 431 (2002)].

Translated by A. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 1040–1044.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1188–1193.

SOLIDS
Electronic Properties

Anomalous Angular Dependence of Critical Current in Niobium Films

Yu. V. Vishnyakov, A. N. Lykov, and A. Yu. Tsvetkov

Lebedev Institute of Physics, Russian Academy of Sciences, Leninskiĭ pr. 53, Moscow, 119991 Russia
e-mail: lykov@mail1.lebedev.ru

Received September 5, 2003

Abstract—An anomalous angular dependence of the critical current is observed in niobium films. This phenomenon manifests itself in the fact that, under small intensities of the external magnetic field, the critical current attains its maximum in a slightly tilted magnetic field. It is found that the position of the maximum depends on the external magnetic field, as well as on the initial conditions under which the samples were kept. A theoretical model is proposed to explain the results obtained. This model takes into account the effect of diamagnetic properties, pinning of vortices, and the initial conditions on the vortex system in Nb films. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Niobium occupies a special position in superconductivity because the Ginzburg–Landau parameter in it is close to unity. In a magnetic field, niobium exhibits the characteristic features of both type-I and type-II superconductors. Therefore, structures based on such a superconductor may have unusual properties. For example, an anomalous angular dependence of the critical current in Nb-based layered structures was observed in [1], where the role of effective pinning centers was played by insulating NbO_x layers. In [2], it was found that the dependence of critical current on the magnitude and the direction of the external magnetic field is varied along with the initial conditions under which these layered structures were kept. In other words, by changing the initial position of a sample, one can vary the angular dependence of the critical current in such structures by rotating them in a magnetic field, and even change the position of the maximum on this dependence. This fact enables one to control the critical parameters of superconductors during the experiment. To understand the nature of the phenomena observed in [1, 2] in greater detail, we carried out precision angular measurements of the critical current in Nb films.

2. METHOD OF MEASUREMENTS

The samples investigated in this work are narrow superconducting strips fabricated from niobium films. The films were produced by a dc magnetron sputtering technique [3]. As the substrates, we used polycrystalline sapphire with a surface polished to a class of at least 12. The temperature of the substrates during the film deposition was about 350°C. Using the photolithography technique, we fabricated narrow strips with bonding pads from these films. Contact wires were bonded to these pads by indium. The critical temperature T_c of these films was 9.2 K and the thickness d was 0.6 μm. The width w of the strips was 10 μm, the length L was about 100 μm, and the ratio of the room-temperature resistance to the resistance at T = 10 K was $R_k/R_{10} = 2.9$.

A magnetic field with an intensity of up to 65 kOe was produced by a superconducting solenoid. The critical current I_c is determined from the current–voltage characteristic as the current at which the voltage drop on a sample reaches 1 μV. The current–voltage characteristics are measured by the four-point probe technique. The detection of the critical current in our experiment was fully automated. This allowed us to determine the voltage drop to an accuracy of 0.1 μV. A sample was rotated with respect to the direction of the external magnetic field by a precision worm-and-worm gear, made of stainless steel, with a ratio of 30. This fact provided an angular resolution of 0.04° for the sample orientation. The rotating gear with a sample was placed in a uniform magnetic field produced by a superconducting solenoid, so that the sample may rotate with respect to the magnetic field. All of the equipment is placed in a helium bath at a temperature of 4.2 K. In our experiments, the current I transmitted through a sample was always directed perpendicular to the direction of the magnetic field.

3. RESULTS OF MEASUREMENTS

Figure 1 shows an example of the angular dependence of the critical current in niobium films. One can see that the critical current attains its maximum when a sample is tilted to the magnetic field and the magnitude of the shift of the peak position increases as the magnetic-field intensity decreases. In the present case, the

sample was initially situated so that its plane made an angle of $\theta = \theta_0 \sim 30°$ with the direction of the magnetic field. Then, the sample was rotated in the counterclockwise direction, when θ decreased up to the angle $\theta = -\theta_0 \sim -30°$, i.e., until the sample occupied a position symmetric with respect to the initial position. Then, the sample was rotated in the opposite, clockwise, direction until it reached the original position. In Fig.1, the curves of $I_c(\theta)$ are plotted for various values of the external magnetic field H = 1, 1.5, 2, 3, and 4 kOe. The curves marked by closed symbols correspond to the clockwise rotation (in the direction of increasing θ, the positive direction), while the curves marked by light symbols correspond to the rotation in the opposite direction (in the counterclockwise, or negative direction). The solid curves correspond to the approximation of the experimental results by the Lorentz function

$$y = y_0 + \frac{2A}{\pi}\frac{w}{4(x - x_0)^2 + w^2},$$

where y_0, A, w, x_0 are fitting parameters. This approximation is used here to determine the exact position of the maximum of the function $I_c(\theta)$. To distinguish between "clockwise" and "counterclockwise" functions, we denote them by $I_c(\theta_{cl})$ and $I_c(\theta_{acl})$, respectively. Introduce the notation $\Delta\theta = \theta_{cl}^m - \theta_{acl}^m$, where θ_{cl}^m is the angle corresponding to the maximum of $I_c(\theta_{cl}^m)$ and θ_{acl}^m is the angle corresponding to the maximum of $I_c(\theta_{acl}^m)$. The positions of the clockwise and counterclockwise maxima do not coincide; moreover, these maxima are different for different values of the external magnetic field. However, if a sample is brought into a normal state before each measurement $I_c(\theta_{cl})$ or $I_c(\theta_{acl})$ by increasing the magnetic-field intensity, then these maxima will coincide and will be observed in a parallel external magnetic field. An example of the dependence obtained in this way is shown in Fig. 1 (the curve marked by crosses for an external field of H = 1 kOe). The difference between the clockwise and counterclockwise maxima as a function of the external magnetic field, $\Delta\theta(H)$, is shown in Fig. 2. Note that the amplitude of the maximum does not depend on the direction of rotation of a sample, and the maximum is always observed before the sample reaches the parallel orientation with respect to the external magnetic field. The position of the maximum is independent of the direction of the transport current.

We found that the position of the maximum depends on the initial conditions under which a sample was kept before the measurement. The difference between the maxima as a function of θ_0—the initial deviation of the sample from the parallel position—is shown in Fig. 3. In this experiment, we measure the clockwise and counterclockwise functions $I_c(\theta)$, which are similar to those shown in Fig. 1, except that we now reduce the initial deviation θ_0 from measurement to measurement, thus changing the initial conditions for each subsequent measurement. Here, we should note the specific feature

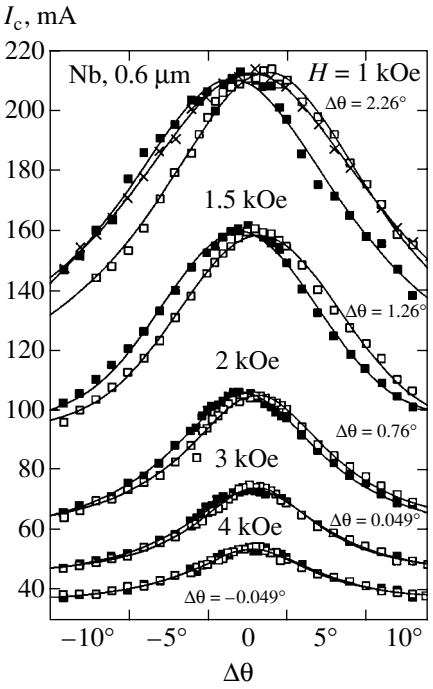


Fig. 1. Critical current through a sample as a function of the angle between the surface of the structure and the direction of the external magnetic field for various values of the field H. Dark symbols correspond to the angular dependence recorded when the sample is rotated in the clockwise direction. Light symbols represent the experimental results recorded when the sample is rotated in the counterclockwise direction.

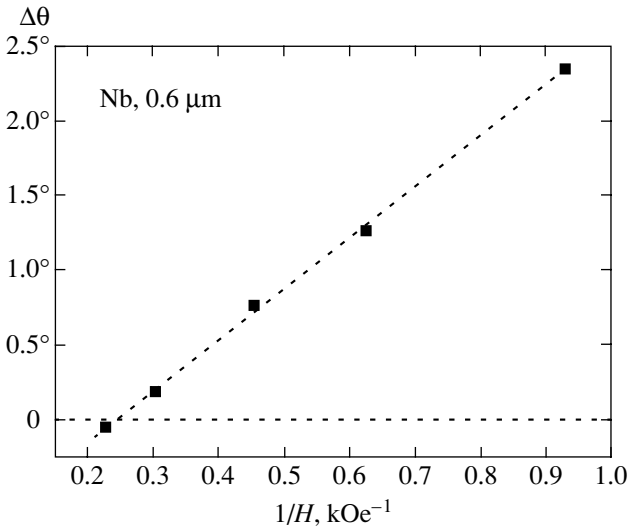


Fig. 2. The angular difference $\Delta\theta$ between the clockwise and counterclockwise peaks as a function of the inverse of the magnetic field intensity.

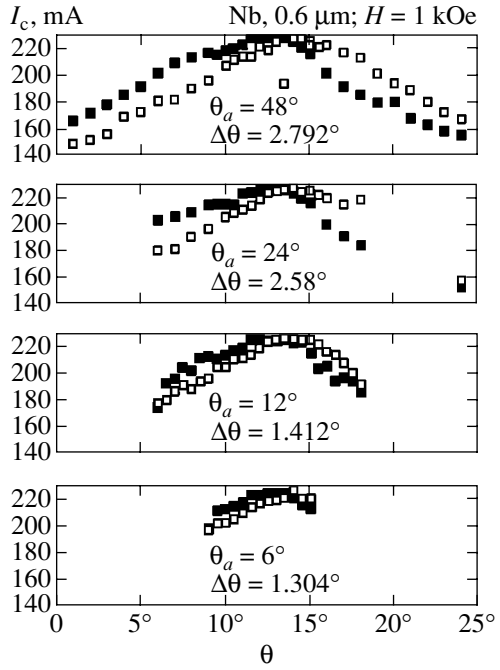


Fig. 3. Angular dependence of the critical current for various values of θ_a.

of the method for measuring these functions, namely, that the initial point of each subsequent measurement lies on the opposite curve of the preceding measurement. Figure 3 shows that $\Delta\theta$ decreases as the sweeping amplitude $\theta_a = 2\theta_0$ decreases. Thus, the positions of the clockwise and counterclockwise maxima depend on the initial conditions (memory effect) of a sample, as well as on the way in which it was brought to this condition. This result is qualitatively analogous to the result obtained on Nb/NbO$_x$ multilayer structures [1, 2].

4. THEORETICAL MODEL

To explain the results, we apply an approach developed in [2–6]. If the external magnetic field is parallel to the film surface, then the flux lines penetrating the sample are parallel to this surface; however, in a slightly tilted magnetic field, the flux lines penetrating the sample may have different shapes due to the presence of effective pinning centers in the films, which may have a considerable effect on the configuration of these lines. Moreover, under the rotation of a film with respect to the external magnetic field, the effective pinning centers entrain the vortices and thus considerably affect the new state. Therefore, the system under investigation partly remembers its previous state. As we mentioned above, niobium, which is the basic element of our samples, occupies a special position in superconductivity because its Ginzburg–Landau parameter is close to unity. This gives rise to diamagnetic properties that are inherent in type-I superconductors.

Taking into account the aforesaid, we write the following equation for the Gibbs free energy density [4]:

$$G = \frac{1}{8\pi}B(B-4\pi M) - \frac{1}{4\pi}HB, \tag{1}$$

where B is the magnetic flux density inside the film, H is the intensity of the external magnetic field, and M is magnetization that determines film's memory of the previous state and the presence of diamagnetic properties exhibited by the structure under investigation. Here, we do not take into account the demagnetization phenomenon because a sample is already magnetized and we do not analyze the process of its magnetization. Moreover, for the parameters d = 0.6 μm, w = 10 μm, and w/d = 16.7 of the samples, the demagnetization factor is close to zero when a sample is oriented parallel to the magnetic field. In a coordinate system that is rigidly bound to a sample so that the z axis is perpendicular to the plane of the sample and the y axis is parallel to the direction of transport current, we can rewrite Eq. (1) as

$$\begin{aligned} G &= \frac{1}{8\pi}(B_x^2 + B_z^2) - \frac{1}{2}B_xM_x \\ -\frac{1}{2}B_zM_z &= \frac{H}{4\pi}(B_x\cos\theta + B_z\sin\theta). \end{aligned} \tag{2}$$

Minimizing Eq. (2), we obtain the following equations for the components B_x and B_z of the magnetic flux density:

$$B_x = H\cos\theta + 2\pi M_x, \tag{3}$$

$$B_z = H\sin\theta + 2\pi M_z, \tag{4}$$

$$B^2 = B_x^2 + B_z^2. \tag{5}$$

The density of the volume pinning force in niobium films placed in a perpendicular magnetic field is usually given by

$$F_p \sim B(1 - B/H_{c2}). \tag{6}$$

When a transport current flows through the film under investigation, the vortex lines are subject to the Lorentz force

$$F_L \approx \frac{1}{c}BJ, \tag{7}$$

where c is the speed of light in vacuum and J is a current through the film. This force tends to depin the vortices from irregularities. The current density at which a vortex is depinned from an irregularity is the critical current density, which corresponds to the transition from a dissipation-free flow of transport current through the sample to a dissipation current. Taking into account that the condition $F_p = F_L$ must hold at the transition point, we can obtain the following expression for the critical

current density for a film placed in a perpendicular magnetic field:

$$J_c = J_{c0}(1 - B/H_{c2}). \quad (8)$$

When deriving an expression for the critical current in a slightly tilted magnetic field, we assume that the interaction between a vortex and a surface that prevents the former from penetrating into the film, the so-called surface pinning, is much greater than the pinning forces in the plane of the film. Therefore, in this case, the critical current is determined by the Lorentz force in the plane of the film. Using Eqs. (3)–(5) and assuming that $M_x = 0$ when the film is parallel to the magnetic field, we can rewrite Eq. (8) as follows:

$$J_c = J_{c0}(1 - K(|\sin\theta - \sin\Delta\theta_t|)^{0.5}), \quad (9)$$

where

$$\sin\Delta\theta_t = -\frac{H^2 + 4\pi^2 M_z^2}{4\pi M_z H}, \quad K = \frac{\sqrt{4\pi M_z H}}{H_{c2}}.$$

Formula (9) represents the critical current density as a function of the angle θ for $\theta \longrightarrow 0$ under the assumption that M_z is independent of θ.

Formula (9) determines the critical current density that is associated with the pinning in the plane of a film when the interaction between vortices is weak and the concentration of pinning centers is high. One can see that the critical current attains its maximum at $\theta = \Delta\theta_t$ rather than at $\theta = 0$. In this case, the magnetic field frozen into the film as if neutralizes due to the diamagnetic effect, and the flux lines are directed along the film surface. Thus, in this case, the critical current is mainly determined by the surface pinning—the interaction between vortices and the film surface. This leads to the boundedness of function (9) for $\theta \sim \Delta\theta_t$. A number of studies have been devoted to the calculation of the surface pinning [7, 8]. However, for our purposes, it is sufficient that this quantity is independent of angle for small deviations of a sample from its parallel orientation in the magnetic field.

The main result obtained in the present paper is the difference between the clockwise and counterclockwise maxima on the angular dependence of the critical current for small external magnetic fields. Therefore, we first consider how this result agrees with the model described above. According to (9), $J_c(\theta)$ attains its maximum at $\theta = \Delta\theta_t > 0$, i.e., before the structure reaches a parallel position. Let us set $\Delta\theta_t = \theta_c$ (where θ_c is the characteristic angular size of a sample, which is determined by the relation $\theta_c = \arctan(D/w)$, $\theta_t = 3.43°$) in Eq. (9) and compare the theoretical function $J_c(\theta)$ with the experimental one. In this case, the parameters J_{c0} and K are determined from the normalization conditions by the experimental dependence. The theoretical dependence is used for the clockwise rotation, while the dependence symmetric to it is used for the counterclockwise rotation. The result obtained is shown in Fig. 4. One can see that the experimental and theoretical curves are in good qualitative agreement. Hence, we can conclude that the positions of the maxima on the experimental dependence $J_c(\theta)$ are indeed determined by the magnetic moment $M_z(H)$ in a sample; the maximal value of this moment is directly related to θ_c, which represents a geometrical characteristic of the sample under investigation. If we associate it only to the existence of frozen vortices in the sample, then $M_z(H)$ should be positive and the peak on the experimental dependence must appear after the sample passes the parallel position. At this angle, the external magnetic field should compensate for the magnetic moment of frozen vortices, and the magnetic field will be parallel to the sample surface. However, $M_z(H) < 0$ in our case because the peak on the experimental angular dependence leads the parallel position. Such behavior of the magnetization can be explained by invoking the diamagnetic properties of niobium. Note that, in the framework of the present approach, one can easily account for the variation in the position of the maximum of $I_c(\theta)$ under the variation of the direction of rotation of a sample because this variation is associated with the variation in the magnitude and the direction of magnetization of a sample.

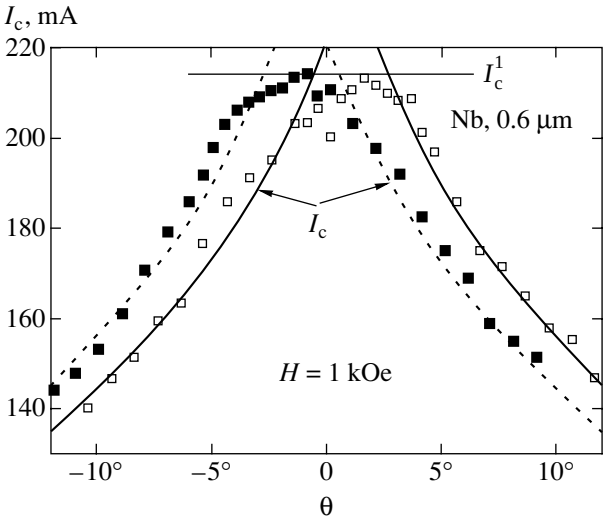


Fig. 4. Experimental and theoretical angular dependence of the critical current. The solid and dashed lines correspond to the theoretical dependence.

Let us return to the experimental dependence shown in Fig. 3. As we noted above, the difference $\Delta\theta$ between the clockwise and counterclockwise peaks decreases as the angle θ_0 that specifies the initial position of a sample with respect to the external magnetic field decreases. Our model also takes this result into account. It is obvious that the magnetization M_z is a function not only of the intensity of the external magnetic field but also of the angle θ_0 – $M_z = M_z(H, \theta_0)$; i.e., it depends on

the initial conditions of the experimental problem. This fact can easily be explained within our model because M_z depends on the magnitude of the frozen magnetic flux. As the sweep amplitude decreases, the magnitude of the flux decreases because a sample approaches a position parallel to the external magnetic field. According to the aforesaid, we can state that the experimental results obtained in this work are in satisfactory agreement with the model proposed.

5. CONCLUSIONS

The anomalous angular dependence of the critical current in niobium films has been observed. This anomaly manifests itself in the fact that, for small intensity of the external magnetic field, the critical current attains its maximum in a slightly tilted magnetic field, and is associated with the complex character of the transition of a vortex system to the unperturbed ground state. In standard measurements of the critical current, the vortex system is in an excited metastable state. Therefore, investigations of the critical current are complicated because theoretical studies deal with the equilibrium properties of the vortex system in superconductors. We can point out the following three basic results of this paper.

1. We have developed a new approach to the study of nonequilibrium properties of the vortex system in superconductors.

2. The anomalous angular dependence of critical current observed in this study should be observed in all superconductors. Here, one should take into account that, in type-II superconductors with a large Ginzburg–Landau parameter, diamagnetic properties are suppressed in strong magnetic fields, and the aforementioned phenomenon should be small.

3. The results obtained show that the conventional method for determining a parallel orientation of a sample in a magnetic field, which is based on finding the maximum of the angular dependence of critical current, may give erroneous results.

ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Industry, Science, and Technology of the Russian Federation (project nos. 40.012.1.1.1355 and 40.012.1.1.1357) and by the Russian Foundation for Basic Research (project no. 02-02-16285).

REFERENCES

1. A. N. Lykov and Yu. V. Vishnyakov, Europhys. Lett. **36**, 625 (1996).
2. Yu. V. Vishnyakov, A. N. Lykov, and A. Yu. Tsvetkov, Preprint No. 39, FIAN (Physical Inst., Russian Academy of Sciences, Moscow, 2001).
3. E. M. Golyamina, V. I. Dedyu, A. N. Lykov, *et al.*, Sverkhprovodimost: Fiz. Khim. Tekh. **4**, 1430 (1991).
4. M. Tachiki, T. Koyama, and S. Takahashi, Physica C (Amsterdam) **185**–**189**, 303 (1991).
5. M. Tachiki and S. Takahashi, Solid State Commun. **70**, 291 (1989).
6. M. Tachiki and S. Takahashi, Solid State Commun. **72**, 1083 (1989).
7. V. V. Schmidt and G. S. Mkrtchyan, Usp. Fiz. Nauk **112**, 459 (1974) [Sov. Phys. Usp. **17**, 170 (1974)].
8. F. F. Ternovskiĭ and L. N. Shekhata, Zh. Éksp. Teor. Fiz. **62**, 2297 (1972) [Sov. Phys. JETP **35**, 1202 (1972)].

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 1045–1053.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1194–1203.

NONLINEAR PHYSICS

Suppression of Chaos in the Vicinity of a Separatrix

A. Yu. Loskutov and A. R. Dzhanoev

Moscow State University, Moscow, 119899 Russia
e-mail: loskutov@moldyn.phys.msu.ru; janoev@polly.phys.msu.ru

Received April 30, 2003

Abstract—The standard Melnikov method for analyzing the onset of chaos in the vicinity of a separatrix is used to explore the possibility of suppressing chaos of dynamical systems of a certain class. Analytical expressions are obtained for external perturbations that eliminate chaotic behavior. These results are supplemented with a numerical analysis of the Duffing–Holmes-oscillator and pendulum equations. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Intensive theoretical and experimental studies of chaotic dynamical systems revealed their unexpected and remarkable property: they are highly susceptible and extremely sensitive to perturbations. This discovery served as a starting point for finding a means to control the behavior of chaotic systems, i.e., to change from chaotic regimes to required regular oscillatory regimes by means of relatively weak perturbations.

Suppression of unstable or chaotic behavior of dynamical systems is generally achieved via stimulated excitation of stable (usually periodic) oscillations by means of multiplicative or additive perturbations. In other words, an external perturbation is required to change from a chaotic state of a system to a regular regime. The statement of the problem is outwardly simple, but its solution is very difficult to find for particular dynamical systems. Moreover, even though the problem has been analyzed in numerous studies, a systematic and rigorous theory of suppression of chaotic behavior has been developed only for some common families of dynamical systems (see [1, 2] and references therein).

Chaotic behavior can be suppressed by two different methods. In one of these, the state of a system is changed from chaotic to regular by perturbation without feedback. In other words, this method does not make use of the current values of dynamic variables. In the other method, the perturbation is adjusted in accordance with the required values of dynamic variables; i.e., feedback is an integral component of the dynamical system. By convention, the former method is called open-loop suppression (or control) of chaotic dynamics. The latter method is called feedback control of chaotic systems. Both methods can be implemented either parametrically or by direct forcing.

To the best of our knowledge, the first analyses of suppression of chaotic dynamics of certain systems were presented in [3, 4]. However, extensive research along these lines was initiated by [5, 6], where it was shown that relatively weak parametric perturbations can be used to regularize a particular saddle orbit embedded in a chaotic attractor. These and other results stimulated studies of suppression of chaotic dynamics and evoked great interest in controlling unstable systems. A vast number of numerical and experimental studies were focused on the possibility of suppression of chaos and implementation of periodic or other required dynamics in various systems and maps (see [1, 2, 7–10] and references therein).

The standard Melnikov method is an effective tool used in analytical treatments of the problem of chaos suppression [11]. It is based on comparison of the first-order terms in the series expansions of the solution in terms of a perturbation parameter on stable and unstable separatrices. In particular, the Melnikov method was applied to explore the possibility of eliminating the chaotic dynamics of the Duffing–Holmes oscillator [12–16] (see also [17]). It was shown that a small parametric perturbation of the system's chaotic dynamics suppresses chaos. Furthermore, the Melnikov method was used in [18] to examine the effects of parametric perturbations in a model of the Josephson junction.

In this paper, the Melnikov method [11, 19] is applied to find analytical expressions for parametric perturbations that suppress chaotic and/or unstable behavior of dissipative dynamical systems. The Duffing–Holmes oscillator and pendulum are considered as examples.

2. THE MELNIKOV METHOD

In this section, we briefly describe the Melnikov analytical method for identifying homoclinic or heteroclinic chaos, relying on the original paper [11] (see also [19–21]).

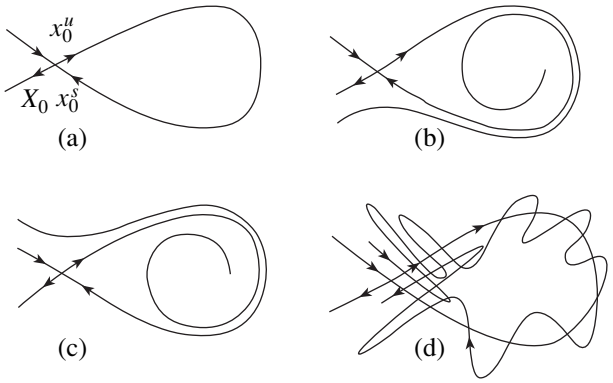


Fig. 1. Split separatrix loops.

Consider a simple autonomous system with a single hyperbolic point X_0 subject to a periodic perturbation:

$$\dot{x} = f_0(x) + \varepsilon f_1(x, t), \tag{1}$$

where $x = (x_1, x_2)$ and f_1 is a periodic function with period T. Suppose that the unperturbed system (with $\varepsilon = 0$) has a single separatrix $x_0(t)$ (see Fig. 1a):

$$\lim_{t \to \pm\infty} x_0(t) = X_0.$$

The separatrix is split by the perturbation; i.e., it has distinct incoming and outgoing branches. Three possibilities arise as a result: the separatrices either do not intersect (in which case one may enclose the other, see Figs. 1b and 1c) or intersect at an infinite number of homoclinic points. Chaotic dynamics are observed only in the latter case (see Fig. 1d).

To find an intersection condition, one must use a perturbation method to calculate the distance $D(t, t_0)$ between the separatrices at an instant t_0. If the outgoing separatrix encloses the incoming one, then $D(t, t_0) < 0$. If the incoming separatrix encloses the outgoing one, then $D(t, t_0) > 0$. Only if there exists t_0 such that the separatrices intersect does the sign of $D(t, t_0)$ alternate.

In the method substantiated in [11], the distance $D(t, t_0)$ between the branches of a split separatrix is determined by performing integration along unperturbed trajectories. The method is based on comparison of the first-order terms in the series expansions of the solution in terms of the perturbation parameter ε on stable and unstable separatrices.

To calculate $D(t, t_0)$, it is sufficient to find the solutions on the stable and unstable manifolds, x^s and x^u. When $\varepsilon = 1$, these solutions differ by the vector

$$r(t, t_0) = x^s(t, t_0) - x^u(t, t_0) = x_1^s(t, t_0) - x_1^u(t, t_0).$$

The Melnikov distance is the projection of r on the direction normal to the unperturbed separatrix x_0 at an instant t.

Omitting intermediate calculations, we write out an expression for D:

$$D(t, t_0) = -\int_{-\infty}^{\infty} f_0 \wedge f_1 dt. \tag{2}$$

This function determines conditions for chaotic behavior of the original system. In the domain where the sign of $D(t, t_0)$ alternates, the separatrices intersect and the system exhibits chaotic dynamics.

3. ELIMINATION OF CHAOTIC DYNAMICS IN THE VICINITY OF A SEPARATRIX

We use the mathematical procedure described above to explore the possibility of suppressing chaotic dynamics for systems with separatrix loops described by Eq. (1).

For this system, the Melnikov function $D(t, t_0)$ can be written as

$$D(t, t_0) = -\int_{-\infty}^{\infty} f_0 \wedge f_1 dt \equiv I[g(x, t)]. \tag{3}$$

Suppose that the sign of $D(t, t_0)$ alternates; i.e., the separatrices intersect (see Fig. 1d). We seek a perturbation $f^*(\omega, t)$ that eliminates the intersection of the separatrices:[1]

$$\dot{x} = f_0(x) + \varepsilon[f_1(x, t) + f^*(\omega, t)], \tag{4}$$

where

$$f^*(\omega, t) = (f_1^*(\omega, t), f_2^*(\omega, t)).$$

We denote by $[s_1, s_2]$ the interval where the sign of $D(t, t_0)$ alternates. Two cases can arise when the system is perturbed by $f^*(\omega, t)$:

$$D^*(t, t_0) > s_2 \tag{5}$$

or

$$D^*(t, t_0) < s_1, \tag{6}$$

where $D^*(t, t_0)$ is the Melnikov distance for system (4). Suppose that (5) is satisfied. (A similar analysis can be

[1] We tentatively call f^* a regularizing perturbation.

performed when inequality (6) holds.) Then,

$$I[g(x, t)] + I[g^*(\omega, x, t)] > s_2, \tag{7}$$

where

$$I[g^*(\omega, x, t)] = -\int_{-\infty}^{+\infty} f_0 \wedge f^* dt.$$

By virtue of (7), there exists χ such that

$$I[g(x, t)] + I[g^*(\omega, x, t)] = s_2 + \chi = \text{const},$$
$$\chi, s_2 \in \mathbb{R}^+.$$

Hence,

$$I[g^*(\omega, x, t)] = \text{const} - I[g(x, t)]. \tag{8}$$

On the other hand,

$$I[g^*(\omega, x, t)] = -\int_{-\infty}^{\infty} f_0 \wedge f^* dt. \tag{9}$$

Suppose that the function $f^*(\omega, t)$ is absolutely integrable over an infinite interval and Fourier transformable. We define $f^*(\omega, t)$ as

$$f^*(\omega, t) = \mathrm{Re}\{\hat{A}(t)e^{-i\omega t}\}$$

with $\hat{A}(t) = (A(t), A(t))$. Therefore,

$$-\int_{-\infty}^{\infty} f_0 \wedge \{\hat{A}(t)e^{-i\omega t}\} dt = \text{const} - I[g(x, t)].$$

The inverse Fourier transform yields

$$f_0 \wedge \hat{A}(t) = \int_{-\infty}^{\infty} (I[g(x, t)] - \text{const}) e^{i\omega t} d\omega.$$

Hence,

$$A(t) = \frac{1}{f_{01}(x) - f_{02}(x)} \times \int_{-\infty}^{\infty} (I[g(x, t)] - \text{const}) e^{i\omega t} d\omega.$$

The quantity $A(t)$ can be interpreted as the amplitude of a regularizing perturbation.

Thus, dynamics of systems that can be represented as (1) are regularized by the perturbation

$$f^*(\omega, t) = \mathrm{Re}\left[\frac{e^{-i\omega t}}{f_{01}(x) - f_{02}(x)} \int_{-\infty}^{\infty} (I[g(x, t)] - \text{const}) e^{i\omega t} d\omega\right].$$

Next, we explore the possibility of suppressing chaotic dynamics for systems governed by equations of the form

$$\begin{aligned} \dot{x} &= P(x, y), \\ \dot{y} &= Q(x, y) + \varepsilon[f(\omega, t) + \alpha F(x, y)], \end{aligned} \tag{10}$$

where $f(\omega, t)$ is a periodic perturbation; $P(x, y)$, $Q(x, y)$, and $F(x, y)$ are smooth functions; and α is a damping parameter.

We consider the most common case when a single hyperbolic point is located at the origin ($x = y = 0$) and $P(x, y) = y$. Let $x_0(t)$ be the solution on the separatrix. For perturbed system (10), the Melnikov distance can be represented as

$$D(t, t_0) = -\int_{-\infty}^{\infty} y_0(t - t_0) \times [f(\omega, t) + \alpha F(x_0, y_0)] dt \equiv I[g(\omega, \alpha)], \tag{11}$$

where $y_0(t) = \dot{x}_0(t)$. As in the case of Eq. (1), we assume that the sign of the Melnikov distance for system (10) alternates; i.e., the separatrices intersect. We seek a perturbation $f^*(\omega, t)$ that eliminates chaotic dynamics:

$$\begin{aligned} \dot{x} &= y, \\ \dot{y} &= Q(x, y) + \varepsilon[f(\omega, t) + \alpha F(x, y) + f^*(\omega, t)]. \end{aligned} \tag{12}$$

Since system (10) is parameterized by α, chaos must be suppressed for each particular value of the parameter. Accordingly, we can write $I[g(\omega)]$ instead of $I[g(\omega, \alpha)]$.

For system (12),

$$f_{01} = y, \quad f_{02} = Q(x, y), \quad \hat{A}(t) = (0, A(t)).$$

Therefore,

$$A(t) = \frac{1}{y_0(t - t_0)} \int_{-\infty}^{\infty} (I[g(\omega)] - \text{const}) e^{i\omega t} d\omega.$$

Thus, a regularizing perturbation for system (12) can be represented as

$$f^*(\omega, t) = \mathrm{Re}\left[\frac{e^{-i\omega t}}{y_0(t - t_0)} \int_{-\infty}^{\infty} (I[g(\omega)] - \text{const}) e^{i\omega t} d\omega\right].$$

Now, let us find a regularizing perturbation in the case when the Melnikov function $D(t, t_0)$ admits an additive shift from its critical value.

Again, we analyze the case when (5) is satisfied. Suppose that α_c corresponds to the critical value of the Melnikov function,

$$I_c = I[g(\omega, \alpha|_{\alpha = \alpha_c})].$$

Then, a subcritical Melnikov distance can be expressed as

$$I_{out} = I_c - \alpha,$$

where $a \in \mathbb{R}^+$ is a constant. Assuming that the system perturbed by $f^*(\omega, t)$ exhibits regular behavior, we have

$$I' + I_{out} + I[g^*(\omega)] > s_2, \tag{13}$$

where

$$I[g^*(\omega)] = -\int_{-\infty}^{+\infty} y_0(t - t_0) f^*(\omega, t) dt.$$

On the other hand, it is obvious that we can take any I' a fortiori greater than I_c:

$$I' = I_c + a > s_2. \tag{14}$$

Now, equating the left-hand sides of (13) and (14), we obtain $I[g^*(\omega)] = 2a$. Substituting

$$f^*(\omega, t) = \mathrm{Re}\{A(t)e^{i\omega t}\},$$

into the expression for $I[g^*(\omega)]$, we find

$$-\int_{-\infty}^{\infty} e^{i\omega t} A(t) y_0(t - t_0) dt = 2a.$$

The inverse Fourier transform yields

$$A(t) y_0(t - t_0) = -2a \int_{-\infty}^{\infty} e^{-i\omega t} d\omega.$$

Hence,

$$A(t) = -\frac{2a}{y_0(t - t_0)} \int_{-\infty}^{\infty} e^{-i\omega t} d\omega = -\frac{4\pi a \delta(t)}{y_0(t - t_0)}.$$

Thus, dynamics of systems that admit additive shift from the critical value of the Melnikov function $D(t, t_0)$ are regularized by the perturbation

$$f^*(\omega, t) = -\frac{4\pi a \delta(t)}{y_0(t - t_0)} \cos(\omega t), \tag{15}$$

where $\delta(t)$ is the Dirac delta function.

In the general case, if $f_0 = (f_{01}(x), f_{02}(x))$, then we obviously obtain

$$f^*(\omega, t) = -\frac{4\pi a \delta(t)}{f_{01}(x) - f_{02}(x)} \cos(\omega t).$$

4. APPLICATION TO PHYSICAL SYSTEMS

Now, we use the approach presented above to analyze the Duffing–Holmes-oscillator and pendulum equations. Transverse intersections of stable and unstable manifolds of these unperturbed systems give rise to homoclinic or heteroclinic orbits.

4.1. Duffing–Holmes Oscillator

The forced Duffing–Holmes oscillator with a parametrically perturbed cubic term is described by the equation

$$\ddot{x} - x + \beta[1 + \eta \cos(\Omega t)]x^3 = \varepsilon[\gamma \cos(\omega t) - \alpha \dot{x}], \tag{16}$$

where η and Ω are the amplitude and frequency of the parametric perturbation, respectively, and $\eta \ll 1$. We rewrite it as

$$\begin{aligned} \dot{x} &= v, \\ \dot{v} &= x - \beta x^3 + \varepsilon[\gamma \cos(\omega t) - \beta \eta x^3 \cos(\Omega t) - \alpha v]. \end{aligned} \tag{17}$$

The corresponding unperturbed Hamiltonian is

$$H_0 = \frac{v^2}{2} - \frac{x^2}{2} + \frac{\beta x^4}{4}.$$

Setting H_0, we find that system (17) has a single hyperbolic point ($x = v = 0$) with a single separatrix. The solution on the separatrix can be represented as [21] (see also [12–15])

$$x_0(t) = \frac{\sqrt{2}}{\sqrt{\beta}} \cosh t, \tag{18}$$

$$v_0(t) = \dot{x}_0(t) = -\frac{\sqrt{2}}{\sqrt{\beta}} \frac{\sinh t}{\cosh^2 t}. \tag{19}$$

Comparing this system with (1), we write

$$f_{01} = v, \quad f_{11} = 0,$$

$$f_{02} = x - \beta x^3,$$

$$f_{12} = \gamma \cos(\omega t) - \eta \beta x^3 \cos(\Omega t) - \alpha v.$$

Therefore,

$$f_0 \wedge f_1 = v_0[\gamma \cos(\omega t) - \eta \beta x_0^3 \cos(\Omega t) - \alpha v_0]$$

and (2) becomes

$$D(t, t_0) = -\int_{-\infty}^{+\infty} dt[\gamma v_0(t-t_0)\cos(\omega t) - \eta\beta x_0^3(t-t_0)v_0(t-t_0)\cos(\Omega t) - \alpha v_0^2(t-t_0)]. \quad (20)$$

Changing to the integration variable $\tau = t - t_0$, we finally obtain [12–15]

$$D(t, t_0) = \frac{2\sqrt{2}}{\sqrt{\beta}}\pi\gamma\omega\frac{\sin(\omega t_0)}{\cosh(\pi\omega/2)} - \frac{\pi\eta}{6\beta}(\Omega^4 + 4\Omega^2)\frac{\sin(\Omega t_0)}{\sinh(\pi\Omega/2)} + \frac{4\alpha}{3\beta}. \quad (21)$$

The sign of $D(t, t_0)$ is preserved if

$$\frac{6\beta d\sinh(\pi\Omega/2)}{\pi(\Omega^4 + 4\Omega^2)} = \eta_{\min} < \eta \le \eta_{\max} = \frac{1}{p^2}\frac{6\sqrt{2\beta}\gamma\omega}{(\Omega^4 + 4\Omega^2)}\frac{\sinh(\pi\Omega/2)}{\cosh(\pi\omega/2)}, \quad (22)$$

where p is an integer (see [12–15]). Using the left-hand inequality in (22), we determine the critical value of the Melnikov function:

$$D_c(t, t_0) = \frac{2\sqrt{2}}{\sqrt{\beta}}\frac{\pi\gamma\omega}{\cosh(\pi\omega/2)}\sin(\omega t_0) + \frac{4\alpha}{3\beta} - d\sin(\Omega t_0).$$

(An analogous calculation can be performed for the right-hand inequality.)

Then, a subcritical value

$$D_{\text{out}}(t, t_0) < D_c(t, t_0).$$

can be represented as

$$D_{\text{out}}(t, t_0) = \frac{2\sqrt{2}}{\sqrt{\beta}}\frac{\pi\gamma\omega}{\cosh(\pi\omega/2)}\sin(\omega t_0) + \frac{4\alpha}{3\beta} - d\sin(\Omega t_0) - a,$$

where $a > 0$ is a constant.

Since the perturbation required to regularize the dynamics of system (16) has the form

$$f^*(\Omega, t) = \text{Re}\{e^{i\Omega t}A(t)\},$$

the corresponding Melnikov distance

$$D^*(t, t_0) = -\int_{-\infty}^{+\infty} v_0(t-t_0)f^*(\Omega, t)dt.$$

is

$$D^*(t, t_0) = -\int_{-\infty}^{+\infty} A(t)v_0(t-t_0)e^{i\Omega t}dt. \quad (23)$$

To find $A(t)$, we define

$$D^*(t, t_0) + D_{\text{out}}(t, t_0) \equiv D'(t, t_0).$$

Since the perturbation $f^*(\Omega, t)$ is regularizing by assumption, it holds that

$$D'(t, t_0) > D_c(t, t_0).$$

It is obvious that we can take any $D'(t, t_0)$ a fortiori greater than $D_c(t, t_0)$:

$$D'(t, t_0) = \frac{2\sqrt{2}}{\sqrt{\beta}}\frac{\pi\gamma\omega}{\cosh(\pi\omega/2)}\sin(\omega t_0) + \frac{4\alpha}{3\beta} - d\sin(\Omega t_0) + a. \quad (24)$$

On the other hand, we can use (23) to write

$$D'(t, t_0) = -\int_{-\infty}^{+\infty} A(t)v_0(t-t_0)e^{-i\Omega t}dt + \frac{2\sqrt{2}}{\sqrt{\beta}}\frac{\pi\gamma\omega}{\cosh(\pi\omega/2)}\sin(\omega t_0) + \frac{4\alpha}{3\beta} - d\sin(\Omega t_0) - a. \quad (25)$$

Equating (24) to (25), we have

$$\int_{-\infty}^{+\infty} A(t)v_0(t-t_0)e^{i\Omega t}dt = -2a.$$

The inverse Fourier transform yields

$$A(t) = -\frac{2a}{v_0(t)}\int_{-\infty}^{+\infty} e^{-i\Omega t}d\Omega.$$

Therefore, dynamics of the forced Duffing–Holmes oscillator are regularized by the perturbation

$$f^*(\Omega, t) = -\frac{4\pi a\delta(t)}{v_0(t-t_0)}\cos(\Omega t). \quad (26)$$

4.2. Pendulum

The analysis presented above can be extended to the classical nonlinear pendulum, whose separatrices make up a heteroclinic orbit in the absence of damping. A periodically forced, damped pendulum is described by the equation [22]

$$\ddot{x} + \alpha\dot{x} + \sin x = \gamma\cos(\omega t). \quad (27)$$

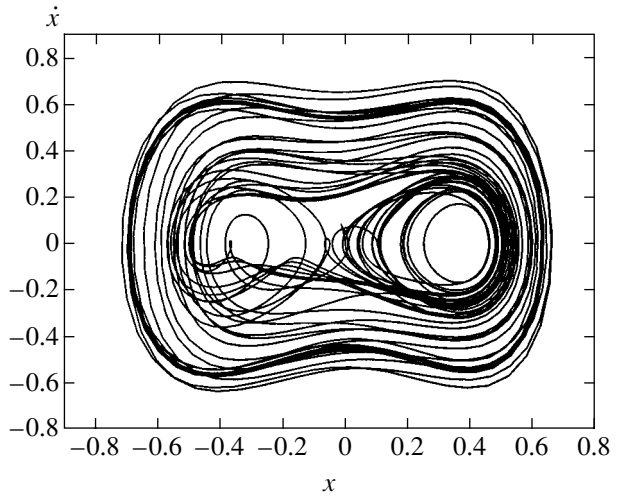


Fig. 2. Phase portrait of Duffing–Holmes oscillator (16): $\alpha = 0.145$, $\beta = 8$, $\eta = 0.03$, $\gamma = 0.14$, $\Omega = \omega = 1.1$.

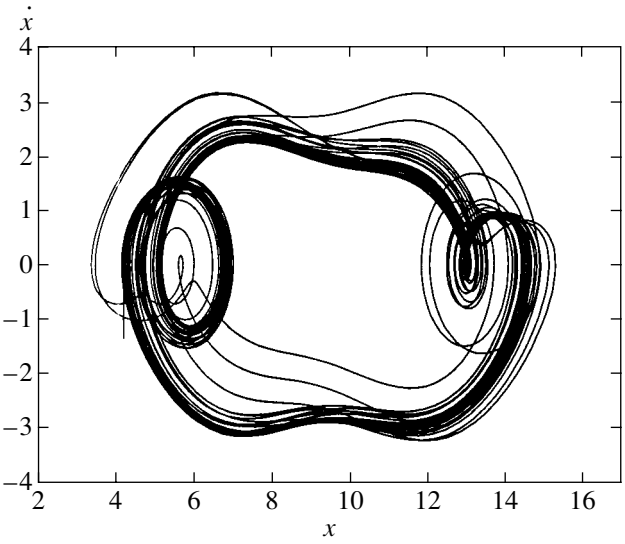


Fig. 3. Phase portrait of pendulum (27): $\alpha = 0.04$, $\gamma = 1.35$, $\omega = 1.0$.

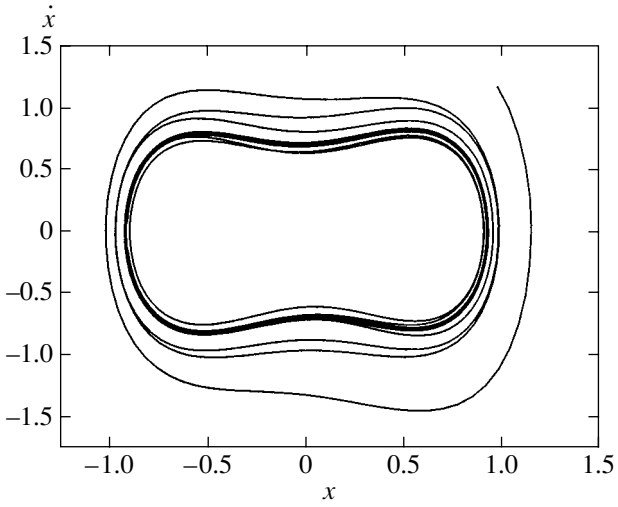


Fig. 4. Phase portrait of Duffing–Holmes oscillator (31): $\alpha = 0.145$, $\beta = 8$, $\eta = 0.03$, $\gamma = 0.14$, $\Omega = \omega = 1.1$, $a = 2$.

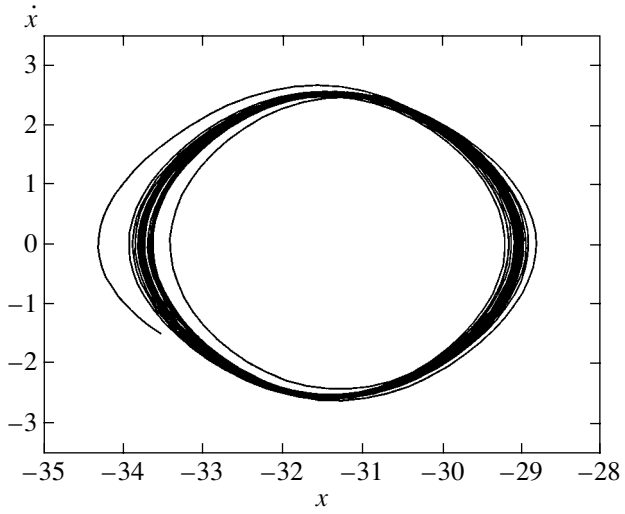


Fig. 5. Phase portrait of pendulum (32): $\alpha = 0.04$, $\gamma = 1.35$, $\omega = 1.0$, $a = 1.2$.

The corresponding unperturbed Hamiltonian is

$$H_0 = \frac{\dot{x}^2}{2} - \cos x.$$

The phase portrait of the pendulum is 2π-periodic in x, with hyperbolic points at $(\pm\pi, x)$ and a center at $(0, 0)$. The system has oscillatory, rotatory, and separatrix solutions. We focus here on solutions of the last type:

$$x_0(t) = \pm\frac{\tanh t}{\cosh t},$$

$$\dot{x}_0(t) = \pm\frac{2}{\cosh t}.$$

The Melnikov distance corresponding to (27) is [22]

$$D(t_0, \omega) = -\alpha\int_{-\infty}^{+\infty}(\dot{x}_0(t))^2 dt \pm \gamma\cos(\omega t_0)\int_{-\infty}^{+\infty}\sin(x_0(t))\dot{x}_0(t)\cos\omega t dt. \quad (28)$$

Calculating the integrals, we obtain

$$D(t_0, \omega) = -4\alpha B\left(\frac{1}{2}, 1\right) \pm \frac{2\pi\gamma}{\cosh\left(\frac{\pi\omega}{2}\right)}\cos(\omega t_0), \quad (29)$$

where $B(r, s)$ is Euler's beta function.

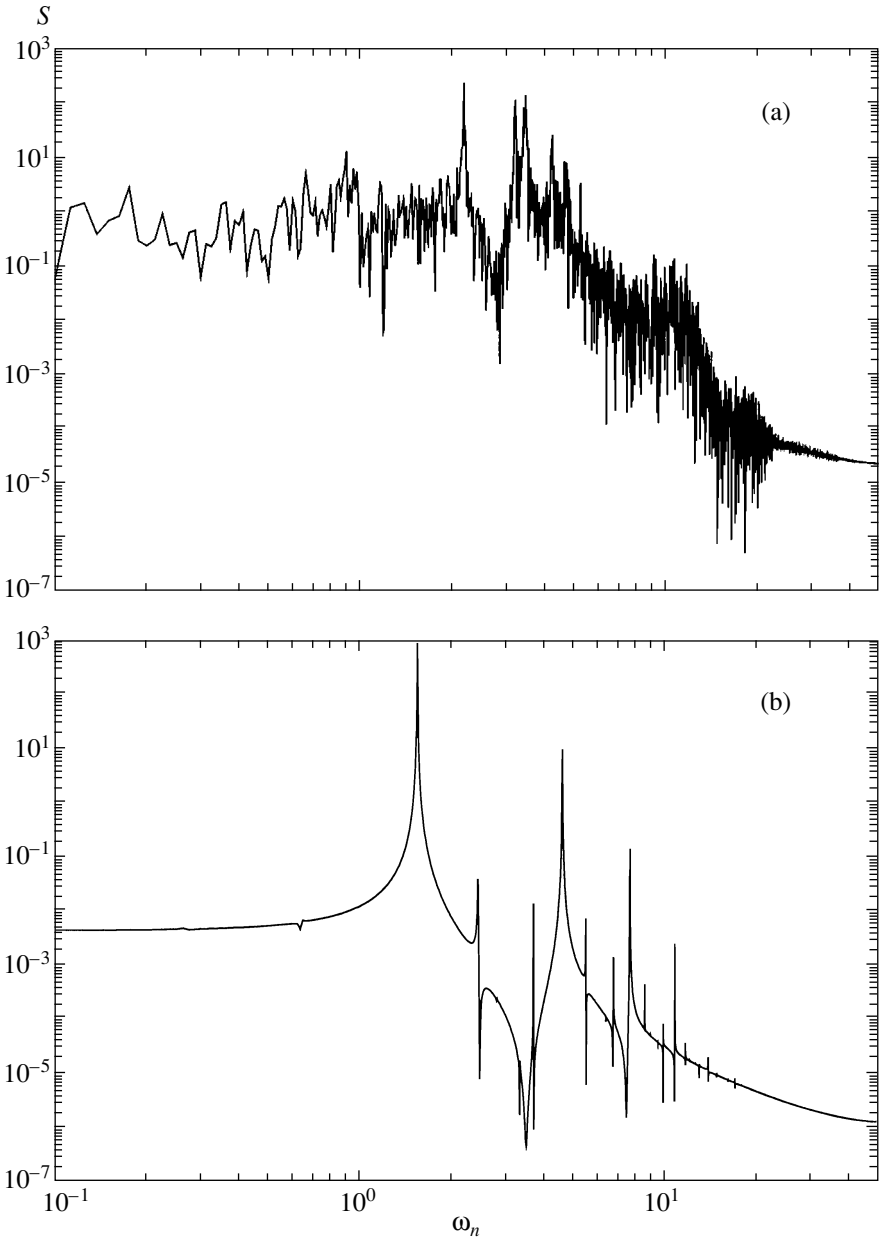


Fig. 6. Spectral density of a realization $x(t)$ for (a) original Duffing–Holmes oscillator (16) with $\alpha = 0.145$, $\beta = 8$, $\eta = 0.03$, $\gamma = 0.14$, and $\Omega = \omega = 1.1$ and (b) regularized Duffing–Holmes oscillator (31) with $a = 2$.

Since this Melnikov function $D(t_0, \omega)$ obviously admits additive shift from its critical values, chaotic behavior of the pendulum is suppressed by the perturbation

$$f^*(\omega, t) = -\frac{4\pi a\delta(t)}{\dot{x}_0(t-t_0)}\cos(\omega t), \qquad (30)$$

where $\dot{x}_0(t)$ is the solution on the unperturbed separatrix.

Physically, the results obtained here mean that dynamics of the Duffing–Holmes oscillator and pendulum are regularized by series of "kicks."

4.3. Numerical Results

In the preceding section, it is shown that chaos in Duffing–Holmes-oscillator and pendulum dynamics can be suppressed by applying perturbations (26) and (30), respectively. In this section, we present the results of a numerical analysis.

We consider Eqs. (16) and (27). In dynamics of the Duffing–Holmes oscillator, the onset of chaos corre-

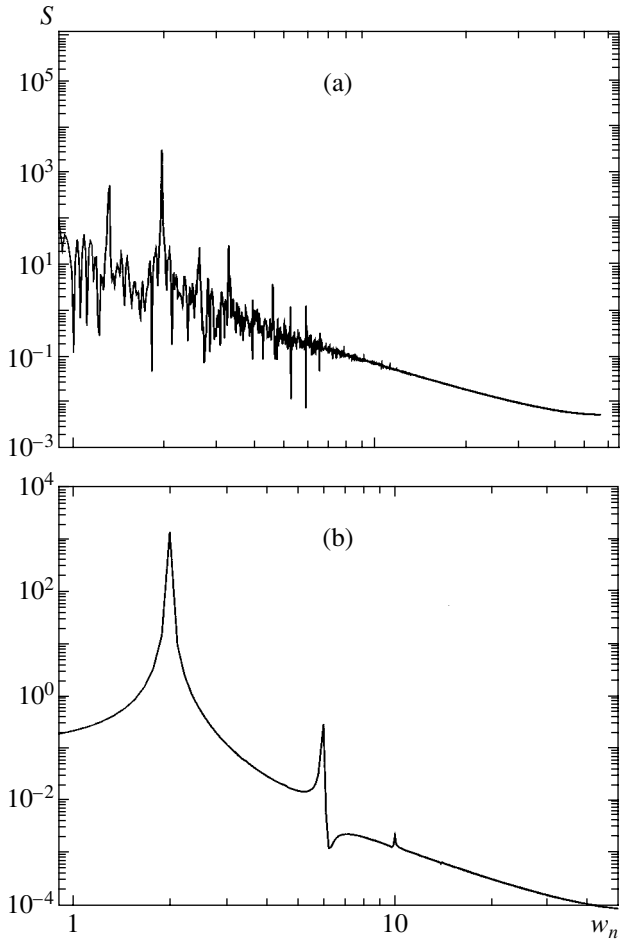


Fig. 7. Spectral density of a realization $x(t)$ for (a) original pendulum equation (27) with $\alpha = 0.04$, $\gamma = 1.35$, and $\omega = 1.0$ and (b) perturbed equation (32) with $a = 1.2$.

sponds to the breakdown of a figure-of-eight separatrix. Figure 2 illustrates the structure of a typical chaotic set obtained in this case. The onset of chaos in pendulum dynamics is associated with the breakdown of a heteroclinic trajectory (see Fig. 3).

Consider the Duffing–Holmes oscillator and pendulum with additional perturbations (26) and (30), respectively. The corresponding equations are

$$\ddot{x} - x + \beta[1 + \eta\cos(\Omega t)]x^3 = \varepsilon\Big[\gamma\cos(\omega t) - \alpha\dot{x} + 2\pi\sqrt{2\beta}\frac{\cosh^2(t-t_0)}{\sinh(t-t_0)}a\delta(t)\cos\Omega t\Big], \quad (31)$$

$$\ddot{x} + \alpha\dot{x} + \sin x = \gamma\cos(\omega t) + 2\pi\cosh(t-t_0)a\delta(t)\cos(\omega t). \quad (32)$$

Figures 4 and 5 show numerical solutions to systems (31) and (32), respectively. It is clear that the dynamics of both oscillator and pendulum approach regular regimes represented by periodic orbits.

To analyze systems (31) and (32) in more detail, we invoke the spectral density defined as

$$S(\omega) = \lim_{T\to\infty}\frac{1}{2\pi T}|X(\omega)|^2,$$

where $X(\omega)$ is the Fourier transform of a solution $x(t)$ to system (16) or (27). The spectral density provides a simple, but reliable characterization of dynamics of a system under study. It can readily be used to find out whether a motion is regular or chaotic.

Figures 6a and 7a show the spectral densities calculated for original systems (16) and (27), respectively; Figs. 6b and 7b, the spectral densities for systems subject to perturbations (26) and (30), respectively. These results demonstrate that chaos is suppressed and dynamics of both systems are regularized.

Taking different parameter values corresponding to chaotic behavior, one can find appropriate regularizing perturbations (see above) and obtain qualitatively similar results, i.e., change from chaotic states to regular oscillations.

Thus, our numerical analysis is consistent with the analytical results obtained in Section 4.

5. CONCLUSIONS

Separatrix splitting is a very convenient method for examining dynamical systems, because it can be used to obtain nonintegrability conditions for many applied problems in analytical form [23]. Currently, the problem of chaos suppression considered in this study is mainly solved by numerical methods (e.g., see [1–10]). However, asymptotic behavior of trajectories can be examined analytically. As a result, the distance between the separatrices split by a perturbation can be found in general form by applying a perturbation method in the vicinity of a homoclinic trajectory.

In this study, separatrix splitting is applied to explore the possibility of chaos suppression in dissipative systems. Analytical expressions are obtained for regularizing perturbations. These results are sufficiently general to be applied to various dynamical systems that admit separatrix splitting.

REFERENCES

1. S. Boccaletti, C. Grebogi, Y.-C. Lai, *et al.*, Phys. Rep. **329**, 103 (2000).
2. A. Loskutov and S. D. Rybalko, in *Proceedings of the 5th International Conference on Difference Equations and Applications, ICDEA'2000, Temuco, Chile, 2000* (Taylor and Francis, London, 2002), p. 207.
3. V. V. Alekseev and A. Yu. Loskutov, Vestn. Mosk. Univ., Ser. 3: Fiz. Astron. **26** (3), 40 (1985).

4. V. V. Alekseev and A. Yu. Loskutov, Dokl. Akad. Nauk SSSR **293**, 1346 (1987) [Sov. Phys. Dokl. **32**, 270 (1987)].
5. E. Ott, C. Grebogi, and J. A. Yorke, Phys. Rev. Lett. **64**, 1196 (1990).
6. F. J. Romeiras, E. Ott, C. Grebogi, and W. P. Dayawansa, Physica D (Amsterdam) **58**, 165 (1992).
7. T. Shinbrot, C. Grebogi, E. Ott, and J. A. Yorke, Nature **363**, 411 (1993).
8. T. Shinbrot, Adv. Phys. **44**, 73 (1995).
9. E. Ott and M. L. Spano, Phys. Today **48**, 34 (1995).
10. A. Loskutov, Comput. Math. Mod. **12**, 314 (2001).
11. V. K. Mel'nikov, Tr. Mosk. Mat. O–va **12**, 3 (1963).
12. R. Lima and M. Pettini, Phys. Rev. A **41**, 726 (1990).
13. F. Cuadros and R. Chacón, Phys. Rev. E **47**, 4628 (1993).
14. R. Lima and M. Pettini, Phys. Rev. E **47**, 4630 (1993).
15. R. Chacón, Phys. Rev. E **51**, 761 (1995).
16. Y. Kivshar, F. Rodelsperger, and H. Benner, Phys. Rev. E **49**, 319 (1994).
17. R. Chacón, Phys. Rev. Lett. **86**, 1737 (2001).
18. G. Cicogna and L. Fronzoni, Phys. Rev. A **42**, 1901 (1990).
19. V. G. Gel'freĭkh and V. F. Lazutkin, Usp. Mat. Nauk **56**, 79 (2001).
20. P. J. Holmes and J. E. Marsden, Commun. Math. Phys. **82**, 523 (1982).
21. J. Guckenheimer and P. J. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields* (Springer, Berlin, 1990).
22. J. L. Trueba, J. Rams, and M. A. F. Sanjuán, Int. J. Bifurcation Chaos Appl. Sci. Eng. **10**, 2257 (2000).
23. V. V. Kozlov, *Symmetry, Topology, and Resonances in the Hamilton Mechanics* (Udmurt. Gos. Univ., Izhevsk, 1995).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 1054–1062.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1204–1212.

MISCELLANEOUS

On Random Transformations of the Wave Function of a Two-Level System

V. A. Antonov[a] and B. P. Kondratyev[b,*]

[a]*Pulkovo Observatory, Russian Academy of Sciences, Pulkovskoe shosse 65, St. Petersburg, 196140 Russia*
[b]*Udmurt State University, Krasnoarmeĭskaya ul. 71, Izhevsk, 426034 Russia*
**e-mail: kond@uni.udm.ru*

Received March 7, 2003

Abstract—The process of random diffusion variation of the wave function of a system with two states is analyzed. A method is developed for calculating the evolution operator and the damping increment of the probability distribution function of the state of the system on the basis of quaternion apparatus. It is proved analytically that the second moments formed from the wave function play the major role since all other statistical characteristics tend to equilibrium at a higher rate. For more general models of a random action, the result remains asymptotically the same, but the relative orders of increments may be different. Exceptional cases of incomplete statistical equilibrium are singled out. The possible role of the given model problem in the actual problem of state splitting in the transition from the microworld to macroworld is discussed. It is shown that, in spite of the views expressed in modern literature, the distribution of finite probabilities in the white noise model does not allow the well-known Schrödinger's Cat paradox to be resolved. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The study of evolution of quantum systems subjected to the action of a nonstationary external perturbing field is undoubtedly interesting from the standpoint of distribution of finite probabilities and, as a consequence, the choice of possible quantum states for the system. In article [1] (see also [2]), such an evolution was studied in detail using the simple example of a two-level system interacting with classical white noise. The probabilities of transition of a particle from one well to another were calculated using a double functional integral along the trajectories, which was expanded into a power series in a small parameter. However, some features of the evolution of the quantum system in [1] remained unclear; we would like to analyze here these questions more rigorously.

It should be recalled that Lesovik *et al.* [1] considered the dynamics of a particle in a two-well potential (symmetric wells). The Hamiltonian of the particle has the form

$$H(t) = -\frac{\hbar}{2}\Delta\sigma_x + \frac{q_0\varphi(t)}{2}\sigma_z, \qquad (1)$$

where σ_x, σ_y, and σ_z are the Pauli matrices, Δ and q_0 are certain constants, and $\varphi(t)$ is classical white noise. The correlators connecting the initial and final states of the particle were calculated by integrating along trajectories. However, it was not rigorously proved that second moments must play the major role in this problem. In addition, higher order moments must also be taken into account in some cases. Finally, we cannot agree with the opinion formulated in [1, 2] that the ultimate inequality of the "weights" of two states emerging as a result of a random action of the thermostat is associated with the solution of the fundamental Schrödinger's Cat problem, i.e., with the choice of one of quantum states upon a transition from the microworld to the macroworld.

To solve these problems, we develop here a new method for analyzing evolutionary quantum problems based on the mathematical apparatus of quaternions. Using this method, a more rigorous and compact representation of the evolution of the state of a two-level system subjected to white noise can be obtained. We will not confine our analysis to a Hamiltonian of type (1) and consider the evolution of the system with more general Hamiltonians in Section 9.

2. FORMULATION OF THE PROBLEM

Let us consider a quantum system with two basic states, which is described by Hamiltonian (1). Constant Δ is reciprocal to the lifetime of the system in the given state and characterizes the stability of this state, while parameter q_0 has the dimension of length and defines the width of the potential barrier. The result of evolution on a time interval can be written in matrix form as

$$\psi(T) = D\psi(0), \qquad (2)$$

where

$$D = D_n D_{n-1} \ldots D_1,$$
$$D_m = \exp\left[-\frac{iH\left(\frac{mT}{n}\right)T}{n\hbar}\right], \quad (3)$$

in the limit $n \longrightarrow \infty$.

It is well known [3] that complex unitary 2×2 matrices are closely related to operations of rigid body rotation. As an intermediate link, it is convenient to use the apparatus of quaternions. Multiplication of quaternions

$$X = d + ai + bj + ck \text{ or } (a, b, c, d), \quad (4)$$

where a, b, c, and d are real-valued parameters, is equivalent [4] to the multiplication of complex matrices

$$D_x = \begin{pmatrix} d + ic & -(b - ia) \\ b + ia & d - ic \end{pmatrix}, \quad (5)$$

which are unitary under the normalization condition

$$a^2 + b^2 + c^2 + d^2 = 1. \quad (6)$$

In the given case, to the required accuracy, we have

$$D_m = 1 - \frac{iT}{n\hbar}\left[-\frac{\hbar\Delta}{2}\sigma_x + \frac{q_0\varphi\left(\frac{mT}{n}\right)}{2}\sigma_z\right] - \frac{q_0^2T^2}{8n^2\hbar^2}\left[\varphi\left(\frac{mT}{n}\right)\right]^2 + \ldots$$

and the corresponding quaternion is given by

$$X_m = 1 + \frac{T}{n\hbar}\left[\frac{\hbar\Delta}{2}i + \frac{q_0}{2}\varphi\left(\frac{mT}{n}\right)k\right] - \frac{q_0^2T^2}{8n^2\hbar^2}\left[\varphi\left(\frac{mT}{n}\right)\right]^2 + \ldots. \quad (7)$$

3. DIFFUSION EQUATION

In accordance with the general principles of using quaternions in the problem of rigid body rotation [5, 6], the previous expression describes a deterministic rotation through angle $T\Delta/n$ about the x axis during time T/n and a random rotation through angle

$$\frac{q_0T}{n\hbar}\varphi\left(\frac{mT}{n}\right)$$

about the z axis. We fix an arbitrary initial position of a test point on a unit sphere. In accordance with the general rules of the diffusion theory [7], the density f of subsequent images of this point at instants $t > 0$ obeys the differential equation

$$\frac{\partial f}{\partial t} = -\Delta A_x f + \frac{\Gamma}{2}A_z^2 f, \quad (8)$$

where the notation

$$\Gamma^2 = \frac{q_0^2}{\hbar^2}\langle\varphi^2\rangle$$

is used; the angle brackets indicate dispersion per unit time and A_x, A_y, and A_z are the operators of rotation about the x, y, and z axes. Equation (8) has the obvious steady-state solution f = const (const = $1/(4\pi)$ for the conventional normalization of probabilities). The remaining particular solutions attenuate with time.

4. SERIES EXPANSION OF FUNCTION f

It can easily be seen that a spherical function of any order l preserves its order after substitution into Eq. (8). Consequently, having fixed l, we can use the ordinary expansion

$$f = \sum_{k=-l}^{l} \alpha_k(t) Y_l^k. \quad (9)$$

Normalization of (complex) spherical functions is immaterial in principle; we choose normalization to simplify equations for quantities α_k. Separating time and introducing decrement λ, we obtain

$$(\lambda - k^2\Gamma)\alpha_k + \frac{\Delta}{2}[(l-k)\alpha_{k+1} - (l+k)\alpha_{k-1}] = 0, \quad -l \le k \le l. \quad (10)$$

If we replace all quantities α_k by $(-1)^k\alpha_{-k}$, the latter will obviously satisfy the same system (10) with the same λ. It follows hence that all solutions can be divided into even solutions with

$$\alpha_{-k} = (-1)^k\alpha_k \quad (11)$$

and odd solutions with

$$\alpha_{-k} = (-1)^{k+1}\alpha_k, \quad \alpha_0 = 0. \quad (11a)$$

5. LOWER BOUNDARY FOR λ

We will now try to determine the lower boundary for λ (for all $\lambda \ge 1$ and nontrivial $\alpha_0, \alpha_1, \ldots, \alpha_l$). To do this, we must prove that

$$\mathrm{Re}\lambda \ge \frac{\Gamma}{2} - \sqrt{\frac{\Gamma^2}{4} - \Delta^2}, \quad \Delta \le \frac{\Gamma}{2}, \quad \mathrm{Re}\lambda \ge \frac{\Gamma}{2}, \quad \Delta \ge \frac{\Gamma}{2}. \quad (12)$$

We begin with a simpler case of an odd solution. Multiplying each equation of system (10) with $k = 1, 2, \ldots, l$ by the quantity

$$\xi_k = \frac{\alpha_k^*}{(l+k)!(l-k)!}, \tag{13}$$

summing the results, and shifting the index of summation in the last sum, we obtain

$$\sum_{k=1}^{l} \frac{(\lambda - k^2\Gamma)|\alpha_k|^2}{(l+k)!(l-k)!} + \frac{\Delta}{2}\sum_{k=1}^{l-1} \frac{\alpha_{k+1}\alpha_k^*}{(l-k-1)!(l+k)!} - \frac{\Delta}{2}\sum_{k=1}^{l-1} \frac{\alpha_k\alpha_{k+1}^*}{(l-k-1)!(l+k)!} = 0. \tag{14}$$

Summing relation (14) with its complex conjugate, we find that

$$\sum_{k=1}^{l} \frac{(\mathrm{Re}\lambda - k^2\Gamma)|\alpha_k|^2}{(l+k)!(l-k)!} = 0.$$

Obviously, at least one of the coefficients in the numerator must be positive or zero. In the given case, this is equivalent to the requirement $\mathrm{Re}\lambda \geq \Gamma$, which is even more stringent than inequality (12).

Let us now consider the even case. Multiplying all equations in (10) by $(-1)^k\xi_k$, summing the results, and shifting the index in the last sum, we obtain

$$\sum_{k=-l}^{l} (-1)^k \frac{(\lambda - k^2\Gamma)|\alpha_k|^2}{(l+k)!(l-k)!} + \frac{\Delta}{2}\sum_{k=-l}^{l-1} (-1)^k \frac{\alpha_{k+1}\alpha_k^*}{(l-k-1)!(l+k)!} + \frac{\Delta}{2}\sum_{k=-l}^{l-1} (-1)^k \frac{\alpha_k\alpha_{k+1}^*}{(l-k-1)!(l+k)!} = 0. \tag{15}$$

Subtracting from this relation its complex conjugate, we obtain

$$\mathrm{Im}\lambda \sum_{k=-l}^{l} (-1)^k \frac{|\alpha_k|^2}{(l+k)!(l-k)!} = 0. \tag{16}$$

Now, we must distinguish between two cases.

(a) $\mathrm{Im}\lambda \neq 0$. In this case, the sum over k in relation (16) vanishes. Multiplying all equations in (10) by ξ_k and summing the results, we obtain a relation similar to (14), but with the lower limit $k = -l$ for all sums. Summing this relation with its complex conjugate, we obtain

$$\sum_{k=-l}^{l} \frac{(\mathrm{Re}\lambda - k^2\Gamma)|\alpha_k|^2}{(l+k)!(l-k)!} = 0. \tag{17}$$

We subtract Eq. (16) from (17), where $\mathrm{Im}\lambda$ is replaced by $\mathrm{Re}\lambda$. Taking into account the parity property (11), we obtain

$$-\sum^{o} \frac{k^2\Gamma|\alpha_k|^2}{(l+k)!(l-k)!} + \sum^{v} \frac{(2\mathrm{Re}\lambda - k^2\Gamma)|\alpha_k|^2}{(l+k)!(l-k)!} = 0, \tag{18}$$

where symbols o and v indicate summation over all even and odd positive $k \leq l$, respectively (the term with $k = 0$ is omitted). Obviously, at least one of the coefficients of the second sum in Eq. (18) must be positive or zero, which is equivalent to the inequality

$$\mathrm{Re}\lambda \geq \frac{\Gamma}{2}. \tag{19}$$

(b) $\mathrm{Im}\lambda = 0$. Assuming that $l \geq 2$, we again multiply Eqs. (10) by ξ_k (but now for $k \geq 2$) and sum the results. After summing with the complex conjugate equality, we obtain

$$\sum_{k=2}^{l} \frac{(\lambda - k^2\Gamma)|\alpha_k|^2}{(l+k)!(l-k)!} - \frac{\Delta(\alpha_1\alpha_2^+ + \alpha_2\alpha_1^+)}{4(l+1)!(l-2)!} = 0. \tag{20}$$

We write separately Eqs. (10) with $k = 1$ and $k = 0$, taking into account the symmetry condition:

$$(\lambda - \Gamma)\alpha_1 + \frac{\Delta}{2}[(l-1)\alpha_2 - (l+1)\alpha_0] = 0,$$
$$\lambda\alpha_0 + l\Delta\alpha_1 = 0.$$

This leads to the equation

$$\frac{\alpha_1}{\alpha_2} = -\frac{\frac{\Delta}{2}(l-1)\lambda}{\lambda^2 - \Gamma\lambda + \frac{l(l+1)\Delta^2}{2}}. \tag{21}$$

If the denominator of this equation has real roots and the value of λ is greater than or equal to the smallest of these roots, we have

$$\Gamma^2 \geq 2l(l+1)\Delta^2 > 4\Delta^2,$$
$$\lambda > \frac{\Gamma}{2} - \sqrt{\frac{\Gamma^2}{4} - \frac{l(l+1)\Delta^2}{2}}$$

and inequality (12) holds. In other cases, the denominator in relation (21) is positive. Substituting relation (21)

into (20), we obtain

$$\left\{\lambda - 4\Gamma + \frac{\Delta^2(l-1)(l+2)\lambda}{4\left[\lambda^2 - \Gamma\lambda + \frac{l(l+1)\Delta^2}{2}\right]}\right\} \times \frac{|\alpha_2|^2}{(l+2)!(l-2)!} + \sum_{k=3}^{l} \frac{(\lambda - k^2\Gamma)|\alpha_k|^2}{(l+k)!(l-k)!} = 0. \quad (22)$$

In this case also, at least one of the coefficients in Eq. (22) must be positive (or equal to zero). If this is valid for one of the coefficients of the sum with $k \geq 3$, we have $\lambda \geq 9\Gamma$, which is even more stringent condition than (19). It remains for us to consider the behavior of the coefficient in the braces. It increases for

$$\lambda < \Delta\sqrt{\frac{l(l+1)}{2}}. \quad (23)$$

Henceforth, we assume that inequality (23) is valid since, otherwise, $\lambda > \Delta$, and this estimate is more stringent than (12). Substituting the real-valued λ from inequality (12), where the sign "≥" is replaced by the equality sign, we see that the coefficient in the braces is negative. Consequently, the actual value of λ must be higher.

In the simples case when $l = 1$, system (10) has the form

$$(\lambda - \Gamma)\alpha_1 - \Delta\alpha_0 = 0, \quad \lambda\alpha_0 + \Delta\alpha_1 = 0.$$

Equating the determinant to zero, we obtain

$$\lambda_{1,2} = \frac{\Gamma}{2} \pm \sqrt{\frac{\Gamma^2}{4} - \Delta^2}. \quad (24)$$

Solution (24) has already been obtained in a different way in [1] and satisfies inequality (12) as exact equality for one or both (for $\Delta \geq 2\Gamma$) roots (24). Thus, inequality (12) holds in all cases.

6. ROTATION OF THE 4-SPACE

From the behavior of a test point, we pass to rotation of the space as a whole, which is described by the product of random quaternions:

$$X = X_n X_{n-1} \ldots X_1. \quad (25)$$

In accordance with the general rules [5, 6], the transformation of the radius vector of the test point, which is formally treated as a purely vectorial quaternion $R = xi + yj + zk$, is described by the quaternion multiplication,

$$R' = XRX^{-1}. \quad (26)$$

Here, quaternion R' (which is obtained in a purely vectorial form) corresponds to a changed position of the test point on the sphere. Quaternion X normalized by condition (6) depends on three parameters, while vector R', which is also normalized to unity, depends only on two parameters. Consequently, X is not defined unambiguously from R and R'; there remains one more degree of freedom, which is manifested, as can easily be seen, in the possibility of replacing X by $X(\cos\sigma + R\sin\sigma)$ for a fixed R. In addition, it is possible to carry out the discrete substitution $X \longrightarrow -X$ in relation (26). To within these transformations, the inverse determination of X from R' or, in expanded form (e.g., for $R = k$), the solution of the equations

$$2(ac + bd) = x', \quad 2(bc - ad) = y',$$
$$d^2 + c^2 - a^2 - b^2 = r'$$

with allowance for relation (6) can be carried out unambiguously. Thus, the relation between the behavior of a random point on the sphere and product (25) of random quaternions is clarified. In both cases, we are dealing with a Markov chain with values on a sphere in a 3D or 4D space, respectively. The density f of distribution in the former case is connected with the density in the latter case via the integral relation

$$f = \langle p \rangle_R, \quad (27)$$

where the angle brackets indicate averaging over the rotational angle σ introduced above and over the inversion of the quaternion (from symmetry considerations, we conclude that the weights are equal). In the particular case, when $R = k$, the rotation in expanded form is expressed via the relations

$$a, b, c, d \longrightarrow (a\cos\psi - d\sin\psi, b\cos\psi + c\sin\psi,$$
$$c\cos\psi - b\sin\psi, d\cos\psi + a\sin\psi),$$
$$0 \leq \psi \leq 2\pi.$$

In the general case, we are dealing with rotation in a 4D space about a movable "axis," which, however, has the form of a 2D manifold constructed on quaternions 1 and R.

Let us prove the following lemma.

Lemma. An arbitrary function defined in the form of an even-degree polynomial of a, b, c, and d on a 4D sphere can be represented in the form of the sum of a finite number of polynomial functions of the same or lower degrees, which are a priori symmetric relative to

the above averaging carried out over in generally different axes.

The lemma will be proved by induction. For a zero-degree polynomial, the statement is obvious.

The expression

$$(d^2 + c^2)^m, \quad m = 1, 2, \ldots \tag{28}$$

has a symmetric structure. Quaternion k playing the role of an axis can be replaced by $\alpha i + \beta j + \gamma k = 1$, where α, β, and γ are connected only via the condition $\alpha^2 + \beta^2 + \gamma^2 = 1$. The relation

$$F = [d^2 + (\alpha a + \beta b + \gamma c)^2]^m$$

is symmetric relative to this new axis.

By virtue of condition (6), we have

$$d^2 = 1 - a^2 - b^2 - c^2; \tag{29}$$

consequently, to within the terms of a degree smaller than $2m$, the expression for F coincides with

$$N = [(\alpha a + \beta b + \gamma c)^2 - a^2 - b^2 - c^2]^m. \tag{30}$$

We assume that polynomials of a degree lower than or equal to $2m - 2$ are already represented in the required form; for this reason, we can consider N instead of F. First, we take the point $\alpha = \beta = 0$, $\gamma = 1$, at which

$$N = (-a^2 - b^2)^m.$$

We define polar angle θ_1 and azimuth ψ_1 on the sphere $\alpha^2 + \beta^2 + \gamma^2 = 1$ and spherical coordinates θ, ψ, and ρ in the abc space. To expand the function

$$(-a^2 - b^2)^m = (-\sin^2\theta)^m = (t^2 - 1)^m$$

in the Legendre polynomials in $t = \cos\theta$, we use the identity

$$\int_{-1}^{1} (1 - x^2)^m P_{2n}(x)\,dx = 2(-1)^n \frac{(2n-1)!![(2m)!!]^2}{(2n)!!(2n + 2m + 1)!!(2m - 2n)!!}, \quad m \geq n.$$

Consequently, in the special case of $\gamma = 1$, we have

$$N = \sum_{n=0}^{m} (-1)^{m-n} h_{nm} \rho^{2m} P_{2n}(\cos\theta_1) P_{2n}(\cos\theta),$$

where

$$h_{nm} = \frac{(2n+1)!![(2m)!!]^2}{(2n)!!(2n + 2m + 1)!!(2m - 2n)!!}, \quad m \geq n,$$

$$h_{mn} = 0, \quad m < n.$$

In the general case, we obtain

$$\begin{aligned} N &= \sum_{n=0}^{m} (-1)^{m-n} h_{nm} \rho^{2m} P_{2n}(\cos\xi) \\ &= \sum_{n=0}^{m} (-1)^{m-n} h_{nm} \rho^{2m} \Bigg\{ P_{2n}(\cos\theta) P_{2n}(\cos\theta_1) \\ &\quad + 2 \sum_{k=0}^{2n} \frac{(2n-k)!}{(2n+k)!} P_{2n}^{(k)}(\cos\theta) \\ &\quad \times P_{2n}^{(k)}(\cos\theta_1) \cos[k(\psi - \psi_1)] \Bigg\}, \end{aligned} \tag{31}$$

where ξ is the angle between the vectors with components (a, b, c) and (α, β, γ). The validity of relation (31) is associated with the obvious invariance of ξ relative to joint rotations of spaces abc and $\alpha\beta\gamma$.

Spherical functions are mutually independent in view of their orthogonality. For each specific m, we are dealing only with a finite system of such functions; their independence must be manifested even on a finite set of points $(\alpha_j, \beta_j, \gamma_j)$ on a sphere. Taking an appropriate linear combination from corresponding functions N, we can always separate in this way the term with a single spherical function of λ_1 and θ_1 in relation (31). If we eliminate nonzero numerical factors, the coefficient of this term can be represented as

$$\rho^{2(m-n)} \times \left[\rho^{2n} P_{2n}^{(k)}(\cos\theta) \begin{cases} \cos(k\psi), & 0 \leq k \leq 2n \\ \sin(k\psi), & 1 \leq k \leq 2n \end{cases} \right], \tag{32}$$

where the brackets contain various harmonic polynomials of degree $n = 0, 1, 2, \ldots, m$. It can easily be seen that these polynomials form a complete system in the class of all homogeneous polynomials of a, b, c of degree $2n$; for all these polynomials, we can obtain the required representation using a superposition of averaged functions.

Analogously to relation (28), the expression

$$(d^2 + c^2)^m (ad + bc) \tag{33}$$

is symmetric. Instead of a, b, and c, we can substitute into this relation the new orthogonal components $\beta a - \alpha b$, $-(\alpha a + \beta b)\gamma + (\alpha^2 + \beta^2)c$, $\alpha a + \beta b + \gamma c$. If we substitute a, b, and c for d^2, expression (33) will contain, in addition to terms of the type that has already been investigated, the term

$$N_1 = dN(\beta a - \alpha b).$$

On the sphere $a^2 + b^2 + c^2 = 1$, we have

$$\cos\xi = \alpha a + \beta b + \gamma c$$

and the application of the rotation operator in the *abc* space gives

$$\frac{\partial}{\partial\psi}P_{2n}(\cos\xi) = \left(a\frac{\partial}{\partial b} - b\frac{\partial}{\partial a}\right)P_{2n}(\alpha a + \beta b + \gamma c)$$

$$= (a\beta - b\alpha)P'_{2n}(\cos\xi).$$

Using the familiar identity

$$P_N(t) = \frac{P'_{N+1} - P'_{N-1}(t)}{2N+1}, \quad N \geq 1,$$

we obtain the following expression for any *a*, *b*, and *c*:

$$(a\beta - b\alpha)P_{2n}(\cos\xi)$$

$$= \frac{\rho}{4n+1}\frac{\partial}{\partial\psi}[P_{2n+1}(\cos\xi) - P_{2n-1}(\cos\xi)].$$

After simple combination of terms, taking into account expression (31), we obtain

$$N_1 = 2d\rho^{2m+1}\sum_{n=0}^{m}(-1)^{m-n}\left(\frac{h_{nm}}{4n+1} + \frac{h_{n+1,m}}{4n+5}\right)$$

$$\times \sum_{k=1}^{2n+1} k\frac{(2n+1-k)!}{(2n+1+k)!} \tag{34}$$

$$\times P^{(k)}_{2n+1}(\cos\theta)P^{(k)}_{2n+1}(\cos\theta_1)\sin[k(\psi_1 - \psi)].$$

Arguing in the same way as in the previous case, we must only stipulate that expression (34) does not contain harmonic polynomials with $k = 0$. However, this does not affect the final result since the "band" harmonic polynomial ($k = 0$) can be obtained from any other polynomial appearing in expression (34) by averaging over $2n$ orientations for rotation about any axis along which this selected polynomial does not vanish. Thus, the polynomials of degree $2(m + 1)$, which contain one factor *d*, also possess the required property. Since higher powers of *d* ($d \geq 2$) can be eliminated, the lemma is completely proved.

7. REPRESENTATION OF ROTATIONS BY MULTIPLICATION OF QUATERNIONS

We can now easily establish the relation between the superposition of rotations of a fixed "tag" and the quaternion multiplication that carried more information. After addition of new random cofactors X_n, product (25) again behaves as a Markov chain. The degree of the polynomial describing the density of distribution over sphere (6) is also preserved. The entire process can be described using the polynomial eigenfunctions $p_j(a, b\ c, d)$ with corresponding decrements λ_j. In accordance with the lemma, for an even-degree polynomial *p*, an expansion of the type

$$p(a, b, c, d) = \sum_{i=1}^{N} q_i(a, b, c, d)$$

must exist, where polynomial q_i is symmetric relative to rotation about the R_i axis. If we denote by $d\omega$ an element of a sphere in the 4D space, we can write

$$\int |p(a, b, c, d)|^2 d\omega$$

$$= \sum_{i=1}^{N}\int p(a, b, c, d)q_i^+(a, b, c, d)d\omega \tag{35}$$

$$= \sum_{i=1}^{N}\int \langle p(a, b, c, d)\rangle_{R_i} q_i^+(a, b, c, d)d\omega.$$

The left-hand side of this expression is positive; consequently, at least one of the functions $\langle p(a, b, c, d)\rangle_{R_i}$ on the right-hand side is not identically equal to zero. Earlier, we demonstrated the relation (which is linear for densities) between the rotation of a sphere with a tagged point R_i and a superposition of quaternions. Consequently, for densities of even powers of *a*, *b*, *c*, and *d*, the values of decrement λ are chosen in the same way as for rotations with a tagged point.

A different situation is observed for polynomials of odd powers of *a*, *b*, *c*, and *d*, which may correspond to a wider set of decrements. If, however, we return to the initial appearance of quaternions in (5), it can be seen that simultaneous sign reversals in *a*, *b*, *c*, and *d* lead only to the substitution of $-\psi$ for ψ. This cannot have any statistical consequences and, hence, the difference between (*a*, *b*, *c*, *d*) and (*–a*, *–b*, *–c*, *–d*) is not informative in the given formulation of the problem.

8. QUANTUM PROBABILITIES

We can easily find the final distribution of the squared modulus for any of two components of the wave function. Since the result cannot depend on the initial distribution, we assume that at least $\psi_1(0) = 1$ and $\psi_2(0) = 0$ for $t = 0$; in this case, from relation (5) we obtain

$$\psi_1(T) = d + ic, \quad |\psi_1(T)|^2 = c^2 + d^2.$$

As $T \longrightarrow \infty$, the distribution on sphere (6) tends to a uniform distribution in view of symmetry. Taking the parametrization in the form

$$a = \cos\eta\cos\delta, \quad b = \cos\eta\sin\delta, \quad c = \sin\eta\cos\varepsilon,$$

$$d = \sin\eta\sin\varepsilon, \quad d\omega = \sin\eta\cos\eta d\eta d\delta d\varepsilon,$$

we obtain for the inequality

$$Q < c^2 + d^2 < Q + dQ, \quad 0 < Q < 1,$$

or

$$\sqrt{Q} < \sin\eta < \sqrt{Q} + \frac{dQ}{2\sqrt{Q}}$$

a probability proportional to dQ; in other words, for $|\psi_1|^2$ as well as, naturally, for $|\psi_2|^2$, we obtain a uniform distribution in the limit. This result was also noted in [1].

9. A MORE GENERAL FORMULATION OF THE PROBLEM

Our method permits a generalization of the problem when the form of the Hamiltonian (see expression (1)) is not specified; we only assume that random parameters appearing in this Hamiltonian are independent of one another on sequential intervals of time.

The relation with quaternions and with rotation of the sphere remains the same. It is sufficient to consider again the density $f(\tilde{\theta}, \tilde{\psi})$ on a sphere in a 3D space in polar coordinates as a superposition of spherical functions of a fixed principal index $l \geq 1$; i.e.,

$$f = \sum_{k=k_1}^{k_2} c_k e^{ik\bar{\psi}} P_l^{(|k|)}(\cos\tilde{\theta}). \tag{36}$$

Here, k_1 and k_2 stand for the largest and smallest values of index k from those for which $c_k \neq 0$. (Normalization of $Y_l^{(k)}$ in this case differs from that in relation (9), but this is immaterial.)

We consider functions (36) as eigenfunctions. As in particular example (1), the corresponding values of $\mathrm{Re}\lambda$ are generally positive. It will be proved below that exceptions (i.e., the existence of undamped solutions (36)) with $\lambda = 0$ or purely imaginary λ are observed only for models from the following two categories.

(a) Rotations of a sphere are reduced to rotations about the same (not random) axis, these rotations being probably combined with its inversion.

(b) Admissible rotations transform a definite regular polyhedron into itself.

To prove these statements, we again denote the tagged point on the sphere by R and possible rotations on each step by $L_1, L_2, \ldots, L_s$; irrespective of the step number, we ascribe probabilities $Q_1, Q_2, \ldots, Q_s$ to these rotations. (This latter discretization is fully immaterial and only makes verbal formulation more convenient.) Eigenfunction (36) must satisfy the basic relation for Markov chains, which in our case has the form

$$\exp\left(-\frac{\lambda T}{n}\right) f(R) = \sum_{q=1}^{s} Q_q f(L_q^{-1}R). \tag{37}$$

In accordance with the Buniakowski inequality, we have

$$\left|\sum_{q=1}^{s} Q_q f(L_q^{-1}R)\right|^2 \leq \sum_{q=1}^{s} Q_q \sum_{q=1}^{s} Q_q \left|f(L_q^{-1}R)\right|^2. \tag{38}$$

Introducing a spherical element $d\Omega$, evaluating the integrals over the entire sphere, and taking into account relation (37), as well as the natural normalization condition

$$\sum_{q=1}^{s} Q_q = 1,$$

we obtain

$$\exp\left(-\frac{2T\mathrm{Re}\lambda}{n}\right)\int |f|^2 d\Omega$$

$$\leq \sum_{q=1}^{s} Q_q \int \left|f(L_q^{-1}R)\right|^2 d\Omega \tag{39}$$

$$= \sum_{q=1}^{s} Q_q \int |f|^2 d\Omega = \int |f|^2 d\Omega.$$

In the case of zero or a purely imaginary λ in this relation, we are ultimately dealing with an equality. However, the exact equality in relation (38) is attained, in accordance with the general rule, only when all functions $f(L_q^{-1}R)$, $q = 1, 2, \ldots, s$ coincide to within numerical (complex-valued) coefficients. By virtue of relation (37), the relations

$$\left|f(L_q^{-1}R)\right| = \sigma_q f(R), \quad q = 1, 2, \ldots, s$$

with complex constants σ_q are valid. In view of conservation of the integral of $|f|^2$ upon rotation, we must have $|\sigma_q| = 1$ so that

$$\left|f(L_q^{-1}R)\right| = |f(R)|, \quad q = 1, 2, \ldots, s.$$

Thus, a certain body is invariant to all admissible rotations $L_1, \ldots, L_s$ (for better visualization, we imagine the surface $r = |f(R)|$. If it is a body of revolution other than a sphere, its axis must remain unchanged to within inversion, and we arrive at case (a). If we are not dealing with a body of revolution, it coincides with itself only in a finite number of ways so that not only $L_1, \ldots, L_q$, but also all their combinations of any number and order belong to the symmetry group of the body. However, such finite groups of rotation (unless they have a

unique invariant axis) are necessarily connected with a regular polyhedron [4] and we arrive at case (b).

It remains for us to consider the possibility of $|f| = \text{const}$. In accordance with relation (36), the expansion of $|f|^2$ into a complex Fourier series contains the last term of the form

$$|f|^2 = \ldots + c_{k_1} c_{k_2}^* e^{i(k_1 - k_2)\tilde{\psi}} P_l^{(|k_1|)}(\cos\tilde{\theta}) P_l^{(|k_2|)}(\cos\tilde{\theta}).$$

This term cannot be reduced to a constant for any $l \geq 1$. For $k_1 \neq k_2$, the dependence on $\tilde{\theta}$ remains, while for $k_1 = k_2$, the dependence on $\tilde{\psi}$ takes place. Thus, this case is irrelevant, which completes the proof.

As before, the result is automatically extended to the multiplication of quaternions X_m for indistinguishable X and $-X$. As applied to the evolution of the wave function, exceptional cases can be reformulated as follows:

(a) all realizations of matrices D_m commute with one another (an analog of rotation about the same axis);

(b) all realizations of matrices D_m belong to a finite group.

Since there are no infinitesimal rotations in case (b), it is irrelevant in the presence of diffusion, as in the case of Hamiltonian (1). Case (a) is formally realized in model (1) for $\Delta = 0$ and $\Gamma = 0$.

The following qualitative remark is also appropriate here. In contrast to particular example (1), other models do not contain an analog of inequality (12) ensuring the prevailing role of deviations from equilibrium with $l = 1$ over long time intervals. On the contrary, zero can now become the point of condensation of values of λ for large values of l. In this case, although densities $f(R)$ converge (except in the above-mentioned cases (a) and (b)) to a constant for $t \longrightarrow \infty$ due to the possibility of indefinitely exact approximation of any function continuous on a sphere by a polynomial, the establishment of statistical equilibrium can generally occur at a rate which is generally slower than that defined by any exponential function.

10. DISCUSSION OF THE PROBLEM

It was proposed in [1, 2, 8] that the inequality of "weights" of two states due to a random action of the thermostat (even if these weights were initially equal), which was noted in these publications and repeatedly discussed here, is related to the problem of choosing one of the possible quantum states in a transition to the macroworld. It was concluded in this connection [2] that the well-known Schrödinger's Cat paradox could be resolved in this way. However, this opinion is erroneous. As a matter of fact, to resolve the paradox in the macroworld, only one of the two components of the wave function, which is responsible for the readings of a macroinstrument, must remain. But the model with white noise obviously does not meet precisely this requirement. Indeed, for large values of t, the probability that, for example, one of the weights $|\psi_1|^2$ and $|\psi_2|^2$ becomes twice as large as the other amounts approximately to 2/3; however, this probability does not increase with time and cannot become equal to unity in any way. Meanwhile, to resolve the paradox of Schrödinger's Cat, the total (or asymptotic) exclusion of states is required.

It was found that a generalization to a wider class of models with a random Hamiltonian does not open any new opportunities since the statistical distribution for large t remains unchanged except for obviously degenerate situations.

By the way, it is mathematically obvious that the limiting probability distribution itself can also be obtained without using any quaternions, just by defining equal probability on the complex circle

$$|\psi_1|^2 + |\psi_2|^2 = 1,$$

analogously to calculations in statistical mechanics [9] (only not for an indefinitely increasing, but for a finite number of degrees of freedom). However, equilibrium probability distribution is presumed in [9] and in calculations based on traditional statistical mechanics in general, while we consider here the evolution to the equilibrium state. However, the results (and this is important) match. Internal correlations in Hamiltonian $H(t)$ can hardly change the situation since the processes that are treated as random can be split into almost independent segments. In principle, we can imagine such a correlation of phenomena in a reservoir with the quantum state of an isolated system, such that the system tends to a "pure" state in the sense of the choice between $\psi_2 \longrightarrow 0$ and $\psi_1 \longrightarrow 0$. However, it can be analogously stated that black and white grains contained in a box can be separated and gathered in opposite parts of the box by shaking it if this shaking correlates in some intricate manner with the initial positions of the grains. In both cases, we would have a thermodynamic miracle, which is ruled out by our knowledge of the role of probability in the world [10, 11].

Consequently, we arrive at the conclusion that the necessary condition stating that "wave functions do not permit a superposition according to macroscopic features" cannot be expressed in a linear form. For this reason, any solution of the problem of transition from the microworld to the macroworld includes nonlinearity. However, the considerations formulated by Menskiĭ [12] and in the discussion concerning this publication do not provide a correct answer to crucial questions. It would be apparently erroneous to state that each quantum experiment should be completed at the boundary of the microworld. At present, this appears as archaic in the light of the studies carried out on an intermediate scale without an abrupt termination or a clearly manifested "observer." This suggests that nonlinearity should be treated as an inherent property of the equation describing

the evolution of the wave function in the intermediate region between the micro- and the macroworld [10, 13].

In the approach developed in [13], the main feature, apart from nonlinearity, is the interaction with a certain reservoir of cosmic origin. It is significant, however, that it cannot be referred to as a thermostat in our case since this cosmic reservoir must have two auxiliary functions; i.e., it must be a source of fluctuations and absorb excessive indeterminacy. From the standpoint of physics, this situation resembles the role of solar radiation and open cosmic space in life on the Earth.

The following is also worth noting. It was shown in [13] that the separation of a superposition of states under these assumptions can be carried out, but the role of the cosmic factor turns out to be only auxiliary and it does not carry any a priori information on the result of splitting. An appropriate analog here can be symmetric diffusion of a molecule between two cold walls to one of which the molecule ultimately sticks. In order to achieve the same result in the interaction between a quantum system and a thermostat, the latter must possess miraculous properties in each (!) specific experiment. This constitutes the principal difference between the model with a thermostat described in [8] and the model considered in [13].

REFERENCES

1. G. B. Lesovik, A. V. Lebedev, and A. O. Imambaev, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 565 (2002) [JETP Lett. **75**, 474 (2002)].
2. G. B. Lesovik, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 528 (2001) [JETP Lett. **74**, 471 (2001)].
3. H. Jeffreys and B. Swirles, *Methods of Mathematical Physics*, 3rd ed. (Cambridge Univ. Press, Cambridge, 1966; Mir, Moscow, 1969).
4. F. Klein, *Lectures on the Icosahedron and the Solution of Equations of the Fifth Degree*, 2nd ed. (Dover, New York, 1956; Nauka, Moscow, 1989).
5. V. N. Branets and N. P. Shmyglevskiĭ, *Application of Quaternions in Problems of Orientation of Solids* (Nauka, Moscow, 1973) [in Russian].
6. S. N. Kirpichnikov and V. S. Novoselov, *Mathematical Aspects of Kinematics of Solids* (Leningr. Gos. Univ., Leningrad, 1986) [in Russian].
7. M. S. Bartlett, *An Introduction to Stochastic Processes, with Special Reference to Methods and Applications*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1966; Inostrannaya Literatura, Moscow, 1958).
8. G. B. Lesovik, Usp. Fiz. Nauk **171**, 449 (2001) [Phys. Usp. **44**, 429 (2001)].
9. A. Ya. Khinchin, *Mathematical Foundations of Quantum Statistics* (Gostekhizdat, Moscow, 1951) [in Russian].
10. B. B. Kadomtsev, *Dynamics and Information* (UFN, Moscow, 1997) [in Russian].
11. N. N. Bogolyubov and Yu. V. Sanochkin, *Ludwig Boltzmann: Papers and Speeches* (Nauka, Moscow, 1970) [in Russian].
12. M. B. Menskiĭ, Usp. Fiz. Nauk **170**, 631 (2000) [Phys. Usp. **43**, 585 (2000)].
13. B. P. Kondratyev and V. A. Antonov, *Solution of the Schrödinger Paradox. An Experiment in Creation of Nonlinear Quantum Mechanics* (Udmurt. Gos. Univ., Izhevsk, 1994) [in Russian].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 837–852.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 963–981.

GRAVITATION, ASTROPHYSICS

Post-post-Newtonian Equations of Motion for Point Particles in the General Theory of Relativity

Cz. Jankewicz* and I. Nowak**

Institute of Physics, Rzeszow University, PL 35-310 Rzeszow, Poland
** e-mail: czjan@univ.rzeszow.pl*
*** e-mail: iwnow@univ.rzeszow.pl*
Received July 11, 2003

Abstract—The post-post-Newtonian equations of motion for point particles are derived from the Einstein gravitational field equations by using the Einstein–Infeld–Hoffmann method with the help of the energy-momentum tensor proposed by Infeld and Plebanski [5, 6]. The obtained equations of motion coincide with the equations derived by Kopeikin [10] by using the Fock method. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Einstein, Infeld, and Hoffmann and, independently, Fock developed two methods for deriving the Newtonian and post-Newtonian equations of motion for point particles from the Einstein gravitational field equations. These methods are described in detail in the monograph by Fock [1] (Chapter 6) and in the book by Infeld and Plebanski [2] (Chapter 3).

The main features of the Fock method are as follows: (i) the field equations with the energy-momentum tensor of the continuous medium are used; (ii) the metric tensor is expanded into series in a small parameter, which formally corresponds to the series expansion in the reciprocal velocity of light; (iii) in harmonic coordinates, these expansions reduce the approximate field equations to wave equations; (iv) in the solutions to these wave equations, retardation corrections are taken into account; (v) the equations of motion are derived from the condition of integrability of the field equations in the form of covariant laws of conservation for the energy-momentum tensor, which follow from the Bianchi identities (in particular, point particles are defined as centers of mass of elastic spherically symmetric nonrotating bodies with sizes smaller than the distances between them); and (vi) the final values of the metric tensor on the world lines of centers of mass are obtained by using the equations describing the internal structure of the bodies. In pioneering works [3, 4], the equations of motion are derived not from the covariant laws of conservation for the energy-momentum tensor, but from the equations defining harmonic coordinates.

In the Einstein–Infeld–Hoffmann (EIH) method, as in the Fock method, the metric tensor is expanded into reciprocal power series in the velocity of light. In addition, it is assumed that the derivatives of the metric tensor expansion coefficients with respect to time and spatial coordinates do not change their order of smallness. This additional assumption, which forms the essence of the EIH method of successive approximations, makes it possible to reduce the approximate field equations to equations of the Poisson type. The equations of motion are derived from covariant laws of conservation for the energy-momentum tensor. In particular, point particles are defined as singularities of the metric tensor, which satisfies the field equations with the energy-momentum tensor containing the Infeld–Plebanski $\hat{\delta}$ functions [5, 6]. These $\hat{\delta}$ functions also ensure regularization of the metric tensor on the world lines of singularities. In their earlier publications [7–9], Einstein, Infeld, and Hoffmann derived the equations of motion not from the covariant laws of conservation for the energy-momentum tensor, but from the integrability conditions for the field equations in the form of two-dimensional surface integrals surrounding singularities.

Kopeikin [10], who extended the Fock method, derived from the field equations the post-post-Newtonian equations of motion for point particles with radiation corrections to these equations.

Here, we derive the post-post-Newtonian equations of motion for point particles with the Infeld–Plebanski energy-momentum tensor using the EIH method of successive approximation.

The equations of motion derived by Kopeikin and by us coincide with the equations of motion obtained earlier in [11, 12], in which the Hadamard *partie finie* method of regularization is employed.

The equations of motion can also be derived using the Arnowitt–Deser–Misner (ADM) method [13, 14]. For example, publications [15, 16] are in line with this approach. Since the harmonic coordinates used by Kopeikin [10] and by us are not admitted in this method, the resultant post-post-Newtonian equations of motion are different.

1063-7761/04/9805-0837$26.00 © 2004 MAIK "Nauka/Interperiodica"

2. BASIC CONCEPTS AND NOTATION

We will write the Einstein equations for the gravitational field in the form

$$R_{\mu\nu} = -\frac{8\pi k}{c^2}T^*_{\mu\nu}, \tag{2.1}$$

where

$$T^*_{\mu\nu} = \left(g_{\mu\rho}g_{\nu\sigma} - \frac{1}{2}g_{\mu\nu}g_{\rho\sigma}\right)T^{\rho\sigma}.$$

In the harmonic coordinates defined by the equations

$$\partial_\alpha(\sqrt{-g}g^{\alpha\beta}) = 0, \tag{2.2}$$

the Ricci tensor has the form

$$R_{\mu\nu} = \frac{1}{2}g^{\alpha\beta}\partial_\alpha\partial_\beta g_{\mu\nu} - g^{\alpha\beta}g^{\rho\sigma}(\Gamma_{\rho\mu\alpha}\Gamma_{\sigma\nu\beta} + \Gamma_{\mu\rho\alpha}\Gamma_{\sigma\nu\beta} + \Gamma_{\nu\rho\alpha}\Gamma_{\sigma\mu\beta}), \tag{2.3}$$

where

$$\Gamma_{\rho\mu\nu} = \frac{1}{2}(\partial_\mu g_{\nu\rho} + \partial_\nu g_{\mu\rho} - \partial_\rho g_{\mu\nu}).$$

For N point particles, the energy-momentum tensor proposed by Infeld and Plebanski [2], Chapter 1, is given by

$$\sqrt{-g}T^{\mu\nu} = \sum_A m_A\hat{\delta}(\bar{x} - \bar{\xi}_A) \times\left(g_{\rho\sigma}\frac{d\xi^\rho_A}{dx^0}\frac{d\xi^\sigma_A}{dx^0}\right)^{-1/2}\frac{d\xi^\mu_A}{dx^0}\frac{d\xi^\nu_A}{dx^0}. \tag{2.4}$$

Field equations (2.1) and Bianchi's identities lead to the following equations of motion for point particles treated as singularities of the metric tensor (see [2], Chapter 1):

$$\frac{d^2\xi^a_A}{(dx^0)^2} + \left(\overset{A}{\Gamma}{}^a_{\mu\nu} - \frac{d\xi^a_A}{dx^0}\overset{A}{\Gamma}{}^0_{\mu\nu}\right)\frac{d\xi^\mu_A}{dx^0}\frac{d\xi^\nu_A}{dx^0} = 0. \tag{2.5}$$

We are using the following notation: k is the gravitational constant; c is the velocity of light; Greek indices assume the values of 0, 1, 2, and 3, while Latin indices (unless other values are specified) assume the values of 1, 2, 3; recurring indices indicating the corresponding summation; indices A, B, and C assume values of 1, 2, …, N, where N is the number of particles; $(\bar{x}) = (x^a)$ and $x^0 = ct$ correspond to spatial and time coordinates of a point in the field, respectively; $(\bar{\xi}_A) = (\xi^a_A)$ and $\xi^0_A = x^0$ are the spatial and time coordinates of point particles, respectively;

$$|\bar{x}| = r, \quad \bar{x} - \bar{\xi}_A = \bar{r}_A, \quad |\bar{x} - \bar{\xi}_A| = r_A,$$

$$\bar{\xi}_A - \bar{\xi}_B = \bar{r}_{AB}, \quad |\bar{\xi}_A - \bar{\xi}_B| = r_{AB},$$

$$N^a_A = r_A^{-1}(x^a - \xi^a_A), \quad N^a_{AB} = r_A^{-1}(\xi^a_A - \xi^a_B),$$

$$V^a_A = \frac{d\xi^a_A}{dt}, \quad W^a_A = \frac{d^2\xi^a_A}{dt^2},$$

$\Delta = \partial_a\partial_a$; ∂_μ are the derivatives with respect to coordinates x^μ; ∂^A_a are the derivatives with respect to coordinates ξ^a_A; the dot on the function indicates the derivative with respect to time t; and the signature of the metric tensor coincides with the signature of the Minkowski tensor, which is equal to (+, –, –, –).

The $\hat{\delta}$ function introduced by Infeld and Plebanski possesses all properties of the Dirac δ function as well as the additional property (see Appendix in [2])

$$\int\hat{\delta}(\bar{r}_A)r_A^{-p}(d\bar{x}) = 0, \quad p = 1, 2, \ldots, L. \tag{2.6}$$

We will also use the following notation:

$$\overset{A}{(\ldots)} = \int(\ldots)\hat{\delta}(\bar{r}_A)(d\bar{x}).$$

We will seek the solutions to field equations (2.1) by the EIH method of successive approximations, assuming that

(a) the metric tensor can be expanded into the power series

$$g_{00} = 1 + c^{-2}\underset{(2)}{h_{00}} + c^{-4}\underset{(4)}{h_{00}} + c^{-6}\underset{(6)}{h_{00}} + \ldots,$$

$$g_{0n} = c^{-3}\underset{(3)}{h_{0n}} + c^{-5}\underset{(5)}{h_{0n}} + \ldots, \tag{2.7}$$

$$g_{mn} = -\delta_{mn} + c^{-2}\underset{(2)}{h_{mn}} + c^{-4}\underset{(4)}{h_{mn}} + \ldots;$$

(b) the derivative of the expansion coefficients of the metric tensor with respect to time t and with respect to spatial coordinates x^a do not change their order of smallness:

$$\underset{(i)}{h_{\mu\nu}} \sim \frac{\partial}{\partial t}\underset{(i)}{h_{\mu\nu}} \sim \frac{\partial}{\partial x^a}\underset{(i)}{h_{\mu\nu}}. \tag{2.8}$$

Propositions (a) and (b) make it possible to reduce the approximate field equations to equations of the Poisson type with a generally unlimited carrier of field sources.

Instead of expansion (2.7), we can use a more general expansion,

$$g_{\mu\nu} = \underset{(0)}{g_{\mu\nu}} + \sum_{i=1}^{\infty} c^{-i} \underset{(i)}{h_{\mu\nu}}; \tag{2.9}$$

however, in the harmonic system of coordinates [17, 18], the field equations give

$$\underset{(0)}{g_{\mu\nu}} = \eta_{\mu\nu},$$

while the equalities

$$\underset{(2i-1)}{h_{00}} = 0, \quad \underset{(2i)}{h_{0n}} = 0, \quad \underset{(2i-1)}{h_{mn}} = 0, \quad i = 1, 2, \ldots$$

are the propositions used in the EIH method, which takes into account only standing waves [7–9, 19].

In order to reduce the approximate field equations to the Poisson equations with a limited carrier of field sources, we will use the identities

$$\partial_a(fg) \equiv f\partial_a g + f\partial_a f, \tag{2.10}$$

$$\begin{aligned}&\partial_a\partial_b(fg)\\ &\equiv f\partial_a\partial_b g + g\partial_a\partial_b f + \partial_a f\partial_b g + \partial_a g\partial_b f.\end{aligned} \tag{2.11}$$

For singular functions f and g, these identities define the derivatives of the products of the functions in terms of the product of the derivatives of these functions. We will apply identities (2.10) and (2.11) (see Appendix A) to the singular functions

$$f = r_A^{-1}, r_A^{-2}, \ldots, \quad g = r_B^{-1}, r_B^{-2}, \ldots.$$

3. POST-NEWTONIAN APPROXIMATION

Using formulas (2.3), (2.4), (2.7), and (2.8), we obtain the following equations from field equations (2.1):

$$\begin{aligned}\underset{(2)}{\Delta h_{00}} &= 8\pi k\sum_A m_A\hat{\delta}(\bar{r}_A),\\ \underset{(2)}{\Delta h_{mn}} &= 8\pi k\sum_A m_A\hat{\delta}(\bar{r}_A)\delta_{mn},\\ \underset{(3)}{\Delta h_{0n}} &= -16\pi k\sum_A m_A\hat{\delta}(\bar{r}_A)V_A^n.\end{aligned} \tag{3.1}$$

The solutions to these equations have the form

$$\begin{aligned}\underset{(2)}{h_{00}} &= -2\Phi, \quad \underset{(2)}{h_{mn}} = -2\Phi\delta_{mn},\\ \underset{(3)}{h_{0n}} &= 4\Phi_n,\end{aligned} \tag{3.2}$$

where

$$\Phi = k\sum_A m_A r_A^{-1}, \quad \Phi_n = k\sum_A m_A V_A^n r_A^{-1}. \tag{3.3}$$

Using relations (2.3), (2.4), (2.7), and (2.8), as well as solutions (3.2), we obtain from Eqs. (2.1)

$$\begin{aligned}&\underset{(4)}{\Delta h_{00}} + 2\ddot{\Phi} + 4\Phi\Delta\Phi - 4\partial_k\Phi\partial_k\Phi\\ &= 16\pi k\sum_A m_A\hat{\delta}(\bar{r}_A)\left(\frac{3}{4}V_A^2 - \frac{5}{2}\Phi\right).\end{aligned} \tag{3.4}$$

Considering that

$$\ddot{\Phi} = \frac{1}{2}\Delta\ddot{\chi}, \quad \chi = k\sum_A m_A r_A, \tag{3.5}$$

$$\Delta\Phi = -4\pi\sum_A m_A\hat{\delta}(\bar{r}_A), \tag{3.6}$$

and taking into account the identity

$$\Delta\Phi^2 \equiv 2\Phi\Delta\Phi + 2\partial_k\Phi\partial_k\Phi,$$

which follows from relations (A.5) and (A.6), we obtain

$$\begin{aligned}\Delta\left(\underset{(4)}{h_{00}} + \ddot{\chi} - 2\Phi^2\right) &= 16\pi k\sum_A m_A\hat{\delta}(\bar{r}_A)\\ &\times\left(\frac{3}{4}V_A^2 - \frac{1}{2}\Phi\right).\end{aligned} \tag{3.7}$$

To solve this equation taking into account the first expression from (3.3), we consider the integral

$$\begin{aligned}&\int\hat{\delta}(\bar{x}' - \bar{\xi}_A)\Phi'|\bar{x} - \bar{x}'|^{-1}(d\bar{x}')\\ &= km_A\int\hat{\delta}(\bar{x}' - \bar{\xi}_A)r_A'^{-1}|\bar{x} - \bar{x}'|^{-1}(d\bar{x}')\\ &+ k\sum_{\substack{B\\B\neq A}} m_B\int\hat{\delta}(\bar{x}' - \bar{\xi}_A)r_B'^{-1}|\bar{x} - \bar{x}'|^{-1}(d\bar{x}').\end{aligned} \tag{3.8}$$

The first integral on the right-hand side gives a nonzero contribution only for $\bar{x}' - \bar{\xi}_A = 0$. To evaluate this integral, we expand $|\bar{x} - \bar{x}'|^{-1}$ into a series in $r_A' = |\bar{x}' - \bar{\xi}_A| \approx 0$:

$$|\bar{x} - \bar{x}'|^{-1} = \sum_{m=0}^{\infty} P_m(N_A^k N_A'^k)r_A^{-(m+1)}r_A'^m, \tag{3.9}$$

where $P_m(N_A^k N_A'^k)$ are the Legendre polynomials. Prop-

erty (2.6) of the $\hat{\delta}$ function leads to the equality

$$\int\hat{\delta}(\bar{x}'-\bar{\xi}_A)r_A'^{-1}|\bar{x}-\bar{x}'|^{-1}(d\bar{x}') = \int\hat{\delta}(\bar{x}'-\bar{\xi}_A)P_1(N_A^kN_A'^k)r_A^{-2}(d\bar{x}') = 0, \tag{3.10}$$

since $P_1(N_A^kN_A'^k) = N_A^kN_A'^k$ is an odd function of variable $\bar{r}_A'$. The second integral on the right-hand side of relation (3.8) can easily be evaluated since the integrand is continuous for $\bar{x}' = \bar{\xi}_A$. Thus, we can write

$$\int\hat{\delta}(\bar{x}'-\bar{\xi}_A)\Phi'|\bar{x}-\bar{x}'|^{-1}(d\bar{x}') = \overset{A}{\bar{\Phi}}r_A^{-1}, \tag{3.11}$$

where

$$\overset{A}{\bar{\Phi}} = k\sum_{\substack{B\\B\neq A}} m_Br_{AB}^{-1}.$$

Taking into account relations (3.10) and (3.11) we obtain from Eq. (3.7)

$$\underset{(4)}{h_{00}} = 2\Phi^2 - \ddot{\chi} + k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)r_A^{-1}. \tag{3.12}$$

Solutions (3.2) and (3.12) are sufficient for deriving the post-Newtonian equations of motion.

From formulas (2.5) and (2.7), in the post-Newtonian approximation, we have

$$\frac{d^2\xi_A^n}{dt^2} = \underset{(0)}{F_A^n} + c^{-2}\underset{(2)}{F_A^n}, \tag{3.13}$$

where

$$\underset{(0)}{F_A^n} = -\frac{1}{2}\underset{(2)}{\overset{A}{\partial_n h_{00}}}, \tag{3.14}$$

$$\underset{(2)}{F_A^n} = -\frac{1}{2}\underset{(4)}{\overset{A}{\partial_n h_{00}}} - \frac{1}{2}\underset{(2)}{\overset{A}{h_{nk}}}\underset{(2)}{\partial_k h_{00}} + \underset{(3)}{\dot{h}_{0n}} + \frac{1}{2}\underset{(2)}{\overset{A}{\dot{h}_{0(0)}}}V_A^n + \underset{(2)}{\dot{h}_{nk}}V_A^k$$

$$- \underset{(3)}{\overset{A}{\partial_n h_{0k}}}V_A^k + \underset{(3)}{\overset{A}{\partial_k h_{0n}}}V_A^k + \underset{(2)}{\overset{A}{\partial_k h_{00}}}V_A^kV_A^n \tag{3.15}$$

$$- \frac{1}{2}\underset{(2)}{\overset{A}{\partial_n h_{ks}}}V_A^kV_A^s + \underset{(2)}{\overset{A}{\partial_k h_{ns}}}V_A^kV_A^s.$$

Substituting solutions (3.2) and (3.12) into Eq. (3.15), we obtain the following integrals with singular functions, which after integration give

$$\int\hat{\delta}(\bar{r}_A)r_A^{-2}N_A^k(d\bar{x}) = 0,$$

$$\int\hat{\delta}(\bar{r}_A)r_A^{-3}N_A^kN_A^bN_A^s(d\bar{x}) = 0,$$

$$\int\hat{\delta}(\bar{r}_A)r_A^{-2}N_A^kN_A^b(d\bar{x}) = \frac{1}{3}\delta_{kb}\int\delta(\bar{r}_A)r_A^{-2}(d\bar{x}) = 0,$$

$$\int\hat{\delta}(\bar{r}_A)r_A^{-1}r_B^{-2}N_B^k(d\bar{x}) = 0, \quad A\neq B,$$

$$\int\hat{\delta}(\bar{r}_A)r_A^{-2}r_B^{-1}N_A^n(d\bar{x}) = 0, \quad A\neq B.$$

The first two integrals are equal to zero because the integrands are odd functions, while the third integral is equal to zero due to property (2.6) of the $\hat{\delta}$ function and the evenness of the integrands. The last two integrals can be evaluated using a series expansion of r_B^{-M} in $r_A \approx 0$:

$$r_B^{-L} = \sum_{m=0}^{\infty} C_m^{\frac{1}{2}L}(N_A^kN_{BA}^k)\frac{r_A^m}{r_{AB}^{m+L}}, \tag{3.16}$$

where $C_m^L(N_A^kN_{BA}^k)$ are the Gegenbauer polynomials.

Thus, using solutions (3.2) and (3.12) and taking into account relations (3.13)–(3.15), we obtain

$$\frac{d^2\xi_A^n}{dt^2} = k\sum_{\substack{B\\B\neq A}} m_A\partial_n^Ar_{AB}^{-1}$$

$$+ c^{-2}\left\{k\sum_{\substack{B\\B\neq A}} m_B\left[\frac{1}{2}V_B^kV_B^s\partial_n^A\partial_k^A\partial_s^Ar_{AB}\right.\right.$$

$$+ (4V_B^nV_A^k - 4V_A^nV_A^k - 4V_B^nV_B^k + 3V_A^nV_B^k)\partial_k^Ar_{AB}^{-1}$$

$$+ \left(V_A^2 - 4V_A^kV_B^k + \frac{3}{2}V_B^2\right)\partial_n^Ar_{AB}^{-1} \tag{3.17}$$

$$\left.- k(5m_A + 4m_B)r_{AB}^{-1}\partial_n^Ar_{AB}^{-1}\right]$$

$$- \frac{1}{2}k^2\sum_{\substack{B\\B\neq A\\B\neq C}}\sum_{\substack{C\\C\neq A}} m_Bm_C(\partial_k^Br_{BC}^{-1}\partial_k^B\partial_n^Br_{AB} - 8r_{AB}^{-1}\partial_n^Br_{BC}^{-1}$$

$$\left.\left.- 2r_{BC}^{-1}\partial_n^Br_{AB}^{-1} - 8r_{AB}^{-1}\partial_n^Cr_{AC}^{-1}\right)\right\}.$$

For two bodies ($N = 2$), we obtain from relation (3.17)

$$\frac{d^2\xi_1^n}{dt^2} = km_2\partial_n^1 r_{12}^{-1} + c^{-2}\left\{km_2\left[\frac{1}{2}V_2^kV_2^s\partial_n^1\partial_k^1\partial_s^2 r_{12}\right.\right.$$
$$+ (4V_2^nV_1^k - 4V_1^nV_1^k - 4V_2^nV_2^k + 3V_1^nV_2^k)\partial_k^1 r_{12}^{-1}$$
$$\left. + \left(V_1^2 - 4V_1^kV_2^k + \frac{3}{2}V_2^2\right)\partial_n^1 r_{12}^{-1}\right. \quad (3.18)$$
$$\left.\left. - k(5m_1 + 4m_2)r_{12}^{-1}\partial_n^1 r_{12}^{-1}\right]\right\}.$$

Substituting index 2 for 1 into this equation, we obtain the equation of motion for the second body. Equations (3.18) were derived for the first time by Einstein, Infeld, and Hoffmann [7] from the field equations in the coordinates defined by the conditions

$$\partial_m\gamma_{mn} = 0, \quad \partial_m\gamma_{m0} - c^{-1}\dot{\gamma}_{00} = 0, \quad (3.19)$$

where

$$\gamma_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta^{\alpha\beta}h_{\alpha\beta}\eta_{\mu\nu};$$

and by Petrova [4], who used the Fock method in the harmonic coordinates.

4. SOLUTION OF FIELD EQUATIONS IN THE POST-NEWTONIAN APPROXIMATION

Taking into account solutions (3.2) and (3.12), as well as formulas (2.3), (2.4), (2.7), and (2.8), we obtain from Eq. (2.1)

$$\Delta\underset{(4)}{h_{mn}} + 4\Phi\Delta\Phi\delta_{mn} + 4\partial_k\Phi\partial_k\Phi\delta_{mn}$$
$$- 4\partial_m\Phi\partial_n\Phi + 2\ddot{\Phi}\delta_{mn} = 16\pi k \quad (4.1)$$
$$\times\sum_A m_A\hat{\delta}(\bar{r}_A)\left[V_A^mV_A^n - \frac{1}{4}V_A^2\delta_{mn} - \frac{1}{2}\Phi\delta_{mn}\right].$$

We transform these equations using the formulas

$$\Delta\ln S_{Ab} = r_A^{-1}r_B^{-1}, \quad S_{AB} = r_A + r_B + r_{AB}, \quad (4.2)$$

$$\Delta\ln r_A = r_A^{-2}, \quad (4.3)$$

$$\partial_a\partial_b r_A^{-1} = r_A^{-3}(3N_A^aN_A^b - \delta_{ab}) - \frac{4}{3}\pi\hat{\delta}(\bar{r}_A)\delta_{ab}, \quad (4.4)$$

as well as the identity following from (A.3) for $L = 2$,

$$\partial_a\partial_b r_A^{-2} = 2\partial_a r_A^{-1}\partial_b r_A^{-1} + 2r_A^{-1}\partial_a\partial_b r_A^{-1}. \quad (4.5)$$

Using expressions (4.2)–(4.5) and the first expression from formula (3.3), as well as relation (A.13), we obtain

$$\partial_m\Phi\partial_n\Phi = \Delta\left[k^2\sum_A\sum_{B\neq A}m_Am_B\partial_m^A\partial_n^B\ln S_{AB}\right.$$
$$\left. + \frac{1}{8}k^2\sum_A m_A^2(\partial_m\partial_n\ln r_A + r_A^{-2}\delta_{mn})\right] \quad (4.6)$$
$$+ \frac{4}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)r_A^{-1}\delta_{mn}.$$

Using expressions (3.5) and (4.6), we obtain from Eq. (4.1)

$$\Delta\left[\underset{(4)}{h_{mn}} + \ddot{\chi}\delta_{mn} + 2\Phi^2\delta_{mn} - 4k^2\sum_A\sum_{B\neq A}m_Am_B\partial_m^A\partial_n^B\ln S_{AB}\right.$$
$$\left. - \frac{1}{2}k^2\sum_A m_A^2(\partial_m\partial_n\ln r_A + r_A^{-2}\delta_{mn})\right] \quad (4.7)$$
$$= \frac{16}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)r_A^{-1}\delta_{mn}$$
$$+ 16\pi k\sum_A m_A\hat{\delta}(\bar{r}_A)\left[V_A^nV_A^m - \frac{1}{4}V_A^2\delta_{mn} - \frac{1}{2}\Phi\delta_{mn}\right].$$

Taking into account formula (3.10), we can write the solution to Eq. (4.7) in the form

$$\underset{(4)}{h_{mn}} = -\ddot{\chi}\delta_{mn} - 2\Phi^2\delta_{mn}$$
$$+ 4k^2\sum_A\sum_{B\neq A}m_Am_B\partial_m^A\partial_n^B\ln S_{AB}$$
$$+ \frac{1}{2}k^2\sum_A m_A^2(\partial_m\partial_n\ln r_A + r_A^{-2}\delta_{mn}) \quad (4.8)$$
$$- k\sum_A m_A\left[4r_A^{-1}V_A^mV_A^n - 2\left(\frac{A}{\Phi} + \frac{1}{2}V_A^2\right)r_A^{-1}\delta_{mn}\right].$$

Using solution (3.2) and formulas (2.3), (2.4), (2.7), and (2.8), we obtain from Eq. (2.1)

$$\Delta\underset{(5)}{h_{0n}} - 4\ddot{\Phi}_n + 16\partial_k\Phi\partial_n\Phi_k + 12\dot{\Phi}\partial_n\Phi - 8\Phi\Delta\Phi_n$$
$$= 16\pi k\sum_A m_A\hat{\delta}(\bar{r}_A)\left[\Phi V_A^n + 2\Phi_n - \frac{1}{2}V_A^2V_A^n\right]. \quad (4.9)$$

Analogously to Eq. (4.6), we transform consecutively the third and fourth terms on the left-hand side of Eq. (4.9):

$$\partial_k\Phi\partial_n\Phi_k = \Delta\Bigg[k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A m_B\partial_k^A\partial_n^B\ln S_{AB}V_A^k + \frac{1}{8}k^2\sum_A m_A^2(\partial_k\partial_n\ln r_A + r_A^{-2}\delta_{kn})V_A^k\Bigg] + \frac{4}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)r_A^{-1}V_A^n, \tag{4.10}$$

$$\dot{\Phi}\partial_n\Phi = \Delta\Bigg[-k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A m_B\partial_k^A\partial_n^B\ln S_{AB}V_A^k - \frac{1}{8}k^2\sum_A m_A^2(\partial_k\partial_n\ln r_A + r_A^{-2}\delta_{kn})V_A^k\Bigg] - \frac{4}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)r_A^{-1}V_A^n. \tag{4.11}$$

In addition, we have

$$\Phi_n = \frac{1}{2}\Delta\chi_n,\quad \chi_n = k\sum_A m_A V_A^n r_A, \tag{4.12}$$

$$\ddot{\Phi}_n = \frac{1}{2}\Delta\ddot{\chi}_n. \tag{4.13}$$

Substituting expressions (4.10), (4.11), and (4.13) into Eq. (4.9), we obtain

$$\Delta\Bigg[\underset{(5)}{h_{0n}} - 2\ddot{\chi}_n + 4k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A m_B\partial_k^A\partial_n^B\ln S_{AB}(4V_B^k - 3V_A^k) + \frac{1}{2}k^2\sum_A m_A^2(\partial_k\partial_n\ln r_A + r_A^{-2}\delta_{kn})V_A^k\Bigg] = -\frac{16}{3}\pi k^2\sum_A m_A^2\delta(\bar{r}_A)r_A^{-1}V_A^n + 16\pi k\sum_A m_A\delta(\bar{r}_A)\Big[2\Phi_n - \Phi V_A^n - \frac{1}{2}V_A^2V_A^n\Big]. \tag{4.14}$$

Using formula (3.11) and similar formulas for Φ_n, we obtain the solution to Eq. (4.14) in the form

$$\underset{(5)}{h_{0n}} = 2\ddot{\chi}_n + 2k\sum_A m_A r_A^{-1}V_A^2V_A^n - \frac{1}{2}k^2\sum_A m_A^2(\partial_k\partial_n\ln r_A + r_A^{-2}\delta_{kn})V_A^k - 4k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A m_B[\partial_k^A\partial_n^B\ln S_{AB}(4V_B^k - 3V_A^k) - r_A^{-1}r_{AB}^{-1}(V_A^n - 2V_B^n)]. \tag{4.15}$$

It is slightly more difficult to evaluate $\underset{(6)}{h_{00}}$. Taking into account expressions (3.2), (3.12), (4.8), and (4.15), as well as (2.2)–(2.4), (2.7), and (2.8), we obtain from Eq. (2.1)

$$\begin{aligned}\Delta\underset{(6)}{h_{00}} &+ 12\dot{\Phi}\dot{\Phi} - 16\partial_k\Phi_s\partial_s\Phi_k + 4\Phi\ddot{\Phi} + 16\Phi_k\partial_k\dot{\Phi} - 4\partial_k\Phi\partial_k\ddot{\chi} + 16\partial_k\Phi_s\partial_k\Phi_s + 8\Phi\partial_k\Phi\partial_k\Phi + (\chi)^{\cdots\cdot} - 12\Phi^2\Delta\Phi + 2\ddot{\chi}\Delta\Phi\\ &+ 8k\sum_A m_A r_A^{-1}\partial_k\partial_s\Phi V_A^kV_A^s + 4k\sum_A m_A(2\overset{A}{\Phi} - 3V_A^2)\partial_k\Phi\partial_k r_A^{-1} - \Bigg[k\sum_A m_A(2\overset{A}{\Phi} - 3V_A^2)r_A^{-1}\Bigg]^{\cdot\cdot} + 4k^3\sum_A m_A^3r_A^{-5}\\ &- 2k\sum_A m_A(2\overset{A}{\Phi} - 3V_A^2)\Phi\Delta r_A^{-1} - 4k\sum_A m_A\Big(\overset{A}{\Phi} + \frac{1}{2}V_A^2\Big)r_A^{-1}\Delta\Phi - k^3\sum_A\sum_{\substack{B\\B\neq A}} m_A^2m_B\partial_k\partial_s r_B^{-1}\partial_k\partial_s\ln r_A\\ &- 8k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A m_B\partial_k\partial_s\Phi\partial_k^A\partial_s^B\ln S_{AB}\end{aligned} \tag{4.16}$$

$$= 16\pi k\sum_A m_A\hat{\delta}(\bar{r}_A)\Bigg\{\frac{25}{4}\Phi^2-\frac{3}{4}\Phi V_A^2-\frac{5}{4}\ddot{\chi}+2\Phi_k V_A^k$$

$$+\frac{7}{16}V_A^4+\frac{1}{2}k\sum_{\substack{B\\B\neq A}} m_B\Big(5\overset{B}{\bar{\Phi}}-\frac{7}{2}V_B^2\Big)r_B^{-1}$$

$$+\frac{1}{2}km_A\Big(5\overset{A}{\bar{\Phi}}-\frac{7}{2}V_A^2\Big)r_A^{-1}+\frac{1}{4}k^2\sum_{\substack{B\\B\neq A}} m_B^2r_B^{-2}+\frac{1}{2}k^2m_A^2r_A^{-2}$$

$$+\frac{1}{2}k^2\sum_B\sum_{\substack{C\\C\neq B}} m_Bm_C(r_B^{-1}r_C^{-1}-r_B^{-1}r_{BC}^{-1}-r_C^{-1}r_{BC}^{-1})\Bigg\}.$$

Using formulas (4.2)–(4.4), as well as (A.9) and (A.13), we can transform the expressions on the left-hand side of Eq. (4.16). Indeed, in the increasing numbers of terms on the left-hand side of Eq. (4.16), we obtain

$$\dot{\Phi}\dot{\Phi} = \Delta\Bigg\{k^2\sum_A\sum_{\substack{B\\B\neq A}} m_Am_B\partial_k^A\partial_s^B\ln S_{AB}V_A^kV_B^s$$

$$+\frac{1}{8}k^2\sum_A m_A^2(\partial_k\partial_s\ln r_AV_A^sV_A^k+r_A^{-2}V_A^2)\Bigg\} \qquad (4.17)$$

$$+\frac{4}{3}\pi k^2\sum_A m_A^2r_A^{-1}\hat{\delta}(\bar{r}_A)V_A^2$$

for the second term,

$$\partial_k\Phi_s\partial_s\Phi_k = \Delta\Bigg\{k^2\sum_A\sum_{\substack{B\\B\neq A}} m_Am_B\partial_k^B\partial_s^A\ln S_{AB}V_A^kV_B^s$$

$$+\frac{1}{8}k^2\sum_A m_A^2(\partial_k\partial_s\ln r_AV_A^kV_A^s+r_A^{-2}V_A^2)\Bigg\} \qquad (4.18)$$

$$+\frac{4}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)V_A^2r_A^{-1}$$

for the third term,

$$\Phi\ddot{\Phi} = \Delta\Bigg\{k^2\sum_A\sum_{\substack{B\\B\neq A}} m_Am_B[\partial_k^B\partial_s^B\ln S_{AB}V_B^kV_B^s$$

$$+\partial_k^B\ln S_{AB}W_B^k]+k^2\sum_A m_A^2\Big[\frac{3}{8}\Big(\partial_k\partial_s\ln r_AV_A^kV_A^s-\frac{1}{3}r_A^{-2}V_A^2\Big)$$

$$-\frac{1}{2}\partial_k\ln r_AW_A^k\Big]\Bigg\}-\frac{4}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)V_A^2r_A^{-1} \qquad (4.19)$$

for the fourth term,

$$\Phi_k\partial_k\dot{\Phi} = \Delta\Bigg[-k^2\sum_A\sum_{\substack{B\\B\neq A}} m_Am_B\partial_k^B\partial_s^B\ln S_{AB}V_A^sV_B^k$$

$$-\frac{3}{8}k^2\sum_A m_A^2\Big(\partial_k\partial_s\ln r_AV_A^kV_A^s-\frac{1}{3}r_A^{-2}V_A^2\Big)\Bigg] \qquad (4.20)$$

$$+\frac{4}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)V_A^2r_A^{-1}$$

for the fifth term, and

$$k\sum_A m_Ar_A^{-1}\partial_k\partial_s\Phi V_A^kV_A^s$$

$$= \Delta\Bigg\{k^2\sum_A\sum_{\substack{B\\B\neq A}} m_Am_B\partial_k^B\partial_s^B\ln S_{AB}V_A^kV_A^s$$

$$+\frac{3}{8}k^2\sum_A m_A^2\Big(\partial_k\partial_s\ln r_AV_A^kV_A^s-\frac{1}{3}r_A^{-2}V_A^2\Big)\Bigg\} \qquad (4.21)$$

$$-\frac{4}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)V_A^2r_A^{-1}$$

for the twelfth term.

We transform the sixth, seventh, eighth, and thirteenth terms on the left-hand side of formula (4.16) tak-

ing into account expressions (3.5) and (4.12). In accordance with formulas (A.1)–(A.6), we have

$$\begin{aligned}
\partial_k\Phi\partial_k\ddot{\chi} &= \frac{1}{2}\Delta(\Phi\ddot{\chi}) - \frac{1}{2}\ddot{\chi}\Delta\Phi - \frac{1}{2}\Phi\Delta\ddot{\chi} \\
&= \frac{1}{2}\Delta(\Phi\ddot{\chi}) - \frac{1}{2}\ddot{\chi}\Delta\Phi - \Phi\ddot{\Phi}, \\
\partial_k\Phi_s\partial_k\Phi_s &= \frac{1}{2}\Delta\Phi_k^2 - \Phi_k\Delta\Phi_k, \\
\Phi\partial_k\Phi\partial_k\Phi &= \frac{1}{6}\Delta\Phi^3 - \frac{1}{2}\Phi\Delta\Phi, \\
k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)\partial_k\Phi\partial_k r_A^{-1} &= \frac{1}{2}\Delta\left\{k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)\Phi r_A^{-1}\right\} \\
&\quad - \frac{1}{2}k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)r_A^{-1}\Delta\Phi \\
&\quad - \frac{1}{2}k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)\Phi\Delta r_A^{-1}.
\end{aligned} \tag{4.22}$$

Using formula (3.5), we obtain the following expression for the fourteenth term on the left-hand side of Eq. (4.16):

$$\begin{aligned}
&\left[k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)r_A^{-1}\right]^{\cdot\cdot} \\
&= \frac{1}{2}\Delta\left[k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)r_A\right]^{\cdot\cdot}.
\end{aligned} \tag{4.23}$$

Taking into account formula (A.13), we can write the ninth and fifteenth terms on the left-hand side of Eq. (4.16) in the form

$$\begin{aligned}
(\chi)^{\cdot\cdot\cdot\cdot} &= -\frac{1}{12}\Delta k\sum_A m_A(r_A^3)^{\cdot\cdot\cdot\cdot}, \\
k^3\sum_A m_A^3 r_A^{-5} &= \frac{1}{6}\Delta k^3\sum_A m_A^3 r_A^{-3} \\
&\quad + 2\pi k^3\sum_A m_A^3 r_A^{-2}\hat{\delta}(\bar{x} - \bar{\xi}_A).
\end{aligned} \tag{4.24}$$

To transform the expression (see [12])

$$\sum_A\sum_{\substack{B\\B\neq A}} m_A^2 m_B\partial_k\partial_s r_B^{-1}\partial_k\partial_s\ln r_A,$$

we again take into account identity (A.14) as well as the dependences

$$\begin{aligned}
\hat{\delta}(\bar{r}_A)\ln r_B &= \hat{\delta}(\bar{r}_A)\ln r_{AB}, \quad A\neq B, \\
N_A^k N_B^k &= \frac{1}{2}(r_A^{-1}r_B + r_B^{-1}r_A - r_{AB}^2 r_A^{-1}r_B^{-1}).
\end{aligned} \tag{4.25}$$

This gives

$$\begin{aligned}
&\sum_A\sum_{\substack{B\\B\neq A}} m_A^2 m_B\partial_k\partial_s r_B^{-1}\partial_k\partial_s\ln r_A \\
&= \frac{1}{2}\Delta\left\{\sum_A\sum_{\substack{B\\B\neq A}} m_A^2 m_B\left[\partial_k^A\partial_k^B(r_B^{-1}\ln r_A - r_B^{-1}\ln r_{AB})\right.\right. \\
&\quad \left.\left. - \frac{1}{2}r_B^{-1}r_A^{-2} + \frac{1}{2}r_{AB}^{-2}r_B^{-1} + \frac{1}{2}r_B r_A^{-2}r_{AB}^{-2}\right]\right\} \\
&\quad + 4\pi\hat{\delta}(\bar{r}_A)r_A^{-1}(r_B r_{AB}^{-2} - r_B^{-1}).
\end{aligned} \tag{4.26}$$

The remaining terms in Eq. (4.16) (except the nineteenth term) contain the $\hat{\delta}$ function in the integrand. Substituting relations (4.17)–(4.24) and (4.26) into Eq. (4.16), we obtain the final equation for $\underset{(6)}{h_{00}}$ in the form

$$\begin{aligned}
&\Delta\left\{\underset{(6)}{h_{00}} + 8\Phi_b^2 + \frac{4}{3}\Phi^3 - 2\Phi\ddot{\chi} + \frac{1}{12}k\sum_A m_A(r_A^3)^{\cdot\cdot\cdot\cdot}\right. \\
&\quad + \frac{2}{3}k^3\sum_A m_A^3 r_A^{-3} - 4k^2\sum_A m_A^2\partial_k\ln r_A W_A^k \\
&\quad \left. - \frac{1}{2}k^2\sum_A m_A^2(\partial_k\partial_s\ln r_A V_A^k V_A^s + r_A^{-2}V_A^2)\right. \\
&\quad + 2k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)\Phi r_A^{-1} \\
&\quad \left. - \frac{1}{2}\left[k\sum_A m_A(2\overset{A}{\bar{\Phi}} - 3V_A^2)r_A\right]^{\cdot\cdot}\right.
\end{aligned}$$

$$-\frac{1}{2}k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A^2 m_B\Big[\partial_k^A\partial_k^B(r_B^{-1}\ln r_A - r_B^{-1}\ln r_{AB})$$

$$-\frac{1}{2}r_B^{-1}r_A^{-2}+\frac{1}{2}r_{AB}^{-2}r_B^{-1}+\frac{1}{2}r_B r_A^{-2}r_{AB}^{-2}\Big]$$

$$-k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A m_B[(16V_A^sV_B^k-12V_A^kV_B^s)\partial_k^A\partial_s^B\ln S_{AB}$$

$$+(16V_B^sV_A^k-8V_B^sV_B^k-8V_A^sV_A^k)\partial_k^B\partial_s^B\ln S_{AB}$$

$$-8\partial_k^B\ln S_{AB}W_B^k]\Big\} = 16\pi k\sum_A m_A\hat{\delta}(\bar{r}_A) \qquad (4.27)$$

$$\times\Big\{\frac{1}{4}\Phi^2+\frac{9}{4}\Phi V_A^2-\frac{1}{4}\ddot{\chi}-2\Phi_n V_A^n+\frac{7}{16}V_A^4$$

$$+\frac{1}{2}k\sum_{\substack{B\\B\neq A}} m_B\Big(\overset{B}{\bar{\Phi}}-\frac{3}{2}V_B^2\Big)r_B^{-1}+\frac{1}{4}k^2\sum_{\substack{B\\B\neq A}} m_B^2 r_B^{-2}$$

$$-2k^2\sum_{\substack{B\\B\neq A}} m_A m_B r_B^{-1}r_{AB}^{-1}$$

$$+\frac{1}{2}k^2\sum_{\substack{B\\B\neq A}}\sum_{\substack{C\\C\neq A\\C\neq B}} m_C m_B(r_B^{-1}r_C^{-1}-r_B^{-1}r_{BC}^{-1}-r_C^{-1}r_{BC}^{-1})$$

$$+\frac{1}{8}k^2\sum_{\substack{B\\B\neq A}} m_A m_B r_A^{-1}(r_B r_{AB}^{-2}+15r_B^{-1}-16r_{AB}^{-1})$$

$$+km_A\Big(\overset{A}{\bar{\Phi}}-\frac{3}{2}V_A^2\Big)r_A^{-1}+\frac{1}{2}k^2m_A^2r_A^{-2}\Big\}$$

$$-16\pi\sum_A m_A\partial_a\hat{\delta}(\bar{r}_A)\dot{\chi}_a+\frac{16}{3}\pi k^2\sum_A m_A^2\hat{\delta}(\bar{r}_A)V_A^2r_A^{-1}$$

$$-8\pi k^3\sum_A m_A^3\hat{\delta}(\bar{r}_A)r_A^{-2}$$

$$+8k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A m_B\partial_k\partial_s\Phi\partial_k^A\partial_s^B\ln S_{AB}.$$

To solve this equation, we must carry out regularization of the following integrals containing singular functions:

$$\int\hat{\delta}(\bar{r}'_A)r'^{-1}_A N'^k_A N'^s_A|\bar{x}-\bar{x}'|^{-1}(d\bar{x}'),$$

$$\int\hat{\delta}(\bar{r}'_A)r'^{-1}_A r'^{-1}_B|\bar{x}-\bar{x}'|^{-1}(d\bar{x}'), \qquad (4.28)$$

$$\int\hat{\delta}(\bar{r}'_A)r'^{-2}_A|\bar{x}-\bar{x}'|^{-1}(d\bar{x}').$$

It follows from expansions (3.9) and (3.16), as well as from property (2.6) of the $\hat{\delta}$ function, that integrals (4.28) are equal to zero.

Thus, the solution to Eq. (4.27) can be written in the form

$$\underset{(6)}{h_{00}} = 2\Phi\ddot{\chi}-8\Phi_k^2-\frac{4}{3}\Phi^3-\frac{1}{12}k\sum_A m_A(r_A^3)^{\cdots\cdot}$$

$$-\frac{2}{3}k^3\sum_A m_A^3r_A^{-3}-\frac{7}{4}k\sum_A m_AV_A^4r_A^{-1}$$

$$-2k\sum_A m_A(2\overset{A}{\bar{\Phi}}-3V_A{}^2)\Phi r_A^{-1}+4k^2\sum_A m_A^2\partial_k\ln r_A W_A^k$$

$$+\frac{1}{2}k^2\sum_A m_A^2(\partial_k\partial_s\ln r_A V_A^kV_A^s+r_A^{-2}V_A^2)$$

$$+8k\sum_A m_A\overset{A}{\bar{\Phi}}_k V_A^kr_A^{-1}-k\sum_A m_A\overset{A}{\bar{\Phi}}(\overset{A}{\bar{\Phi}}+9V_A^2)r_A^{-1}$$

$$+\frac{1}{2}\Big[k\sum_A m_A(2\overset{A}{\bar{\Phi}}-3V_A^2)r_A\Big]^{\cdot\cdot}+k\sum_A m_A\overset{A}{\ddot{\bar{\chi}}}r_A^{-1}$$

$$-2k^2\sum_A\sum_{\substack{B\\B\neq A}} m_Am_B\Big(\overset{B}{\bar{\Phi}}-\frac{3}{2}V_B^2\Big)r_{AB}^{-1}r_A^{-1}$$

$$+4k^3\sum_A\sum_{\substack{B\\B\neq A}} m_A^2m_Br_{AB}^{-2}r_A^{-1}-2k^3\sum_A\sum_{\substack{B\\B\neq A}} m_Am_B^2r_{AB}^{-2}r_A^{-1} \qquad (4.29)$$

$$+k^2\sum_A\sum_{\substack{B\\B\neq A}} m_Am_B[(16V_A^sV_B^k-12V_A^kV_B^s)\partial_k^A\partial_s^B\ln S_{AB}$$

$$+(16V_B^sV_A^k-8V_B^sV_B^k-8V_A^sV_A^k)\partial_k^B\partial_s^B\ln S_{AB}$$

$$-8\partial_k^B\ln S_{AB}W_B^k]+\frac{1}{2}k^3\sum_A\sum_{\substack{B\\B\neq A}} m_A^2m_B$$

$$\times\Big[\partial_k^A\partial_k^B(r_B^{-1}\ln r_A - r_B^{-1}\ln r_{AB}) - \frac{1}{2}r_B^{-1}r_A^{-2} + \frac{3}{2}r_{AB}^{-2}r_B^{-1} + \frac{1}{2}r_B r_A^{-2}r_{AB}^{-2}\Big]$$

$$+ 4k^3\sum_A\sum_{\substack{B\\B\neq A}}\sum_{\substack{C\\C\neq A\\C\neq B}} m_A m_B m_C r_{AB}^{-1}r_{BC}^{-1}r_A^{-1} - 2k^3\sum_A\sum_{\substack{B\\B\neq A}}\sum_{\substack{C\\C\neq A}} m_A m_B m_C r_{AB}^{-1}r_{AC}^{-1}r_A^{-1} + H,$$

where H is the solution to the equation

$$\Delta H = 8k^2\sum_A\sum_{\substack{B\\B\neq A}} m_A m_B \partial_k\partial_s\Phi\partial_k^A\partial_s^B\ln S_{AB}. \quad (4.30)$$

We will carry out subsequent calculations for a system of two bodies ($N = 2$). In this case, function H has the form [12]

$$H = k^3 m_1^2 m_2\{4\Delta_1\partial_k^1\partial_k^2[(r_1 + r_{12})\ln S_{12}] + 8\partial_k^1\ln S_{12}\partial_k^1 r_{12}^{-1} - 4r_2 r_1^{-2}r_{12}^{-2} + 4r_{12}^{-2}r_1^{-1} - 2r_{12}^{-1}r_1^{-2} + 4r_2^2 r_1^{-3}r_{12}^{-2} + 6r_2^2 r_1^{-2}r_{12}^{-3} - 4r_1^{-3} - 6r_{12}^{-13}\} + (1\longleftrightarrow 2). \quad (4.31)$$

The above expression for two bodies was derived taking into account identity (A.14) and formulas (4.2) and (4.3), as well as the expression $\hat\delta(\bar r_1)r_2 = \hat\delta(\bar r_1)r_{12}$ and the dependences

$$\Delta_1\ln S_{12} = r_1^{-1}r_{12}^{-1},$$

$$S_{12}^{-1}(1 + N_1^k N_{12}^k) = \frac{1}{2}(r_2^{-1} + r_1^{-1} - r_{12}r_A^{-1}r_2^{-1}),$$

$$\partial_k^1\partial_k^2\ln S_{12} = \frac{1}{2}(r_1^{-1}r_2^{-1} - r_1^{-1}r_{12}^{-1} - r_2^{-1}r_{12}^{-1}).$$

Finally, for the two bodies in question, solutions (3.2), (3.12), (4.8), (4.15), and (4.30) assume the form

$$h_{00}^{(2)} = -2km_1k_1^{-1} - 2km_2r_2^{-1},$$

$$h_{mn}^{(2)} = -2km_1r_1^{-1}\delta_{mn} - 2km_2r_2^{-1}\delta_{mn}, \quad (4.32)$$

$$h_{0n}^{(3)} = 4km_1r_1^{-1}V_1^n + 4km_2r_2^{-1}V_2^n,$$

$$h_{00}^{(4)} = km_1r_1^{-1}[(N_1V_1)^2 - 4V_1^2] + 2k^2m_1^2r_1^{-2} + k^2m_1m_2\Big[2r_1^{-1}r_2^{-1} + \frac{1}{2}r_1r_{12}^{-3} + \frac{5}{2}r_2^{-1}r_{12}^{-1} - \frac{1}{2}r_1^2r_2^{-1}r_{12}^{-3}\Big] + (1\longleftrightarrow 2), \quad (4.33)$$

$$h_{mn}^{(4)} = km_1r_1^{-1}(N_1V_1)^2\delta_{mn} - 4km_1r_1^{-1}V_1^mV_1^n - k^2m_1^2r_1^{-2}[\delta_{mn} + N_1^mN_1^n] + k^2m_1m_2\Big[\Big(\frac{1}{2}r_1r_{12}^{-3} + \frac{5}{2}r_1^{-1}r_{12}^{-1} - 2r_1^{-1}r_2^{-1} - \frac{1}{2}r_1^2r_2^{-1}r_{12}^{-3} - 4r_{12}^{-1}S_{12}^{-1}\Big)\delta_{mn} + 4N_{12}^mN_{12}^n(S_{12}^{-2} + r_{12}^{-1}S_{12}^{-1}) - 4S_{12}^{-2}(N_1^mN_2^n + N_1^nN_2^m + 2N_1^mN_{12}^n + 2N_1^nN_{12}^m)\Big] + (1\longleftrightarrow 2), \quad (4.34)$$

$$h_{0n}^{(5)} = 4km_1r_1^{-1}V_1^nV_1^2 + k^2m_1^2r_1^{-2}[(N_1V_1)N_1^n - V_1^n] + k^2m_1m_2\{N_1^nS_{12}^{-2}[16(N_{12}V_1) + 16(N_2V_1) - 12(N_{12}V_2) - 12(N_2V_2)] + N_{12}^n[16S_{12}^{-2}(N_1V_2) - 12S_{12}^{-2}(N_1V_1) - 4S_{12}^{-2}(N_{12}V_1) - 4S_{12}^{-1}r_{12}^{-1}(N_{12}V_1)] + V_1^n[r_2^2r_1^{-1}r_{12}^{-3} + 3r_1^{-1}r_{12}^{-1} - 8r_2^{-1}r_{12}^{-1} + 4r_{12}^{-1}S_{12}^{-1}]\} - k^2m_1m_2[6(N_{12}V_1)^2N_{12}^n + 4r_{12}^{-2}N_{12}^n] + k^2m_1m_2V_1^n(2r_2r_{12}^{-3} - 3r_1^{-1}r_{12}^{-3}) + (1\longleftrightarrow 2), \quad (4.35)$$

$$h_{00}^{(6)} = km_1r^{-1}\Big[3(N_1V_1)^2V_1^2 - 4V_1^4 - \frac{3}{4}(N_1V_1)^4\Big] + k^2m_1^2r_1^{-2}[V_1^2 - 3(N_1V_1)^2] + k^2m_1m_2\Big\{V_1^2\Big[\frac{3}{8}r_1^2r_2r_{12}^{-5} - \frac{3}{8}r_1^3r_{12}^{-15} + \frac{3}{8}r_1r_2^2r_{12}^{-5} - \frac{3}{8}r_2^3r_{12}^{-5} + \frac{37}{8}r_1r_{12}^{-3} - r_1^2r_2^{-1}r_{12}^{-3} - \frac{3}{8}r_2r_{12}^{-3} - 2r_2^2r_1^{-1}r_{12}^{-3} - 6r_1^{-1}r_{12}^{-1} + 5r_2^{-1}r_{12}^{-1} + 8r_{12}r_1^{-1}r_2^{-1}S_{12}^{-1} - 16r_{12}^{-1}S_{12}^{-1}\Big] + (V_1V_2)\Big[\frac{3}{4}r_1^3r_{12}^{-5} - 8r_1^{-1}r_2^{-1} - \frac{3}{4}r_1^2r_2r_{12}^{-5} - \frac{13}{4}r_1r_{12}^{-3} + 2r_1^2r_2^{-1}r_{12}^{-3} + 6r_1^{-1}r_{12}^{-1} + 16r_1^{-1}S_{12}^{-1} + 12r_{12}^{-1}S_{12}^{-1}\Big] + (N_{12}V_1)^2\Big[\frac{15}{8}r_1^3r_{12}^{-5} - \frac{15}{8}r_1^2r_2r_{12}^{-5} - \frac{15}{8}r_1r_2^2r_{12}^{-5} + \frac{15}{8}r_2^3r_{12}^{-5} - \frac{57}{8}r_1r_{12}^{-3} + \frac{3}{4}r_1^2r_2^{-1}r_{12}^{-3} + \frac{33}{8}r_2r_{12}^{-3}$$

$$-\frac{7}{4}r_2^{-1}r_{12}^{-1}+16S_{12}^{-2}+16r_{12}^{-1}S_{12}^{-1}\Big]$$

$$+(N_{12}V_1)(N_{12}V_2)\Big[\frac{15}{4}r_1^2r_2r_{12}^{-5}-\frac{15}{4}r_1^3r_{12}^{-5}$$

$$+\frac{9}{4}r_1r_{12}^{-3}-12S_{12}^{-2}-12r_{12}S_{12}^{-1}\Big]$$

$$+(N_1V_1)^2\Big[\frac{3}{4}r_2^2r_1^{-1}r_{12}^{-3}-2r_2^{-1}r_1^{-1}+\frac{1}{4}r_1r_{12}^{-3}$$

$$-\frac{7}{4}r_1^{-1}r_{12}^{-1}+8S_{12}^{-2}+8r_1^{-1}S_{12}^{-1}\Big] \qquad (4.36)$$

$$-(N_1V_1)(N_1V_2)[r_1r_{12}^{-3}+16S_{12}^{-2}+16r_1^{-1}S_{12}^{-1}]$$

$$+(N_1V_2)^2[8S_{12}^{-2}+8r_1^{-1}S_{12}^{-1}]$$

$$+(N_{12}V_1)(N_1V_1)\Big[3r_1^2r_{12}^{-4}-\frac{3}{4}r_2^2r_{12}^{-4}-\frac{3}{2}r_{12}^{-2}-16S_{12}^{-2}\Big]$$

$$+(N_{12}V_2)(N_1V_1)\Big[-3r_1^2r_{12}^{-4}+\frac{3}{2}r_2^2r_{12}^{-4}-\frac{13}{2}r_{12}^{-2}+40S_{12}^{-2}\Big]$$

$$+(N_{12}V_1)(N_1V_2)\Big[-\frac{3}{2}r_1^2r_{12}^{-4}-4r_{12}^{-2}-16S_{12}^{-2}\Big]$$

$$+(N_{12}V_2)(N_1V_2)\Big[\frac{3}{2}r_1^2r_{12}^{-4}+3r_{12}^{-2}-16S_{12}^{-2}\Big]$$

$$-16S_{12}^{-2}(N_1V_2)(N_2V_1)+12S_{12}^{-2}(N_1V_1)(N_2V_2)$$

$$-2k^3m_1^3r_1^{-3}+k^3m_1^2m_2\Big[\frac{1}{4}r_1^3r_{12}^{-6}-4r_1^{-3}-\frac{1}{2}r_2^{-3}-\frac{9}{2}r_1^{-2}r_2^{-1}$$

$$-\frac{3}{16}r_1^4r_2^{-1}r_{12}^{-6}+\frac{1}{8}r_1^2r_2r_{12}^{-6}-\frac{1}{4}r_2^2r_1r_{12}^{-6}+\frac{1}{16}r_2^3r_{12}^{-6}-\frac{5}{4}r_1r_{12}^{-4}$$

$$+\frac{23}{8}r_1^2r_2^{-1}r_{12}^{-4}-\frac{43}{8}r_2r_{12}^{-4}+\frac{5}{2}r_2^2r_1^{-1}r_{12}^{-4}+3r_{12}^{-3}-3r_1r_2^{-1}r_{12}^{-3}$$

$$-r_2r_1^{-1}r_{12}^{-3}+5r_2^2r_1^{-2}r_{12}^{-3}-4r_2^3r_1^{-3}r_{12}^{-3}+\frac{3}{2}r_1^{-1}r_{12}^{-2}+\frac{1}{4}r_1^2r_2^{-3}r_{12}^{-2}$$

$$-\frac{3}{16}r_2^{-1}r_{12}^{-2}-\frac{15}{4}r_2r_1^{-2}r_{12}^{-2}+4r_2^2r_1^{-3}r_{12}^{-2}-5r_1^{-2}r_{12}^{-1}$$

$$-5r_1^{-1}r_2^{-1}r_{12}^{-1}+4r_2r_1^{-3}r_{12}^{-1}+\frac{1}{4}r_1^{-2}r_2^{-3}r_{12}^2\Big]\Big\}+(1\longleftrightarrow 2),$$

where

$$(N_AV_A)\equiv N_A^kV_A^k,\quad (N_AV_B)\equiv N_A^kV_B^k,$$

$$(N_{AB}V_B)\equiv N_{AB}^kV_B^k,\quad (V_AV_B)\equiv V_A^kV_B^k.$$

Solutions (4.32)–(4.36) obtained here coincide with the solutions given in [12].

5. POST-POST-NEWTONIAN EQUATIONS OF MOTION

The equations of motion in the post-post-Newtonian approximation can be written in the form

$$\frac{d^2\xi_1^n}{dt^2}=\underset{(0)}{F_1^n}+c^{-2}\underset{(2)}{F_1^{(n)}}+c^{-4}[\underset{(4)}{F_1^{'n}}+\underset{(4)}{F_1^{''n}}+\underset{(4)}{F_1^{'''n}}], \qquad (5.1)$$

where the equalities

$$\underset{(4)}{F_1^{'n}}=-\frac{1}{2}km_2\overset{1}{\overline{\partial_n\partial_kr_2\underset{(2)}{W_2^k}}},\quad \underset{(4)}{F_1^{''n}}=4km_2\overset{1}{\overline{r_2^{-1}\underset{(2)}{W_2^n}}} \qquad (5.2)$$

follow from the expressions for $\overset{1}{\overline{\partial_nh_{00}}}$ and $\overset{1}{\overline{\dot h_{0n}}}$, respectively, appearing in F_1^n.

For $\underset{(4)}{F_1^{'''n}}$, we obtain from Eqs. (2.5) and (2.7)

$$\underset{(4)}{F_1^{'''n}}=-\frac{1}{2}\overset{1}{\overline{\partial_n\underset{(6)}{h_{00}}}}-\frac{1}{2}\overset{1}{\overline{\underset{(2)}{h_{nk}}\,\partial_k\underset{(4)}{h_{00}}}}+\overset{1}{\overline{\underset{(5)}{\dot h_{0n}}}}+\frac{1}{2}\overset{1}{\overline{\underset{(4)}{\dot h_{00}}}}\,V_1^n$$

$$+\overset{1}{\overline{\underset{(4)}{\dot h_{nk}}}}\,V_1^k-\overset{1}{\overline{\partial_n\underset{(5)}{h_{0k}}}}\,V_1^k+\overset{1}{\overline{\partial_k\underset{(5)}{h_{0n}}}}\,V_1^k+\overset{1}{\overline{\partial_k\underset{(4)}{h_{00}}}}\,V_1^kV_1^n$$

$$-\frac{1}{2}\overset{1}{\overline{\underset{(4)}{\partial_nh_{ks}}}}\,V_1^kV_1^s+\overset{1}{\overline{\underset{(4)}{\partial_kh_{ns}}}}\,V_1^kV_1^s-\frac{1}{2}\overset{A}{\overline{\underset{(3)}{h_{0n}}\,\underset{(2)}{\dot h_{00}}}}+\overset{1}{\overline{\underset{(2)}{h_{nk}}\,\underset{(3)}{\dot h_{0k}}}}$$

$$+\overset{1}{\overline{\underset{(2)}{h_{nk}}\,\underset{(2)}{\dot h_{ks}}}}\,V_1^s-\frac{1}{2}\overset{1}{\overline{\underset{(4)}{h_{nk}}\,\partial_k\underset{(2)}{h_{00}}}}-\frac{1}{2}\overset{1}{\overline{\underset{(2)}{h_{00}}\,\underset{(2)}{\dot h_{00}}}}\,V_1^n$$

$$-\frac{1}{2}\overset{1}{\overline{\underset{(3)}{h_{0k}}\,\partial_k\underset{(2)}{h_{00}}}}\,V_1^n-\overset{1}{\overline{\underset{(2)}{h_{00}}\,\partial_k\underset{(2)}{h_{00}}}}\,V_1^kV_1^n-\overset{1}{\overline{\underset{(2)}{h_{0n}}\,\partial_k\underset{(2)}{h_{00}}}}\,V_1^k$$

$$+\overset{1}{\overline{\underset{(2)}{h_{nk}}\,\partial_s\underset{(3)}{h_{0k}}}}\,V_1^s-\overset{1}{\overline{\underset{(2)}{h_{nk}}\,\partial_k\underset{(3)}{h_{0s}}}}\,V_1^s+\overset{1}{\overline{\underset{(2)}{h_{na}}\,\partial_k\underset{(2)}{h_{as}}}}\,V_1^kV_1^s$$

$$+\overset{1}{\overline{\partial_k\underset{(3)}{h_{0s}}}}\,V_1^sV_1^kV_1^n-\frac{1}{2}\overset{1}{\overline{\underset{(2)}{\dot h_{ks}}}}\,V_1^kV_1^sV_1^n$$

$$-\frac{1}{2}\overset{1}{\overline{\underset{(2)}{h_{nk}}\,\underset{(2)}{h_{ks}}\,\partial_s\underset{(2)}{h_{00}}}}-\frac{1}{2}\overset{1}{\overline{\underset{(2)}{h_{na}}\,\partial_a\underset{(2)}{h_{ks}}}}\,V_1^kV_1^s. \qquad (5.3)$$

Using formula (3.18) and regularizing the diverging

integrals (see Appendix B), we obtain

$$F_{1(4)}^{\prime n} = -\frac{1}{2}k^2 m_1 m_2 [3N_{12}^n (N_{12}V_1)^2 + 4N_{12}^n (N_{12}V_2)^2$$
$$-7N_{12}^n (N_{12}V_1)(N_{12}V_2) + 4V_1^n (N_{12}V_2) - 4V_2^n (N_{12}V_2)$$
$$-3V_1^n (N_{12}V_1) + 3V_2^n (N_{12}V_1)],$$

$$F_{1(4)}^{\prime\prime n} = -16k^3 m_1^2 m_2 r_{12}^{-4} N_{12}^n - 20k^3 m_2^2 m_1 r_{12}^{-4} N_{12}^n \quad (5.4)$$
$$-k^2 m_1 m_2 r_{12}^{-3} [-8N_{12}^n V_1^2 - 4N_{12}^n V_2^2 + 16N_{12}^n (V_1 V_2)$$
$$+6N_{12}^n (N_{12}V_1)^2 + 16V_2^n (N_{12}V_2) - 16V_1^n (N_{12}V_2)$$
$$-12V_2^n (N_{12}V_1) + 12V_1^n (N_{12}V_1)].$$

To evaluate $F_{1(4)}^{\prime\prime\prime n}$, we will use solutions (4.33)–(4.37) as well as formula (5.3). After regularization (see Appendix B), we have

$$\frac{1}{2}\overline{\partial_n h_{00}}^{1}_{(6)} = 3k^3 m_2^3 r_{12}^{-4} N_{12}^n - \frac{7}{4}k^3 m_1^2 m_2 r_{12}^{-4} N_{12}^n$$
$$+\frac{23}{2}k^3 m_2^2 m_1 r_{12}^{-4} N_{12}^n$$
$$+k^2 m_2^2 r_{12}^{-3} [-N_{12}^n V_2^2 + 6N_{12}^n (N_{12}V_2)^2 - 3V_2^n (N_{12}V_2)]$$
$$+k^2 m_1 m_2 r_{12}^{-3} \Big[\frac{3}{4}N_{12}^n V_1^2 + \frac{35}{4}N_{12}^n V_2^2 - \frac{5}{2}N_{12}^n (V_1 V_2)$$
$$-\frac{23}{2}N_{12}^n (N_{12}V_1)^2 - \frac{37}{2}N_{12}^n (N_{12}V_2)^2$$
$$+25N_{12}^n (N_{12}V_2)(N_{12}V_1) + \frac{31}{4}V_2^n (N_{12}V_2)$$
$$-\frac{39}{4}V_1^n (N_{12}V_2) - \frac{3}{4}V_2^n (N_{12}V_1) + \frac{31}{4}V_1^n (N_{12}V_1)\Big]$$
$$+km_2 r_{12}^{-2} \Big[-\frac{9}{2}N_{12}^n (N_{12}V_2)^2 V_2^2 + 3V_2^n (N_{12}V_2)V_2^2$$
$$+2N_{12}^n V_A^4 + \frac{15}{8}N_{12}^n (N_{12}V_2)^4 - \frac{3}{2}V_2^n (N_{12}V_2)^3\Big],$$

$$\frac{1}{2}\overline{h_{nk}\partial_k h_{00}}^{1}_{(2)\ (4)} = 4k^3 m_2^3 r_{12}^{-4} N_{12}^n + 2k^3 m_2^2 m_1 r_{12}^{-4} N_{12}^n$$
$$+k^2 m_2^2 r_{12}^{-3} [3N_{12}^n (N_{12}V_2)^2 - 2V_2^n (N_{12}V_2) - 4N_{12}^n V_2^2],$$

$$\frac{1}{2}\overline{h_{nk}\partial_k h_{00}}^{1}_{(4)\ (2)} = -2k^3 m_2^3 r_{12}^{-4} N_{12}^n + 7k^3 m_2^2 m_1 r_{12}^{-4} N_{12}^n$$
$$+k^2 m_2^2 r_{12}^{-3} [N_{12}^n (N_{12}V_2)^2 - 4V_2^n (N_{12}V_2)],$$

$$\frac{1}{2}\overline{h_{nm}h_{mb}\partial_b h_{00}}^{1}_{(2)\ (2)\ (2)} = 4k^3 m_2^3 r_{12}^{-4} N_{12}^n,$$

$$\overline{h_{nk}\dot{h}_{0k}}^{1}_{(2)\ (3)} = 8k^3 m_2^2 m_1 r_{12}^{-4} N_{12}^n - 8k^2 m_2^2 r_{12}^{-3} V_2^n (N_{12}V_2),$$

$$\frac{1}{2}\overline{h_{0n}\dot{h}_{00}}^{1}_{(2)\ (2)} = -4k^2 m_2^2 r_{12}^{-3} V_2^n (N_{12}V_2),$$

$$-\overline{\dot{h}_{0n}}^{1}_{(5)} = -14k^3 m_2^2 m_1 r_{12}^{-4} N_{12}^n$$
$$+k^2 m_2^2 r_{12}^{-3} [N_{12}^n V_2^2 - 4N_{12}^n (N_{12}V_2)^2 + 3V_2^n (N_{12}V_2)]$$
$$+k^2 m_2 m_1 r_{12}^{-3} [[-6N_{12}^n V_2^2 + 10N_{12}^n (N_{12}V_2)^2]$$
$$-10V_2^n (N_{12}V_2) - 17N_{12}^n (V_2 V_1) - 42N_{12}^n (N_{12}V_1)^2$$
$$+12N_{12}^n V_1^2 + 46N_{12}^n (N_{12}V_2)(N_{12}V_1)$$
$$+12V_1^n (N_{12}V_2) - 21V_2^n (N_{12}V_1) + 6V_1^n (N_{12}V_1)]$$
$$+km_2 r_{12}^{-2} [-8V_2^n (N_{12}V_2)V_1^2 + 6V_2^n (N_{12}V_2)^3],$$

$$-\frac{1}{2}\overline{\dot{h}_{00}}^{1}_{(4)} V_1^n = k^2 m_1 m_2 r_{12}^{-3} V_1^n (N_{12}V_1)$$
$$-2k^2 m_2^2 r_{12}^{-3} V_1^n (N_{12}V_2)$$
$$+km_2 r_{12}^{-2} \Big[3V_1^n V_2^2 (N_{12}V_2)^2 + \frac{3}{2}V_1^n (N_{12}V_2)\Big],$$

$$\frac{1}{2}\overline{h_{00}\dot{h}_{00}}^{1}_{(2)\ (2)} V_1^n = 2k^2 m_2^2 r_{12}^{-3} V_1^n (N_{12}V_2),$$

$$\frac{1}{2}\overline{h_{0k}\partial_k h_{00}}^{1}_{(3)\ (2)} V_1^n = 4k^2 m_2^2 r_{12}^{-3} V_1^n (N_{12}V_2),$$

$$\overline{h_{0n}\partial_k h_{00}}^{A}_{(3)\ (2)} V_1^k = 8k^2 m_2^2 r_{12}^{-3} V_2^n (N_{12}V_1),$$

$$-\overline{h_{nb}\partial_k h_{0b}}^{1}_{(2)\ (3)} V_1^k = -8k^2 m_2^2 r_{12}^{-3} V_2^n (N_{12}V_1),$$

$$\overline{h_{nb}\partial_b h_{0k}}^{1}_{(2)\ (3)} V_1^k = 8k^2 m_2^2 r_{12}^{-3} N_{12}^n (V_1 V_2),$$

$$-\overline{h_{nk}\dot{h}_{kb}}^{1}_{(2)\ (2)} V_1^b = -4k^2 m_2^2 r_{12}^{-3} V_1^n (N_{12}V_2),$$

$$-\overline{\dot{h}_{nk}}^{1}_{(4)} V_1^k = k^2 m_2^2 r_{12}^{-3} [-V_2^n (N_{12}V_1) - N_{12}^n (V_1 V_2)$$
$$+4N_{12}^n (N_{12}V_2)(N_{12}V_1) + 2V_1^n (N_{12}V_2)]$$

$$
\begin{aligned}
&+k^2m_1m_2r_{12}^{-3}[-5N_{12}^nV_1^2+12N_{12}^n(V_1V_2)\\
&+24N_{12}^n(N_{12}V_1)^2-32N_{12}^n(N_{12}V_2)(N_{12}V_1)\\
&-7V_1^n(N_{12}V_1)+12V_2^n(N_{12}V_1)]\\
&+km_2r_{12}^{-2}[-3V_1^n(N_{12}V_2)^3\\
&+2V_1^n(N_{12}V_2)V_2^2+4V_2^n(N_{12}V_2)(V_1V_2)],\\
&\overline{\frac{1}{\partial_n h_{0k}}}_{(5)}V_1^k-\overline{\frac{1}{\partial_k h_{0n}}}_{(5)}V_1^k\\
&=k^2m_2^2r_{12}^{-3}[-V_2^n(N_{12}V_1)+N_{12}^n(V_1V_2)]\\
&+k^2m_1m_2r_{12}^{-3}[7N_{12}^nV_1^2-6N_{12}^n(V_2V_1)-7V_1^n(N_{12}V_1)\\
&+6V_2^N(N_{12}V_1)]+km_2r_{12}^{-2}[4V_2^n(N_{12}V_1)V_2^2\\
&-4N_{12}^nV_2^2(V_1V_2)+6N_{12}^n(N_{12}V_2)(V_1V_2)\\
&-6V_2^n(N_{12}V_1)(N_{12}V_2)-2N_{12}^nV_2^2(V_1V_2)],\\
&-\overline{\frac{1}{\partial_k h_{00}}}_{(4)}V_1^kV_1^n=km_2r_{12}^{-2}[-4V_1^n(N_{12}V_1)V_2^2\\
&-2V_1^n(N_{12}V_2)(V_1V_2)+3V_1^n(N_{12}V_1)(N_{12}V_2)^2]\\
&+4k^2m_2^2r_{12}^{-3}V_1^n(N_{12}V_1)+2k^2m_2m_1r_{12}^{-3}V_1^n(N_{12}V_1),\\
&\overline{\frac{1}{h_{00}\,\partial_k h_{00}}}_{(2)\quad(2)}V_1^kV_1^n=-4k^2m_2m_1r_{12}^{-3}V_1^n(N_{12}V_1),\\
&\frac{1}{2}\overline{\frac{1}{\partial_n h_{mk}}}_{(4)}V_1^kV_1^m\\
&=k^2m_2^2r_{12}^{-3}[-V_1^n(N_{12}V_1)+N_{12}^nV_1^2+2N_{12}^n(N_{12}V_1)^2]\\
&+k^2m_1m_2r_{12}^{-3}[-4N_{12}^n(N_{12}V_1)^2+3V_1^n(N_{12}V_1)]\\
&+km_2r_{12}^{-2}\Big[V_2^n(N_{12}V_2)V_1^2\\
&-\frac{3}{2}N_{12}^n(N_{12}V_2)^2V_1^2+2N_{12}^n(V_1V_2)^2\Big],\\
&-\overline{\frac{1}{\partial_b h_{nk}}}_{(4)}V_1^kV_1^b\\
&=k^2m_2^2r_{12}^{-3}[-V_1^n(N_{12}V_1)+N_{12}^nV_1^2-4N_{12}^n(N_{12}V_1)^2]\\
&+k^2m_1m_2r_{12}^{-3}[8N_{12}^n(N_{12}V_1)^2-3V_1^n(N_{12}V_1)-3N_{12}^nV_1^2]\\
&+km_2r_{12}^{-2}[3V_1^n(N_{12}V_2)^2(N_{12}V_1)\\
&-4V_2^n(N_{12}V_1)(V_1V_2)-2V_1^n(V_1V_2)(N_{12}V_2)],\\
&\frac{1}{2}\overline{\frac{1}{h_{ns}\,\partial_s h_{kb}}}_{(2)\quad(2)}V_1^kV_1^b=-2k^2m_2^2r_{12}^{-3}N_{12}^nV_1^2,\\
&-\overline{\frac{1}{h_{ns}\,\partial_m h_{ks}}}_{(2)\quad(2)}V_1^kV_1^m=4k^2m_2^2r_{12}^{-3}V_1^n(N_{12}V_1),\\
&-\overline{\frac{1}{\partial_s h_{0b}}}_{(3)}V_1^bV_1^sV_1^n=4km_2r_{12}^{-2}(V_1V_2)V_1^n(N_{12}V_1),\\
&-\frac{1}{2}\overline{\frac{1}{\dot h_{mk}}}_{(2)}V_1^mV_1^nV_1^k=-km_2r_{12}^{-2}V_1^2V_1^n(N_{12}V_2).
\end{aligned}
\tag{5.5}
$$

We have introduced the following notation:

$$(N_{12}V_2)\equiv N_{12}^kV_2^k,\quad (V_1V_2)\equiv V_1^kV_2^k.$$

Expressions (3.18), (5.1) and (5.4), (5.5) lead to the post-post-Newtonian equations of motion for two point particles:

$$
\begin{aligned}
\frac{d^2\xi_1^n}{dt^2}&=-km_2r_{12}^{-2}N_{12}^n\\
&+c^{-2}\Bigg\{km_2r_{12}^{-2}\Big[\frac{3}{2}N_{12}^n(N_{12}V_2)^2-2N_{12}^nV_2^2\\
&+4N_{12}^n(V_1V_2)-N_{12}^nV_1^2+3V_2^n(N_{12}V_2)-4V_2^n(N_{12}V_1)\\
&+4V_1^n(N_{12}V_1)-3V_1^n(N_{12}V_2)\Big]\\
&+k^2m_2(5m_1+4m_2)r_{12}^{-3}N_{12}^n\Bigg\}\\
&+c^{-4}\Bigg\{km_2r_{12}^{-2}\Big[4N_{12}^nV_2^2(V_1^kV_2^k)-2N_{12}^nV_2^4\\
&-2N_{12}^n(V_1^kV_2^k)^2+\frac{3}{2}N_{12}^nV_1^2(N_{12}V_2)^2+\frac{9}{2}N_{12}^nV_2^2(N_{12}V_2)^2\\
&-6N_{12}^n(V_1^kV_2^k)(N_{12}V_2)^2-\frac{15}{8}N_{12}^n(N_{12}V_2)^4\\
&+5V_1^nV_2^2(N_{12}V_2)-3V_2^nV_1^2(N_{12}V_2)+V_1^nV_1^2(N_{12}V_2)\\
&+4V_1^nV_2^2(N_{12}V_1)-4V_2^nV_2^2(N_{12}V_1)+5V_2^nV_2^2(N_{12}V_2)\\
&-4V_1^n(V_1^kV_2^k)(N_{12}V_1)+4V_2^n(V_1^kV_2^k)(N_{12}V_1)\\
&+4V_1^n(V_1^kV_2^k)(N_{12}V_2)-4V_1^n(V_1^kV_2^k)(N_{12}V_2)\\
&+6V_2^n(N_{12}V_1)(N_{12}V_2)^2-6V_1^n(N_{12}V_1)(N_{12}V_2)^2
\end{aligned}
\tag{5.6}
$$

$$+\frac{9}{2}V_1^n(N_{12}V_2)^3-\frac{9}{2}V_2^n(N_{12}V_2)^3\Big]$$

$$+k^2m_1m_2r_{12}^{-3}\Big[-\frac{15}{4}N_{12}^nV_1^2+\frac{5}{4}N_{12}^nV_2^2-\frac{5}{2}N_{12}^nV_2^kV_1^k$$

$$+\frac{39}{2}N_{12}^n(N_{12}V_1)^2+\frac{17}{2}N_{12}^n(N_{12}V_2)^2$$

$$-39N_{12}^n(N_{12}V_2)(N_{12}V_1)-\frac{55}{4}V_2^n(N_{12}V_2)$$

$$+\frac{55}{4}V_1^n(N_{12}V_2)+\frac{63}{4}V_2^n(N_{12}V_1)-\frac{63}{4}V_1^n(N_{12}V_1)\Big]$$

$$+k^2m_2^2r_{12}^{-3}[4N_{12}^nV_2^2-8N_{12}^n(V_1V_2)-6N_{12}^n(N_{12}V_2)^2$$

$$+2N_{12}^n(N_{12}V_1)^2-4N_{12}^n(N_{12}V_2)(N_{12}V_1)$$

$$+2V_2^n(N_{12}V_2)-2V_1^n(N_{12}V_2)+2V_2^n(N_{12}V_1)$$

$$-2V_1^n(N_{12}V_1)]-9k^3m_2^3r_{12}^{-4}N_{12}^n$$

$$-\frac{57}{4}k^3m_1^2m_2r_{12}^{-4}N_{12}^n-\frac{69}{2}k^3m_2^2m_1r_{12}^{-4}N_{12}^n\Big\}.$$

The equations of motion for the second particle can be obtained by substituting subscript 2 for 1.

6. CONCLUSIONS

Equations of motion (5.6), derived here by the EIH method, coincide with the equations obtained by Kopeikin [10], who used the Fock method. Thus, we have proved that both methods (the EIH method and the Fock method), which can be regarded as classical, give not only identical post-Newtonian, but also post-post-Newtonian equations of motion for point particles in the harmonic coordinates.

It should be noted that the definition of point particles in the Fock method is very clear from the standpoint of physics. In the EIH method, clarity is lost, but mathematical calculations are simplified. We give here almost all calculations (except very simple ones).

If we take retarded waves instead of standing waves in the expansions of metric tensor (2.9), the post-post-Newtonian equations of motion for point particles will not change; a difference will appear only in the next approximations.

APPENDIX A

Derivatives of the Products of Singular Functions

We apply identities (2.10) and (2.11) to singular functions r_A^{-L}, r_B^{-M}, $L = 1, 2, 3, \ldots, M = 1, 2, 3, \ldots$.

Substituting $f = r_A^{-1}$ and, successively, $g = r_A^{-1}$, r_A^{-2}, … into identity (2.10), we obtain

$$\partial_a r_A^{-L} \equiv L r_A^{-L+1}\partial_a r_A^{-1}. \tag{A.1}$$

Substituting r_A^{-L} and r_B^{-M} into identity (2.10) and using (A.1), we obtain

$$\begin{aligned}\partial_a(r_A^{-L}r_B^{-M}) &\equiv L r_B^{-M}r_A^{-L+1}\partial_a r_A^{-1}\\ &+ M r_A^{-L}r_B^{-M+1}\partial_a r_B^{-1}.\end{aligned} \tag{A.2}$$

A similar substitution of functions $f = r_A^{-L}$ and $g = r_A^{-1}$, r_A^{-2}, … into identity (A.1) gives

$$\begin{aligned}\partial_a\partial_b r_A^{-L} &\equiv L(L-1)r_A^{-L+2}\partial_a r_A^{-1}\partial_b r_A^{-1}\\ &+ L r_A^{-L+1}\partial_a\partial_b r_A^{-1}.\end{aligned} \tag{A.3}$$

Using identities (2.11) and (A.1)–(A.3), we obtain the following expressions for r_A^{-L} and r_B^{-M}:

$$\begin{aligned}\partial_a\partial_b(r_A^{-L}r_B^{-M}) &\equiv L r_B^{-M}r_A^{-L+1}\partial_a\partial_b r_A^{-1}\\ &+ M r_A^{-L}r_B^{-M+1}\partial_a\partial_b r_B^{-1} + L(L-1)r_B^{-M}r_A^{-L+2}\partial_a r_A^{-1}\partial_b r_A^{-1}\\ &+ M(M-1)r_A^{-L}r_B^{-M+2}\partial_a r_B^{-1}\partial_b r_B^{-1}\\ &+ LM r_A^{-L+1}r_B^{-M+1}[\partial_a r_A^{-1}\partial_b r_B^{-1} + \partial_a r_B^{-1}\partial_b r_A^{-1}].\end{aligned} \tag{A.4}$$

In particular, formulas (A.3) and (A.4) give

$$\Delta r_A^{-L} \equiv L(L-1)r_A^{-L+2}\partial_a r_A^{-1}\partial_a r_A^{-1} + L r_A^{-L+1}\Delta r_A^{-1}, \tag{A.5}$$

$$\begin{aligned}\Delta(r_A^{-L}r_B^{-M}) &\equiv L r_B^{-M}r_A^{-L+1}\Delta r_A^{-1} + M r_A^{-L}r_B^{-M+1}\Delta r_B^{-1}\\ &+ L(L-1)r_B^{-M}r_A^{-L+2}\partial_a r_A^{-1}\partial_a r_A^{-1}\\ &+ M(M-1)r_A^{-L}r_B^{-M+2}\partial_a r_B^{-1}\partial_a r_B^{-1}\\ &+ 2LM r_A^{-L+1}r_B^{-M+1}\partial_a r_A^{-1}\partial_a r_B^{-1}.\end{aligned} \tag{A.6}$$

Substituting the formulas

$$\partial_a r_A^{-1} = -r_A^{-2}N_A^a, \tag{A.7}$$

$$\partial_a\partial_b r_A^{-1} = r_A^{-3}[3N_A^aN_A^b - \delta_{ab}] - \frac{4}{3}\pi\delta(\bar{r}_A)\delta_{ab} \tag{A.8}$$

(which are borrowed, as in [10, 16], from the theory of generalized functions [20]) into identities (A.1)–(A.6), we obtain the following formulas defining the deriva-

tives of the products of functions r_A^{-L} and r_B^{-M} in terms of the products of the derivatives of these functions:

$$\partial_a r_A^{-L} = -Lr_A^{-L-1}N_A^a, \tag{A.9}$$

$$\begin{aligned}\partial_a(r_A^{-L}r_B^{-M}) &= -Lr_B^{-M}r_A^{-L-1}N_A^a \\ &\quad - Mr_A^{-L}r_B^{-M-1}N_B^a,\end{aligned} \tag{A.10}$$

$$\begin{aligned}\partial_a\partial_b r_A^{-L} &= Lr_A^{-L-2}[(L+2)N_A^aN_A^b - \delta_{ab}] \\ &\quad - \frac{4}{3}\pi Lr_A^{-L+1}\delta(\bar{r}_A)\delta_{ab},\end{aligned} \tag{A.11}$$

$$\begin{aligned}\partial_a\partial_b(r_A^{-L}r_B^{-M}) &= LMr_A^{-L-1}r_B^{-M-1}(N_A^aN_B^b + N_A^bN_B^a) \\ &\quad + Mr_A^{-L}r_B^{-M-2}[(M+2)N_B^aN_B^b - \delta_{ab}] \\ &\quad + Lr_B^{-M}r_A^{-L-2}[(L+2)N_A^aN_A^b - \delta_{ab}] \\ &\quad - \frac{4}{3}\pi Mr_A^{-L}r_B^{-M+1}\delta(\bar{r}_B)\delta_{ab} - \frac{4}{3}\pi Lr_B^{-M}r_A^{-L+1}\delta(\bar{r}_A)\delta_{ab},\end{aligned} \tag{A.12}$$

$$\Delta r_A^{-L} = L(L-1)r_A^{-L-2} - 4\pi Lr_A^{-L+1}\delta(\bar{r}_A), \tag{A.13}$$

$$\begin{aligned}\Delta(r_A^{-L}r_B^{-M}) &= L(L-1)r_B^{-M}r_A^{-L-2} + M(M-1)r_A^{-L}r_B^{-M-2} \\ &\quad - 4\pi Lr_B^{-M}r_A^{-L+1}\delta(\bar{r}_A) - 4\pi Mr_A^{-L}r_B^{-M+1}\delta(\bar{r}_B).\end{aligned} \tag{A.14}$$

APPENDIX B

Regularization of Diverging Integrals

To derive equations of motion (5.6), we must evaluate the following diverging integrals in formulas (5.2) and (5.3):

$$\begin{gathered}\int\hat{\delta}(\bar{r}_A)r_A^Kr_B^Md(\bar{x}), \quad \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_A^nd(\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_B^nd(\bar{x}), \quad \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_A^nN_A^kd(\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_A^nN_B^kd(\bar{x}), \quad \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_B^nN_B^kd(\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_A^nN_A^kN_A^sd(\bar{x}), \quad \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_A^nN_A^kN_B^sd(\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_A^nN_B^kN_B^sd(\bar{x}), \quad \int\hat{\delta}(\bar{r}_A)r_A^Kr_B^MN_B^nN_B^kN_B^sd(\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)\partial_n\partial_b^A\ln S_{AB}(d\bar{x}), \quad \int\hat{\delta}(\bar{r}_A)\partial_n\partial_b^B\ln S_{AB}(d\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)\partial_n\partial_k^A\partial_b^B\ln S_{AB}(d\bar{x}), \quad \int\hat{\delta}(\bar{r}_A)\partial_n^B\partial_k^A\partial_b^A\ln S_{AB}(d\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)\partial_n^B\partial_k^B\partial_b^A\ln S_{AB}(d\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)r_A^{-2}N_A^a\partial_n^B\partial_a^A\ln S_{AB}(d\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)r_B^{-2}N_B^a\partial_n^B\partial_a^A\ln S_{AB}(d\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)r_A^{-2}N_A^a\partial_n^A\partial_a^B\ln S_{AB}(d\bar{x}), \\ \int\hat{\delta}(\bar{r}_A)r_B^{-2}N_A^a\partial_n^A\partial_a^B\ln S_{AB}(d\bar{x})\end{gathered} \tag{B.1}$$

for $K = -4, -3, -2, -1, 0, 1$ and $M = -4, -3, -2, -1, 0, 1$. Taking into account relation (3.16), we obtain the expansion of $r_B^KN_B^a$ and S_{AB}^M into power series in $r_A \approx 0$:

$$\begin{aligned}r_B^KN_B^a &= r_{AB}^KN_{AB}^a + \{r_{AB}^{(K-1)}N_A \\ &\quad + (K-1)r_{AB}^{(K-1)}N_{AB}^aN_{AB}^bN_A^b\}r_A \\ &\quad + \left\{(K-1)r_{AB}^{(K-2)}N_{AB}^bN_{AB}^aN_A^bN_A^b + \frac{1}{2}(K-1)(K-3)\right. \\ &\quad \left.\times r_{AB}^{(K-2)}N_{AB}^aN_{AB}^bN_{AB}^sN_{AB}^bN_A^sN_A^s + \frac{1}{2}(K-1)r_{AB}^{(K-2)}N_{AB}^a\right\}r_A^2 \\ &\quad + \left\{\frac{1}{6}(K-1)(K-3)(K-5)r_{AB}^{(K-3)}N_{AB}^aN_{AB}^bN_{AB}^sN_{AB}^j\right. \\ &\quad \times N_A^bN_A^sN_A^j + \frac{1}{2}(K-1)(K-3)r_{AB}^{(K-3)}N_{AB}^aN_{AB}^bN_A^b \\ &\quad + \frac{1}{2}(K-1)r_{AB}^{(K-3)}N_{AB}^a + \frac{1}{2}(K-1)(K-3) \\ &\quad \left.\times r_{AB}^{(K-3)}N_{AB}^bN_{AB}^sN_A^bN_A^sN_A^a\right\}r_A^3 + O(r_A^4),\end{aligned} \tag{B.2}$$

$$\begin{aligned}S_{AB}^M &= (2r_{AB})^M + M(2r_{AB})^{M-1}(1 + N_{AB}^KN_A^k)r_A \\ &\quad + \frac{1}{2}\{M(M-1)(2r_{AB})^{M-2}(1 + 2N_{AB}^kN_A^k + N_{AB}^kN_{AB}^sN_A^kN_A^s) \\ &\quad + M(2r_{AB})^{M-1}(r_{AB}^{-1} - r_{AB}^{-1}N_{AB}^kN_{AB}^sN_A^kN_A^s)\}r_A^2 \\ &\quad + \left\{\frac{1}{6}M(M-1)(M-2)(2r_{AB})^{M-3}(1\right. \\ &\quad + N_{AB}^kN_{AB}^sN_{AB}^bN_A^kN_A^sN_A^b + 3N_{AB}^kN_A^k \\ &\quad + 3N_{AB}^kN_{AB}^kN_A^sN_A^s) + \frac{1}{2}M(M-1)(2r_{AB})^{M-2} \\ &\quad \times (r_{AB}^{-1} + r_{AB}^{-1}N_{AB}^kN_A^k - r_{AB}^{-2}N_{AB}^kN_{AB}^sN_A^kN_A^s \\ &\quad - r_{AB}^{-1}N_{AB}^kN_{AB}^sN_{AB}^bN_A^kN_A^sN_A^b) + \frac{1}{2}(2r_{AB})^{M-1} \\ &\quad \left.\times (r_{AB}^{-2}N_{AB}^kN_{AB}^sN_{AB}^bN_A^kN_A^sN_A^b - r_{AB}^{-2}N_{AB}^kN_{AB}^k)\right\}r_A^3 + O(r_A^4).\end{aligned} \tag{B.3}$$

Using the formulas [1]

$$\int\hat{\delta}(\bar{r}'_A)(N^k_A N'^k_A)^{2p}(d\bar{x}') = \frac{1}{2p+1}, \quad p = 1, 2, \ldots,$$

$$\int\hat{\delta}(\bar{r}'_A)(N^k_A N'^k_A)^{2p-1}(d\bar{x}') = 0, \quad p = 1, 2, \ldots,$$

and property (2.6) of the $\hat{\delta}$ function, we obtain the following integral making nonzero contributions to equations of motion (5.6):

$$\int\hat{\delta}(\bar{r}_A)r_A^{-1}r_B^{-2}N^n_A N^b_A N^s_A(d\bar{x})$$
$$= -\frac{2}{15}r^{-3}_{AB}[N^n_{AB}\delta_{bs} + N^b_{AB}\delta_{ns} + N^s_{AB}\delta_{nb}],$$

$$\int\hat{\delta}(\bar{r}_A)r_A^{-2}N^n_B(d\bar{x}) = -\frac{1}{3}r^{-2}_{AB}N^n_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)r_A^{-3}r_B N^n_A(d\bar{x}) = -\frac{1}{15}r^{-2}_{AB}N^n_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)r_A^{-3}r_B^{-3}N^n_A(d\bar{x}) = -r^{-6}_{AB}N^n_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)r_A^{-2}r_B^{-4}N^n_B(d\bar{x}) = \frac{5}{3}r^{-6}_{AB}N^n_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)\partial_n\partial^A_b\ln S_{AB}(d\bar{x}) = \left[\frac{1}{4}\delta_{nb} - \frac{1}{4}N^n_{AB}N^b_{AB}\right]r^{-2}_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)\partial_n\partial^B_b\ln S_{AB}(d\bar{x}) = \left[N^n_{AB}N^b_{AB} - \frac{1}{2}\delta_{nb}\right]r^{-2}_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)\partial_n\partial^A_a\partial^B_b\ln S_{AB}(d\bar{x}) \tag{B.4}$$
$$= \left[\frac{1}{2}N^b_{AB}\delta_{an} + \frac{1}{4}N^n_{AB}\delta_{ab} + \frac{1}{4}N^a_{AB}\delta_{nb} - N^n_{AB}N^a_{AB}N^b_{AB}\right]r^{-3}_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)\partial^B_n\partial^A_a\partial^A_b\ln S_{AB}(d\bar{x})$$
$$= \left[\frac{3}{4}N^b_{AB}\delta_{an} + \frac{3}{4}N^a_{AB}\delta_{bn} + \frac{1}{2}N^n_{AB}\delta_{ab} - 3N^n_{AB}N^a_{AB}N^b_{AB}\right]r^{-3}_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)\partial^B_n\partial^B_b\partial^A_a\ln S_{AB}(d\bar{x})$$
$$= [-N^b_{AB}\delta_{an} - N^n_{AB}\delta_{ab} - N^a_{AB}\delta_{nb} + 4N^n_{AB}N^a_{AB}N^b_{AB}]r^{-3}_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)r_B^{-2}N^a_B\partial^B_n\partial^A_a\ln S_{AB}(d\bar{x}) = \frac{1}{2}r^{-4}_{AB}N^n_{AB},$$

$$\int\hat{\delta}(\bar{r}_A)r_B^{-2}N^a_A\partial^A_n\partial^B_a\ln S_{AB}(d\bar{x}) = \frac{1}{2}r^{-4}_{AB}N^n_{AB}.$$

REFERENCES

1. V. A. Fok, *The Theory of Space, Time and Gravitation* (Fizmatgiz, Moscow, 1961; Pergamon Press, Oxford, 1964).
2. L. Infeld and J. Plebanski, *Motion and Relativity* (Pergamon, New York, 1960).
3. V. A. Fok, Zh. Éksp. Teor. Fiz. **9**, 375 (1939).
4. N. M. Petrova, Zh. Éksp. Teor. Fiz. **19**, 989 (1949).
5. L. Infeld and J. Plebanski, Bull. Acad. Pol. Sci. Cl. 3 **4**, 689 (1956).
6. L. Infeld and J. Plebanski, Bull. Acad. Pol. Sci. Cl. 3 **5**, 51 (1957).
7. A. Einstein, L. Infeld, and B. Hoffmann, Ann. Math. **39**, 65 (1938).
8. A. Einstein and L. Infeld, Ann. Math. **41**, 455 (1940).
9. A. Einstein and L. Infeld, Can. J. Math. **1**, 209 (1949).
10. S. M. Kopeĭkin, Astron. Zh. **62**, 889 (1985) [Sov. Astron. **29**, 516 (1985)].
11. T. Damour, *Gravitational Radiation*, Ed. by N. Deruelle and T. Piran (North-Holland, Amsterdam, 1983).
12. L. Blanchet, G. Faye, and B. Ponsot, Phys. Rev. D **58**, 124002 (1998).
13. R. Arnowitt, S. Deser, and C. M. Misner, Phys. Rev. **120**, 313 (1960).
14. R. Arnowitt, S. Deser, and C. M. Misner, in *Gravitation: An Introduction to Current Research*, Ed. by L. Witten (Wiley, New York, 1962), p. 227.
15. T. Ohta, H. Okamura, T. Kimura, and K. Hiida, Prog. Theor. Phys. **51**, 1220 (1974).
16. G. Schafer, in *Mathematics of Gravitation*, Ed. by A. Królak (Banach Center, Polish Academy of Sciences, Warsaw, 1997), Vol. 41, Part 2, p. 43.
17. Cz. Jankewicz, Zh. Éksp. Teor. Fiz. **44**, 649 (1963) [Sov. Phys. JETP **17**, 440 (1963)].
18. G. Voevoda, Zh. Éksp. Teor. Fiz. **45**, 2051 (1963) [Sov. Phys. JETP **18**, 1409 (1964)].
19. L. Infeld, Rev. Mod. Phys. **29**, 398 (1957).
20. N. M. Gel'fand and G. E. Shilov, *Generalized Functions* (Fizmatgiz, Moscow, 1959; Academic, New York, 1964).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 853–869.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 982–998.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Radiative Corrections to Polarization Observables in Elastic Electron–Deuteron Scattering in Hadronic Variables¶

G. I. Gakh and N. P. Merenkov

National Science Center Kharkov Institute of Physics and Technology, Kharkov, 61108 Ukraine
e-mail: merenkov@kipt.kharkov.ua

Received November 17, 2003

Abstract—Model-independent QED radiative corrections to polarization observables in the elastic scattering of unpolarized and longitudinally polarized electron beam by a deuteron target are calculated. Two experimental setups are considered: the deuteron target is arbitrarily polarized, or the vector and/or tensor polarization of the recoil deuteron is measured. The calculations are based on taking all essential Feynman diagrams into account and using the covariant parametrization of the deuteron polarization state. The radiative corrections are calculated for the hadronic variables using invariant integration of the leptonic tensor. Numerical estimates of the radiative corrections to the polarization observables are made for various values of the kinematical variables. *© 2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Recent progress in electron-scattering experiments has made it possible to measure various polarization observables in the region of momentum transfers, where they can help to discriminate between different theoretical predictions. Much of this progress has been made possible by modern high-energy electron accelerators with a high duty cycle, such as MAMI or JLAB, and by the development of polarized sources, targets, and polarimeters.

Electron scattering by few-body systems has shown that the two-body terms of the nuclear electromagnetic operators make important contributions to the observables.

The deuteron, the only bound two-nucleon system, is one of the fundamental systems of nuclear physics. Accordingly, many studies, both experimental and theoretical, have been devoted to it. Of particular interest today is the degree to which the deuteron can be understood as a system of two nucleons interacting via the known nucleon–nucleon interaction.

When addressing the electromagnetic properties of the deuteron more specifically, the corresponding question concerns the ability to predict the three deuteron form factors starting from the calculated deuteron wavefunction and the nucleon form factors known from the electron–nucleon scattering. At low momentum transfers, predictions and data agree quite well when only one-body terms are taken into account; at higher momentum transfers, two-body contributions are important. Whether it is necessary to make allowance for quark degrees of freedom is still a matter of debate. An up-to-date status of the experimental and theoretical research into the deuteron can be found in reviews [1].

The deuteron electromagnetic form factors most often are studied in order to check our understanding of the two-nucleon system. In parallel, however, the deuteron form factors are also exploited to get a better handle on the neutron form factors. In the past, much of our knowledge on the neutron charge form factor $G_{En}(q^2)$ came from precision studies of the deuteron structure function $A(q^2)$ (see Eq. (16) for the definition). Only very recently have experiments involving both polarized electrons and polarized target/recoil nuclei made it possible to access G_{En} via other observables. At large q^2, however, G_{En} is still largely unknown, which represents a serious handicap to quantitative understanding of the deuteron charge form factors.

Elastic electron–deuteron scattering has been investigated in many experiments, and the cross-sectional data today cover a large range of momentum transfers (see review [2]). Some of these data are obviously not very precise; other data, mainly more recent, have reached accuracies that have achieved a level 1%. Over the last few years, it has increasingly become possible to measure not only cross sections, but also spin observables. The knowledge of these spin observables is imperative if one wants to separate the contributions of the different form factors to the $A(q^2)$ structure function. In terms of experiment, good progress has been made. In particular, we now have a reasonably complete set of polarization data for electron–deuteron scattering that allows us to separate the deuteron charge and quadrupole form factors.

¶This article was submitted by the authors in English.

Two techniques are basically available to measure such spin observables.

(1) With storage rings, one can use polarized, internal deuteron gas targets from an atomic beam source [3]. The high intensity of the circulating electron beam allows one to achieve acceptable luminosities despite the very low thickness of the gas target.

(2) At facilities with external beams, one can use polarimeters to measure the polarization of the recoil deuterons. High beam intensities are a prerequisite because the measurement of polarization, which requires a second reaction of the deuteron, involves the loss of a few orders of magnitude in the counting rate.

Current experiments at modern accelerators have reached a new level of precision; this requires a new approach to data analysis and inclusion of all possible systematic uncertainties. An important source of such uncertainties is the electromagnetic radiative effects caused by physical processes that occur in higher orders of perturbation theory with respect to electromagnetic interaction. Previously, we calculated the radiative corrections to the polarization observables in deep inelastic scattering (due to a tensor-polarized deuteron target) [4] and in semi-inclusive deep inelastic scattering (due to the vector polarization of the target and/or outgoing hadron) [5].

In the present paper, we calculate the model-independent $O(\alpha)$ QED corrections to the polarization observables in the scattering of the unpolarized or longitudinally polarized electron beam off a vector- or tensor-polarized deuteron target (or production of an arbitrarily polarized final deuteron),

$$e^-(k_1) + D(p_1) \longrightarrow e^-(k_2) + D(p_2). \qquad (1)$$

The experimental setups also make it possible to measure the tensor polarization observables under scattering off the polarized deuteron target as well as by determination of the recoil deuteron polarizations. Different aspects of respective approaches [6] in JLAB have been discussed recently in [7].

For polarized-target experiments, a scattered electron is usually detected, although the measurement of the recoil deuteron is also possible. In the first case, leptonic variables, and in the second case, hadronic ones are used to calculate radiative corrections. In the leptonic variables, the virtuality of the heavy intermediate photon is not fixed due to the possibility to radiate a photon by the initial or scattered electron. As a result, the corresponding radiative correction involves certain integrals with deuteron form factors over the intermediate photon mass that cannot be computed in a model-independent way (without knowing the form factors). On the contrary, in hadron variables, the heavy photon mass is fixed and the respective radiative correction caused by electromagnetic effects in the lepton part of the interaction can be calculated, in principle, in a model-independent way in any order of perturbation theory.

The measurement of the recoil-deuteron polarization requires the analysis of the second scattering, which, in turn, suggests knowledge of the recoil-deuteron 3-momentum. Therefore, calculation of the radiative correction in this experimental setup requires using the hadronic variables, which we consider in this work. Our approach is based on the covariant parametrization of the polarization state of the deuteron target or recoil deuteron in terms of the 4-momenta of the particles in process (1), used first in [8–10] and recently in [4, 5]. In addition, we use invariant integration of the leptonic tensor to calculate the contribution to the radiative correction caused by the hard-photon radiation. Derived this way, the first-order QED correction is generalized by exponentiation of the most singular terms in the limiting case where the real photon energy is small. Our analytical final results are simple enough and have a physically transparent form.

2. BORN APPROXIMATION

Different polarization observables in the electron–deuteron elastic scattering have been studied in [11–16] and other papers, where the results were expressed in terms of the deuteron electromagnetic form factors. Here, we reproduce most of these results using the method of covariant parametrization of the deuteron polarization state in terms of the particle 4-momenta and demonstrate the advantage of this approach.

We first consider the scattering off the polarized deuteron target. In the one-photon exchange approximation, we define the cross section of process (1) in terms of the contraction of the leptonic $L_{\mu\nu}$ and hadronic $H_{\mu\nu}$ tensors (we neglect the electron mass wherever possible),

$$d\sigma = \frac{\alpha^2}{2Vq^4} L^B_{\mu\nu} H_{\mu\nu} \frac{d^3k_2}{\varepsilon_2} \frac{d^3p_2}{E_2} \delta(k_1 + p_1 - k_2 - p_2), \quad (2)$$

where $V = 2k_1p_1$, ε_2 and E_2 are the respective energies of the scattered electron and the recoil deuteron, and $q = k_1 - k_2 = p_2 - p_1$ is the 4-momentum of the heavy virtual photon that probes the deuteron. For a longitudinally polarized electron beam, the leptonic tensor in the Born approximation is given by

$$L^B_{\mu\nu} = q^2 g_{\mu\nu} + 2(k_{1\mu}k_{2\nu} + k_{2\mu}k_{1\nu}) + 2i\lambda(\mu\nu qk_1),$$
$$(\mu\nu\alpha b) = \varepsilon_{\mu\nu\lambda\rho} a_\lambda b_\rho, \qquad (3)$$

where λ is the degree of the beam polarization (in what follows, we assume that the electron beam is completely polarized, and, consequently, $\lambda = 1$).

The hadronic tensor can be expressed via the deuteron electromagnetic current J_μ describing the transition $\gamma^* d \longrightarrow d$ as

$$H_{\mu\nu} = J_\mu J_\nu^*. \tag{4}$$

Because the deuteron is a spin-one nucleus, its electromagnetic current is completely described by three form factors. Assuming the P- and C-invariance of the hadron electromagnetic interaction, we can write this current as [17]

$$J_\mu = (p_1 + p_2)_\mu \times\left[-G_1 U_1 U_2^* + \frac{G_3}{M^2}\left(U_1 q U_2^* q - \frac{q^2}{2} U_1 U_2^*\right)\right] + G_2(U_{1\mu} U_2^* q - U_{2\mu}^* U_1 q), \tag{5}$$

where $U_{1\mu}(U_{2\mu})$ is the wavefunction of the initial (recoil) deuteron, M is the deuteron mass, and G_i ($i = 1, 2, 3$) are the deuteron electromagnetic form factors. Due to the current hermiticity, the form factors $G_i(q^2)$ are real functions in the region of spacelike momentum transfer. They can be related to the standard deuteron form factors, G_C (the charge monopole), G_M (the magnetic dipole), and G_Q (the quadrupole), as

$$\begin{aligned} G_M &= -G_2, \quad G_Q = G_1 + G_2 + 2G_3, \\ G_C &= \frac{\rho}{6\tau}(G_2 - G_3) + \left(1 + \frac{\rho}{6\tau}\right)G_1, \\ \rho &= -\frac{q^2}{V}, \quad \tau = \frac{M^2}{V}. \end{aligned} \tag{6}$$

The standard form factors have the normalizations

$$\begin{aligned} G_C(0) &= 1, \quad G_M(0) = \frac{M}{m_n}\mu_d, \\ G_Q(0) &= M^2 Q_d, \end{aligned} \tag{7}$$

where m_n is the nucleon mass, $\mu_d(Q_d)$ is deuteron magnetic (quadrupole) moment, and their values are

$$\mu_d = 0.857, \quad Q_d = 0.2859\, fm^2.$$

In calculating the expression for the hadron tensor $H_{\mu\nu}$ in terms of the deuteron electromagnetic form factors, using the explicit form of electromagnetic current (5), one has to use the spin-density matrix of the initial and final deuterons

$$\begin{aligned} U_{1\alpha} U_{1\beta}^* &= -\frac{1}{3}\left(g_{\alpha\beta} - \frac{p_{1\alpha} p_{1\beta}}{M^2}\right) - \frac{i}{2M}(\alpha\beta W p_1) + Q_{\alpha\beta}, \\ U_{2\alpha} U_{2\beta}^* &= -\left(g_{\alpha\beta} - \frac{p_{2\alpha} p_{2\beta}}{M^2}\right) \end{aligned} \tag{8}$$

if the deuteron target is polarized and the polarization of the recoil deuteron is not measured. Here, W_α and $Q_{\alpha\beta}$ are the target-deuteron polarization 4-vector and the quadrupole tensor, respectively.

Taking Eqs. (4), (5), and (8) into account, we can write the hadronic tensor in the general case as

$$H_{\mu\nu} = H_{\mu\nu}(0) + H_{\mu\nu}(V) + H_{\mu\nu}(T), \tag{9}$$

where $H_{\mu\nu}(0)$ corresponds to the unpolarized case and $H_{\mu\nu}(V)$ ($H_{\mu\nu}(T)$) corresponds to the case of the vector (tensor) polarization of the deuteron target. The $H_{\mu\nu}(0)$ term has the form

$$\begin{aligned} H_{\mu\nu}(0) &= -W_1 \tilde{g}_{\mu\nu} + \frac{W_2}{M^2}\tilde{p}_{1\mu}\tilde{p}_{1\nu}, \\ \tilde{g}_{\mu\nu} &= g_{\mu\nu} - \frac{q_\mu q_\nu}{q^2}, \quad \tilde{p}_{1\mu} = p_{1\mu} - \frac{p_1 q}{q^2} q_\mu, \\ W_1 &= -\frac{2q^2}{3}\left(1 + \frac{\rho}{4\tau}\right)G_\mu^2, \\ W_2 &= 4M^2\left(\frac{\rho}{6\tau}G_M^2 + G_C^2 + \frac{\rho^2}{18\tau^2}G_Q^2\right). \end{aligned} \tag{10}$$

In the case under consideration, the term $H_{\mu\nu}(V)$, responsible for the vector polarization of the deuteron target, can be written as

$$H_{\mu\nu}(V) = \frac{iG_M}{2M}\Big[(G_M - G)(Wp_2)(\mu\nu q p_1) + 2M^2\left(1 + \frac{\rho}{4\tau}\right)G(\mu\nu q W)\Big], \tag{11}$$

$$G = 2G_C + \frac{\rho}{6\tau}G_Q,$$

where the 4-vector of the target deuteron polarization satisfies the conditions

$$W^2 = -1, \quad Wp_1 = 0.$$

For the tensor-polarized deuteron target, $H_{\mu\nu}(T)$ can be written in terms of the electromagnetic form factors as

$$\begin{aligned} H_{\mu\nu}(T) &= -\overline{Q}G_M^2\tilde{g}_{\mu\nu} + \frac{\overline{Q}}{M^2} \\ &\times\left[G_M^2 + 4(1+\eta)^{-1}G_Q\left(G_C + \frac{\eta}{3}G_Q + \eta G_M\right)\right] \\ &\times\tilde{p}_{1\mu}\tilde{p}_{1\nu} - 2\eta G_M(G_M + 2G_Q)(\tilde{p}_{1\mu}\tilde{Q}_\nu + \tilde{p}_{1\nu}\tilde{Q}_\mu) \\ &- q^2(1+\eta)G_M^2\tilde{Q}_{\mu\nu}, \end{aligned} \tag{12}$$

where

$$\tilde{Q}_\mu = Q_{\mu\nu}q_\nu - \frac{q_\mu}{q^2}\bar{Q}, \quad \tilde{Q}_\mu q_\mu = 0,$$
$$\tilde{Q}_{\mu\nu} = Q_{\mu\nu} + \frac{q_\mu q_\nu}{q^4}\bar{Q} - \frac{q_\nu q_\alpha}{q^2}Q_{\mu\alpha} - \frac{q_\mu q_\alpha}{q^2}Q_{\nu\alpha}, \quad (13)$$
$$\tilde{Q}_{\mu\nu}q_\nu = 0, \quad \bar{Q} = Q_{\alpha\beta}q_\alpha q_\beta.$$

The target deuteron quadrupole polarization tensor $Q_{\mu\nu}$ satisfies the conditions

$$Q_{\mu\nu} = Q_{\nu\mu}, \quad Q_{\mu\mu} = 0, \quad p_{1\mu}Q_{\mu\nu} = 0. \quad (14)$$

Using the definitions of cross section (2) and leptonic (3) and hadronic (9) tensors, we can easily derive the expression for the unpolarized differential cross section in terms of the invariant variables suitable for the calculation of the radiative corrections,

$$\frac{d\sigma_b^{un}}{dQ^2} = \frac{\pi\alpha^2}{Q^4}\left\{\frac{2\rho}{V}W_1 + \frac{W_2}{V\tau}[1-\rho(1+\tau)]\right\}, \quad (15)$$
$$Q^2 = -q^2 = 2k_1k_2.$$

In the laboratory system, this expression can be written in a more familiar form,

$$\frac{d\sigma_b^{un}}{d\Omega} = \sigma_{NS}\left\{A(Q^2) + B(Q^2)\tan^2\frac{\theta_e}{2}\right\}, \quad (16)$$

where θ_e is the electron scattering angle, σ_{NS} is the Mott cross section multiplied by the deuteron recoil factor

$$\left(1 + 2(\varepsilon_1/M)\sin^2\frac{\theta_e}{2}\right)^{-1},$$

and ε_1 is the electron beam energy. The two structure functions $A(Q^2)$ and $B(Q^2)$ are quadratic combinations of the three electromagnetic form factors describing the deuteron structure,

$$A(Q^2) = G_C^2(Q^2) + \frac{8}{9}\eta^2 G_Q^2(Q^2) + \frac{2}{3}\eta G_M^2(Q^2),$$

$$B(Q^2) = \frac{4}{3}\eta(1+\eta)G_M^2(Q^2), \quad \eta = \frac{Q^2}{4M^2}.$$

Before writing similar distributions for the scattering of polarized particles, we note that, in this case, there may exist, in general, an azimuthal correlation between the reaction plane and the plane ($\mathbf{p}_2$, $\mathbf{W}$) if the recoil deuteron is detected (here, $\mathbf{W}$ is 3-vector of the deuteron polarization). However, in the Born approximation, with the P-invariance of the electromagnetic interaction taken into account, such a correlation is absent. In what follows, we consider the situation where the vector $\mathbf{W}$ belongs to the reaction plane and the corresponding azimuthal angle equals to zero. Therefore, there exist only two independent components of $\mathbf{W}$, which we call the longitudinal and transverse ones. It is convenient to use the covariant parametrization of the deuteron polarization 4-vector in terms of the 4-momenta of the particles in the reaction. This parametrization is ambiguous and depends on the directions along which the longitudinal and transverse components of the deuteron polarization in its rest frame are defined.

As mentioned above, we have to define the longitudinal W^L and transverse W^T 4-vectors. In our case, it is natural to choose the longitudinal direction in the laboratory system along the 3-momentum $\mathbf{q}$ and the transverse direction perpendicular to the longitudinal one in the reaction plane. The corresponding 4-vectors can be written as [5]

$$W_\mu^{(L)} = \frac{2\tau q_\mu - \rho p_{1\mu}}{M\sqrt{\rho(4\tau+\rho)}},$$
$$W_\mu^{(T)} = \frac{(4\tau+\rho)k_{1\mu} - (1+2\tau)q_\mu - (2-\rho)p_{1\mu}}{\sqrt{V(4\tau+\rho)(1-\rho-\rho\tau)}}. \quad (17)$$

This leads to simple expressions for the corresponding part of the hadronic tensor,

$$H_{\mu\nu}^L(V) = -\frac{iG_M^2}{4\tau}(\mu\nu qp_1)\sqrt{\rho(4\tau+\rho)},$$
$$H_{\mu\nu}^T(V) = -\frac{iG_MG}{4}[(4\tau+\rho)(\mu\nu qk_1) \quad (18)$$
$$-(2-\rho)(\mu\nu qp_1)]\sqrt{\frac{(4\tau+\rho)}{\tau(1-\rho-\rho\tau)}}.$$

The polarization-dependent parts of the cross section, due to the vector polarization of the deuteron target, are given by

$$\frac{d\sigma_b^L}{dQ^2} = \frac{\pi\alpha^2}{4\tau V^2}\frac{2-\rho}{\rho}\sqrt{\rho(4\tau+\rho)}G_M^2, \quad (19)$$

$$\frac{d\sigma_b^T}{dQ^2} = \frac{\pi\alpha^2}{VQ^2}\sqrt{\frac{(4\tau+\rho)(1-\rho-\rho\tau)}{\tau}}G_MG, \quad (20)$$

where we assumed that λ in Eq. (3) is equal to one and the deuteron-target polarization degree (longitudinal or transverse) is 100%.

In the laboratory system, these parts of the cross section can be written as

$$\frac{d\sigma_b^L}{dQ^2} = \frac{\pi}{\varepsilon_2^2}\eta\sigma_{NS}\sqrt{(1+\eta)\left(1+\eta\sin^2\frac{\theta_e}{2}\right)} \times \tan\frac{\theta_e}{2}\sec\frac{\theta_e}{2}G_M^2, \quad (21)$$

$$\frac{d\sigma_b^T}{dQ^2} = 2\frac{\pi}{\varepsilon_2^2}\sigma_{NS}\tan\frac{\theta_e}{2}\sqrt{\eta(1+\eta)}G_M\left(G_C+\frac{\eta}{3}G_Q\right), \quad (22)$$

where ε_2 is the scattered electron energy.

It is worth noting that the ratio of the longitudinal polarization asymmetry $A^L = d\sigma_b^L/d\sigma_b$ to the transverse one, $A^T = d\sigma_b^T/d\sigma_b$,

$$\frac{A^L}{A^T} = \frac{2-\rho}{4}\sqrt{\frac{\rho}{\tau(1-\rho-\rho\tau)}}\frac{G_M}{G} \quad (23)$$

is expressed in terms of the deuteron form factors G_M and G in the same way as the corresponding ratio in the case of elastic electron–proton scattering is expressed via the proton electromagnetic form factors G_{Mp} and G_{Ep} [17, 18]. This is a direct consequence of the relation between the proton $H^p_{\mu\nu}(V)$ and deuteron $H_{\mu\nu}(V)$ hadronic tensors, which depend on the proton and deuteron vector polarization, respectively,

$$H_{\mu\nu}(V)(G_M, G) = -\frac{4\tau+\rho}{8\tau}H^p_{\mu\nu}(V)(G_{Mp}, G_{Ep}). \quad (24)$$

We now consider the tensor-polarized deuteron target. For completeness, we introduce the 4-vector

$$W_\mu^{(N)} = \frac{2\varepsilon_{\mu\lambda\rho\sigma}p_{1\lambda}k_{1\rho}k_{2\sigma}}{V\sqrt{V\rho(1-\rho-\rho\tau)}}, \quad (25)$$

which is orthogonal to the reaction plane. It can then be verified that the set of the 4-vectors $W_\mu^{(I)}$, $I = L, T, N$, satisfies the conditions

$$W_\mu^{(\alpha)}W_\mu^{(\beta)} = -\delta_{\alpha\beta}, \quad W_\mu^{(\alpha)}p_{1\mu} = 0, \quad \alpha, \beta = L, T, N.$$

If one more 4-vector $W_\mu^{(0)} = p_{1\mu}/M$ is added to the set of the 4-vectors defined by Eqs. (17) and (25), we obtain the complete set of orthogonal 4-vectors with the properties

$$W_\mu^{(m)}W_\nu^{(m)} = g_{\mu\nu}, \quad W_\mu^{(m)}W_\mu^{(n)} = g_{mn}, \quad m, n = 0, L, T, N. \quad (26)$$

This allows us to express the deuteron quadrupole polarization tensor in the general case as

$$Q_{\mu\nu} = W_\mu^{(m)}W_\nu^{(n)}R_{mn} \equiv W_\mu^{(\alpha)}W_\nu^{(\beta)}R_{\alpha\beta}, \quad R_{\alpha\beta} = R_{\beta\alpha}, \quad R_{\alpha\alpha} = 0, \quad (27)$$

because the components R_{00}, $R_{0\alpha}$, and $R_{\alpha 0}$ are identically equal to zero due to the condition $Q_{\mu\nu}p_{1\nu} = 0$. The quantities $R_{\alpha\beta}$ are in fact the degrees of the tensor polarization of the deuteron target. In the Born approximation, the components R_{NL} and R_{NT} do not contribute and expansion (27) can be rewritten in the standard form

$$Q_{\mu\nu} = \left[W_\mu^{(L)}W_\nu^{(L)} - \frac{1}{2}W_\mu^{(T)}W_\nu^{(T)}\right]R_{LL} + \frac{1}{2}W_\mu^{(T)}W_\nu^{(T)}(R_{TT}-R_{NN}) + (W_\mu^{(L)}W_\nu^{(T)} + W_\mu^{(T)}W_\nu^{(L)})R_{LT}, \quad (28)$$

where we took into account that

$$R_{LL} + R_{TT} + R_{NN} = 0.$$

The part of the cross section that depends on the tensor polarization of the deuteron target can be written as

$$\frac{d\sigma_b^Q}{dQ^2} = \frac{d\sigma_b^{LL}}{dQ^2}R_{LL} + \frac{d\sigma_b^{TT}}{dQ^2}(R_{TT}-R_{NN}) + \frac{d\sigma_b^{LT}}{dQ^2}R_{LT}, \quad (29)$$

where

$$\frac{d\sigma_b^{LL}}{dQ^2} = \frac{\pi\alpha^2}{Q^4}2(1-\rho-\tau\rho) \times \eta\left\{8G_CG_Q + \frac{8}{3}\eta G_Q^2 + \frac{2-2\rho+2\tau\rho+\rho^2}{2(1-\rho-\tau\rho)}G_M^2\right\}, \quad (30)$$

$$\frac{d\sigma_b^{TT}}{dQ^2} = \frac{\pi\alpha^2}{Q^4}2\eta(1-\rho-\tau\rho)G_M^2, \quad (31)$$

$$\frac{d\sigma_b^{LT}}{dQ^2} = -\frac{\pi\alpha^2}{Q^4}4\eta(2-\rho)\sqrt{\frac{\rho(1-\rho-\tau\rho)}{\tau}}G_QG_M. \quad (32)$$

In the laboratory frame, this part of the cross section can be written as

$$\frac{d\sigma_b^Q}{dQ^2} = \frac{\pi}{\varepsilon_2^2} \times \sigma_{NS}[S_{LL}R_{LL} + S_{TT}(R_{TT}-R_{NN}) + S_{LT}R_{LT}], \quad (33)$$

where

$$S_{LL} = \frac{1}{2}\left\{8\eta G_C G_Q + \frac{8}{3}\eta^2 G_Q^2 + \eta\left[1 + 2(1+\eta)\tan^2\frac{\theta_e}{2}\right]G_M^2\right\}, \quad (34)$$

$$S_{TT} = \frac{1}{2}\eta G_M^2, \quad (35)$$

$$S_{LT} = -4\eta\sqrt{\eta + \eta^2\sin^2\frac{\theta_e}{2}}\sec\frac{\theta_e}{2}G_Q G_M. \quad (36)$$

If the longitudinal direction is determined by the recoil deuteron 3-momentum, relations (18) and (21) are not affected by hard photon radiation in the lepton part of the interaction (this corresponds to the use of the so-called hadronic variables, see below) because

$$\mathbf{q} = \mathbf{p}_2 - \mathbf{p}_1.$$

However, when this direction is reconstructed using the 3-momentum of the scattered electron (lepton variables), these relations break down because

$$\mathbf{q} \neq \mathbf{k}_1 - \mathbf{k}_2$$

in this case. This means that, in the leptonic variables, parametrization (17) is unstable and radiation of a hard photon by the electron leads to a mixture of the longitudinal and transverse polarizations.

This mixture can be eliminated by taking the longitudinal direction along the 3-momentum of the initial electron. The corresponding parametrization of the 4-vector polarizations is [19]

$$W_\mu^{(l)} = \frac{2\tau k_{1\mu} - p_{1\mu}}{M}, \qquad W_\mu^{(t)} = \frac{k_{2\mu} - (1-\rho-2\rho\tau)k_{1\mu} - \rho p_{1\mu}}{\sqrt{V\rho(1-\rho-\rho\tau)}}. \quad (37)$$

The hadronic tensors $H_{\mu\nu}^{l,t}$ then have the form

$$H_{\mu\nu}^l = -i\frac{4\tau+\rho}{4\tau}\times\left\{G\left[-2\tau(\mu\nu qk_1) + \frac{2\tau(2-\rho)}{4\tau+\rho}(\mu\nu qp_1)\right] + G_M\frac{\rho(1+2\tau)}{4\tau+\rho}(\mu\nu qp_1)\right\}G_M, \quad (38)$$

$$H_{\mu\nu}^t = -i\sqrt{\frac{\rho\tau}{1-\rho-\rho\tau}}\times\left\{G(1+2\tau)\left[\frac{2-\rho}{4\tau}(\mu\nu qp_1) - \frac{4\tau+\rho}{4\tau}(\mu\nu qk_1)\right] - G_M\frac{1-\rho-\rho\tau}{2\tau}(\mu\nu qp_1)\right\}G_M. \quad (39)$$

In the case of scattering off a polarized target, the tensors $H_{\mu\nu}^{L,T}$ and $H_{\mu\nu}^{l,t}$ are connected by the trivial relations

$$H_{\mu\nu}^L = \cos\theta H_{\mu\nu}^l + \sin\theta H_{\mu\nu}^t, \qquad H_{\mu\nu}^T = -\sin\theta H_{\mu\nu}^l + \cos\theta H_{\mu\nu}^t,$$
$$\cos\theta = -(W^{(L)}W^{(l)}) = \frac{\rho(1+2\tau)}{\sqrt{\rho(4\tau+\rho)}}, \qquad \sin\theta = -(W^{(L)}W^{(t)}) = -2\sqrt{\frac{\tau(1-\rho-\rho\tau)}{4\tau+\rho}}. \quad (40)$$

Using these relations, we can write the polarization-dependent parts of the Born cross section, which correspond to parametrization (37), as

$$\frac{d\sigma_b^l}{dQ^2} = \cos\theta\frac{d\sigma_b^L}{dQ^2} - \sin\theta\frac{d\sigma_b^T}{dQ^2}, \qquad \frac{d\sigma_b^t}{dQ^2} = \sin\theta\frac{d\sigma_b^L}{dQ^2} + \cos\theta\frac{d\sigma_b^T}{dQ^2}, \quad (41)$$

where $d\sigma_b^L/dQ^2$ and $d\sigma_b^T/dQ^2$ are defined by Eqs. (19) and (20). Therefore, we can write

$$\frac{d\sigma_b^l}{dQ^2} = \frac{\pi\alpha^2}{V^2}\times\left[\frac{1+2\tau}{4\tau}(2-\rho)G_M + \frac{2}{\rho}(1-\rho-\rho\tau)G\right]G_M, \quad (42)$$

$$\frac{d\sigma_b^t}{dQ^2} = \frac{\pi\alpha^2}{VQ^2}\sqrt{\frac{\rho(1-\rho-\rho\tau)}{\tau}}\times\left[-\frac{1}{2}(2-\rho)G_M + (1+2\tau)G\right]G_M. \quad (43)$$

In the case of the tensor polarization, the relations that represent an analog of Eq. (41) become

$$\frac{d\sigma_b^Q}{dQ^2} = X_{IJ}\frac{d\sigma_b^{IJ}}{dQ^2}, \quad I, J = L, T, \quad (44)$$

where the partial cross sections $d\sigma_b^{IJ}/dQ^2$ are defined by Eq. (29) as the coefficients in front of the respective

quantities R_{LL}, $R_{TT} - R_{NN}$, and R_{LN}, and the entries of the matrix X_{IJ} are

$$X_{LL} = \frac{1}{4}(1 + 3\cos 2\theta)R_{ll} + \frac{1}{4}(1 - \cos 2\theta)(R_{tt} - R_{nn}) + \sin 2\theta R_{lt},$$

$$X_{TT} = \frac{3}{4}(1 - \cos 2\theta)R_{ll} + \frac{1}{4}(3 + \cos 2\theta)(R_{tt} - R_{nn}) - \sin 2\theta R_{lt}, \quad (45)$$

$$X_{lt} = -\frac{1}{4}\sin 2\theta[3R_{ll} - (R_{tt} - R_{nn})] + \cos 2\theta R_{lt}.$$

As we can see, the polarization-dependent part of the cross section is now expressed in terms of the new polarization parameters R_{ll}, $R_{tt} - R_{nn}$, and R_{lt} defined in accordance with the directions given by Eq. (37), and the coefficients in front of these quantities in the right-hand side of Eq. (44) determine the corresponding partial cross sections $d\sigma_b^{ij}/dQ^2$.

We now consider the scattering off the unpolarized target in the case where the recoil deuteron polarization is measured. In this case, we can obtain both the vector and tensor polarizations of the recoil deuteron using the results given above. For this, we note that the longitudinal and transverse 4-vectors $S^{(L)}$ and $S^{(T)}$, which satisfy the relations $S^2 = -1$ and $(Sp_2) = 0$, are

$$S^{(L)} = \frac{2\tau q_\mu + \rho p_{2\mu}}{M\sqrt{\rho(4\tau + \rho)}}, \quad S^{(T)} = W^{(T)}. \quad (46)$$

The part $H_{\mu\nu}(V)$ of the hadronic tensor can be derived from Eq. (11) by the substitution $W \longrightarrow S$, $p_1 \longleftrightarrow -p_2$. This actually means that we have to replace (Wp_2) in the right-hand side of Eq. (11) with (Sp_1). The vector polarization of the recoil deuteron (longitudinal P^L or transverse P^T) is defined as the ratio of the polarization-dependent part of the cross section to the unpolarized part. Taking into account that $(S^L p_1) = -(W^L p_2)$, we conclude that

$$P^L = -A^L, \quad P^T = A^T, \quad (47)$$

where A^L and A^T are the respective asymmetries for the scattering off the 100%-polarized deuteron target.

Here, we want to draw the reader's attention to the fact that determination of G_M/G by measuring the ratio A^L/A^T in the scattering off a polarized deuteron target is more attractive than by measuring the ratio P^L/P^T in the polarization transfer process because the second scattering is necessary in the latter case. This decreases the corresponding event number by about two orders [20], substantially increasing the statistical error. The problem with the depolarization effect that appears in the scattering of a high-intensity electron beam on the polarized solid target can be avoided using the polarized gas deuteron target [3].

By analogy, the components of the tensor polarization of the recoil deuteron are defined by the ratios of the corresponding partial cross sections to the unpolarized one,

$$\tilde{R}_{LL} = \frac{d\sigma_b^{LL}}{d\sigma_b^{un}}, \quad \tilde{R}_{LT} = \frac{d\sigma_b^{LT}}{d\sigma_b^{un}}, \quad \tilde{R}_{TT} - \tilde{R}_{NN} = \frac{d\sigma_b^{TT}}{d\sigma_b^{un}}. \quad (48)$$

Part $H_{\mu\nu}(T)$ of the hadronic tensor can be derived from Eq. (12) by changing the sign in the term proportional to $G_M(G_M + 2G_Q)$. Straightforward calculations using this updated tensor and parametrization (46) lead to the following results. First, both diagonal partial cross sections in the right-hand side of Eq. (48) are the same as defined by Eq. (29) for the scattering off the polarized target, and second, the partial cross section $d\sigma_b^{LT}/dQ^2$ changes sign compared with the cross section in Eq. (29).

3. RADIATIVE CORRECTIONS

The total radiative correction can be divided into model-independent and model-dependent contributions. The model-independent radiative correction includes all QED corrections to the lepton part of the interaction and insertion of the vacuum polarization into the exchange photon propagator. The model-dependent radiative correction involves additional couplings of the photon with the off-mass-shell hadron and comes from box-type diagrams, hadronic vertex functions, hadron contribution to vacuum polarization, etc. It can be analyzed at the level allowed by the current knowledge of the hadronic structure; as a rule, the corresponding contribution is added to the systematic error.

The standard practice of data analysis in *ep* and *ed* scatterings is that the model-independent radiative correction is taken into account with the accuracy allowed by theoretical calculations. The reason is that it makes the main contribution due to the smallness of the electron mass and can be calculated without any additional assumptions. Therefore, the model-independent radiative correction is calculated theoretically and simply subtracted from the observed quantities, or Monte Carlo generators constructed on the basis of these calculations are implemented into the codes of the data analysis. In this paper, we calculate only the model-independent radiative correction; we bear this in mind in what follows.

There exist two sources of radiative corrections when the corrections of order α are taken into account. The first is caused by virtual and soft photon emission that cannot affect the kinematics of process (1). The second arises due to the radiation of a hard photon,

$$e^-(k_1) + D(p_1) \longrightarrow e^-(k_2) + \gamma(k) + D(p_2), \qquad (49)$$

because cuts on the event selection used in the current experiments allow photons to be radiated with an energy of about 100 MeV or even more [6, 20]. Such photons cannot be interpreted as "soft." The form of the radiative correction caused by the contribution due to the hard photon emission depends strongly on the choice of variables used to describe process (49) [21].

The hadronic variables were used formerly to compute the radiative correction in the elastic and deep-inelastic polarized electron–proton scattering [21, 22]. As noted in [21], the form and value of the radiative correction in the hadronic variables differ substantially from the radiative correction calculated in the leptonic variables. We want to point out that the results in [22] can be used for the elastic *ep* scattering and relations (10) and (22) can be used to calculate the radiative correction in the elastic unpolarized and polarized *ed* scattering in the case of the deuteron vector polarization. Here, we also calculate the radiative correction in the case of the deuteron tensor polarization, which is absent in [22] because the proton has spin 1/2. Our goal is to obtain physically transparent formulas for the radiatively corrected cross sections, which are absent in [22], and to generalize them with the higher orders of the coupling constant α taken into account by simple exponentiation of the leading contributions.

In contrast to [22], we assume at the very beginning that, in reaction (49), the recoil deuteron is detected and the 4-momentum

$$q = p_2 - p_1 = k_1 - k_2 - k$$

is fixed. Because neither the scattered electron nor the hard photon is detected, the complete integration over the 3-momenta of these undetected particles must be performed.

In calculating the radiative correction using the hadronic variables, it is very convenient to use the method of invariant integration. In this method, integration of the leptonic tensor $L^{\gamma}_{\mu\nu}$ (with the emission of an additional photon taken into account) over the variables of the scattered electron and the emitted additional photon is performed before the contraction of the leptonic and hadronic tensors. At the beginning, we use the overall 4-dimension δ-function to eliminate the k_2 momentum and then perform analytic integration with respect to the photon 3-momentum in the special system where

$$\mathbf{k}_1 + \mathbf{p}_1 - \mathbf{p}_2 = 0.$$

It is convenient to introduce the dimensionless hadronic variables

$$x = \frac{q^2}{2k_1q}, \quad y = -\frac{2k_1q}{V}, \quad xy = \rho, \qquad (50)$$

which characterize inelasticity due to the hard-photon emission in the lepton block: if the photon is not radiated, then $x = 1$ and $y = \rho$. The quantity $1 - x$ actually represents the energy fraction of the collinear photon radiated by the initial-state electron. It is easy to verify this statement because

$$q = xk_1 - k_2, \quad k_2^2 = k_1^2 \approx 0$$

in this case.

The use of these variables in the framework of our approach allows us to bypass the complication that comes from the Gram determinant and appears in the standard method developed in [23] and used later in [22, 24]. This makes it possible to simplify the calculations and write physically transparent expressions for both polarized and unpolarized cross sections.

Using the above strategy, we start from the following expression for the cross section of process (49) in the hadronic variables:

$$d\sigma = \frac{\alpha^2}{VQ^4}L^{\gamma}_{\mu\nu}H_{\mu\nu}\frac{d^3p_2}{E_2}\frac{\alpha}{4\pi^2}\frac{d^3k}{\omega}\delta(k_2^2 - m^2). \qquad (51)$$

Here, ω is the photon energy and m is the electron mass. The leptonic tensor corresponding to the hard-photon radiation is well known [25, 26]. It can be written as

$$L^{\gamma}_{\mu\nu} = L^{un}_{\mu\nu} + L^{p}_{\mu\nu}.$$

Its unpolarized symmetric part is

$$L^{un}_{\mu\nu} = 2\left[\frac{(q^2-t)^2 + (q^2-s)^2}{st} - 2m^2q^2\left(\frac{1}{s^2} + \frac{1}{t^2}\right)\right]\tilde{g}_{\mu\nu} + 8\left(\frac{q^2}{st} - \frac{2m^2}{s^2}\right)\tilde{k}_{1\mu}\tilde{k}_{1\nu} + 8\left(\frac{q^2}{st} - \frac{2m^2}{t^2}\right)\tilde{k}_{2\mu}\tilde{k}_{2\nu}, \qquad (52)$$

and the antisymmetric part, arising due to the longitudinal beam polarization, is

$$L^{p}_{\mu\nu} = 4i(\mu\nu q\rho)\left\{k_{1\rho}\left[\frac{q^2-s}{st} - 2m^2\left(\frac{1}{s^2} + \frac{1}{t^2}\right)\right] + k_{2\rho}\left[\frac{q^2-t}{st} - \frac{2m^2s}{(q^2-s)t^2}\right]\right\}, \qquad (53)$$

$$t = -2kk_1, \quad s = 2kk_2 = -Q^2 + Vy,$$

$$\tilde{a}_{\mu} = a_{\mu} - \frac{aq}{q^2}q_{\mu}.$$

After removing the overall δ-function, it is necessary to

calculate the quantity

$$L^i_{\mu\nu} = \int L^\gamma_{\mu\nu}\frac{d^3k}{\omega}\delta(k_2^2 - m^2).$$

We calculate it using the method of invariant integration. We first consider the case of the unpolarized electron beam. With the *P*-invariance of the electromagnetic interaction and gauge invariance of the quantity $L^i_{\mu\nu}$ taken into account, we can represent this quantity in the general form as

$$L^{iun}_{\mu\nu} = A\tilde{g}_{\mu\nu} + B\tilde{k}_{1\mu}\tilde{k}_{1\nu}. \tag{54}$$

The two functions *A* and *B* can be obtained by contracting the left- and right-hand sides of this equation with the respective tensors $\tilde{g}_{\mu\nu}$ and $\tilde{k}_{1\mu}\tilde{k}_{1\nu}$. As a result, we obtain two equations for the two unknowns *A* and *B*,

$$\begin{aligned} I_1 &= 3A - \frac{(q^2 - s)^2}{4q^2}B, \\ I_2 &= \frac{(q^2 - s)^2}{4q^2}\left[-A + \frac{(q^2 - s)^2}{4q^2}B\right], \end{aligned} \tag{55}$$

where we introduce the notation

$$I_1 = \int L^{un}_{\mu\nu}\tilde{g}_{\mu\nu}\frac{d^3k}{\omega}\delta(k_2^2 - m^2),$$

$$I_2 = \int L^{un}_{\mu\nu}\tilde{k}_{1\mu}\tilde{k}_{1\nu}\frac{d^3k}{\omega}\delta(k_2^2 - m^2).$$

Next, we must integrate over the photon phase space in the integrals I_1 and I_2. Because the quantity $\delta(k_2^2 - m^2)d^3k/\omega$ that enters the integrands in I_1 and I_2 is Lorentz invariant, we can take any coordinate system to do this integration. The most convenient one is the coordinate system where

$$\mathbf{k}_1 + \mathbf{p}_1 - \mathbf{p}_2 = 0.$$

In fact, this is the center-of-mass system for the radiated photon and the scattered electron, and, therefore, the polar (ϑ) and azimuthal (ϕ) angles of the radiated photon cover the entire phase space. We therefore have

$$\begin{aligned} \frac{d^3k}{\omega}\delta(k_2^2 - m^2) &= \frac{1 - x}{4R_x}d\phi\, d\cos\vartheta, \\ 0 < \phi < 2\pi, \quad -1 < \cos\vartheta < 1, \quad R_x &= 1 - x + \frac{m^2}{Q^2}, \end{aligned} \tag{56}$$

where the *z* axis is chosen along the direction of the initial electron 3-momentum. In writing the quantity R_x, we set $x = 1$ in the coefficient in front of m^2/Q^2.

The energies of all particles and the polar angle of the initial deuteron in this system can be expressed in terms of the invariant variables as

$$\begin{aligned} \omega &= \frac{Vy(1 - x)}{2\sqrt{R}}, \quad \varepsilon_1 = \frac{Vy + 2m^2}{2\sqrt{R}}, \\ \varepsilon_2 &= \frac{Vy(1 - x) + 2m^2}{2\sqrt{R}}, \quad E_1 = \frac{V(1 - \rho)}{2\sqrt{R}}, \\ E_2 &= \frac{V(1 - y + \rho)}{2\sqrt{R}}, \quad \cos\theta_1 = \frac{2E_1\varepsilon_1 - V}{2|\mathbf{p}_1||\mathbf{k}_1|}, \\ R &= VyR_x, \end{aligned} \tag{57}$$

where ε_1 and E_1 (ε_2 and E_2) are the respective energies of the initial (final) electron and deuteron.

The necessary angular integrals are given by

$$\begin{aligned} \int\frac{1}{-t} &= \frac{2R_x}{Vy(1 - x)}L, \\ \int\frac{2\chi}{-t} &= \frac{2}{y}[R_x(L - 2) + 1 - \rho], \\ \int t &= -\frac{Vy(1 - x)}{R_x}, \\ \int\frac{4\chi^2}{-t} &= \frac{2V(1 - x)}{y} \\ &\times\left[R_x(L - 3) + \frac{(1 - \rho)^2}{2R_x} + 1 - \rho - y\tau\right], \\ \int 2\chi &= \frac{V(1 - x)(1 - \rho)}{R_x}, \\ \int\frac{2m^2}{t^2} &= \frac{4R_x}{Vy(1 - x)^2}, \quad \chi = kp_1, \end{aligned} \tag{58}$$

where we use the short notation for definite integral and the quantity *L*,

$$\int \equiv \int_0^{2\pi}\frac{d\phi}{2\pi}\int_{-1}^{1}d\cos\vartheta, \quad L = \ln\frac{Q^2}{m^2xR_x},$$

and neglect the terms of order (m^2/Q^2) whenever possible.

Calculating the integrals I_1 and I_2 as described above and solving the system of two equations (55), we find

that the functions A and B are given by

$$A = -2\pi[F^p(x, Q^2/m^2) + 3(1-x)],$$

$$B = 8\pi\frac{x^2}{Q^2}[F^p(x, Q^2/m^2) + 3 + x],$$

$$F^p(x, Q^2/m^2) = \frac{1+x^2}{1-x}\ln\frac{Q^2}{m^2 x R_x} - \frac{4}{1-x} + \frac{1}{2R_x} - \frac{m^2}{2Q^2R_x^2} + 1 + 4x. \quad (59)$$

The contraction of the unpolarized parts of the leptonic tensor (integrated over the photon phase space) and the hadronic tensor for radiative process (49) is given by

$$L^{iun}_{\mu\nu}H_{\mu\nu}(0) = -\left(3A + \frac{Vy^2}{4\rho}B\right)W_1 + \left[(1+\eta)A + \frac{V(2-y)^2}{16\tau}B\right]W_2. \quad (60)$$

To write the respective contractions in the polarized case, we have to take into account that the parametrization of the polarization 4-vectors $S^{(L,T)}$ in the radiative process differs somewhat from the Born expressions given by Eq. (46) (we here consider the polarized recoil deuteron for definiteness). Formally, they can be derived by the substitution

$$\rho \longrightarrow y, \quad k_1 \longrightarrow xk_1, \quad \tau \longrightarrow \tau/x, \quad V \longrightarrow xV \quad (61)$$

in Eq. (46).

The contraction of the unpolarized part of the leptonic tensor (integrated over the photon phase space) and the hadronic $H_{\mu\nu}(T)$ tensor for radiative process (49) is given by

$$L^{iun}_{\mu\nu}H_{\mu\nu}(T) = \left\{-A\bar{Q} + VB\left[\frac{(1-y)^2}{4\tau}\bar{Q} + (2\eta(1-\eta)-y)Q_1 + \rho(1+\eta)Q_{11}\right]\right\}G_M^2 + (2-y)VB\left[2\eta Q_1 - \frac{y-2\eta(1-y)}{4\tau+\rho}\bar{Q}\right]G_M G_Q + \frac{\bar{Q}}{2}\left[4A + \frac{(2-y)^2}{4\tau+\rho}VB\right]G_Q G, \quad (62)$$

where

$$Q_1 = Q_{\mu\nu}q_\mu k_{1\nu}, \quad Q_{11} = Q_{\mu\nu}k_{1\mu}k_{1\nu}.$$

Taking into account the relation

$$\frac{d^3p_2}{E_2} = \pi\rho dQ^2\frac{dx}{x^2},$$

which holds for the recoil-deuteron phase space, after some algebra, we derive the following representation for unpolarized cross section of reaction (49):

$$\frac{d\sigma^{un}_H}{dQ^2}(k_1) = \frac{\alpha}{2\pi}\int\limits_{x_m}^{1-\Delta}\left[\frac{d\sigma^{un}_b}{dQ^2}(xk_1)F^{un}(x, Q^2/m^2) + \frac{\pi\alpha^2}{Q^4}\frac{W_2}{M^2}f(x,\rho,\tau)\right]dx, \quad (63)$$

$$F^{un}(x, Q^2/m^2) = F^p(x, Q^2/m^2) + 3 - 4x,$$

$$f(x,\rho,\tau) = 3(x-\rho) + \frac{\rho}{x}\left(\frac{\rho}{2} - \tau\right).$$

Here, Δ is the minimum-energy fraction of the hard photon and x_m depends on the experimental cuts for the photon energy.

For the partial cross sections in the case of tensor polarization of the recoil deuteron in radiative process (49), we obtain the expressions

$$\frac{d\sigma^{LL}_H}{dQ^2}(k_1) = \frac{\alpha^3}{V^2}(1+\eta)\ln x_m(G_M^2 - 2G_Q G) + \frac{\alpha}{2\pi}\int\limits_{x_m}^{1-\Delta}\frac{d\sigma^{LL}_b}{dQ^2}(xk_1)F^T(x, Q^2/m^2)dx,$$

$$\frac{d\sigma^{TT}_H}{dQ^2}(k_1) = \frac{\alpha}{2\pi}\int\limits_{x_m}^{1-\Delta}\frac{d\sigma^{TT}_b}{dQ^2}(xk_1)F^T(x, Q^2/m^2)dx, \quad (64)$$

$$\frac{d\sigma^{LT}_H}{dQ^2}(k_1) = \frac{\alpha}{2\pi}\int\limits_{x_m}^{1-\Delta}\frac{d\sigma^{LT}_b}{dQ^2}(xk_1)F^T(x, Q^2/m^2)dx,$$

where

$$F^T(x, Q^2/m^2) = F^P(x, Q^2/m^2) + 3 - x.$$

We now consider the case of the longitudinally polarized electron beam and calculate the necessary integral, where $L^\gamma_{\mu\nu} = L^p_{\mu\nu}$, using the method of invariant integration. Taking the P-invariance of the electromagnetic interaction and gauge invariance of the quan-

tity $L^{i}_{\mu\nu}$ into account, we can represent this quantity in the general form as

$$L^{ip}_{\mu\nu} = iC(\mu\nu qk_1). \quad (65)$$

The unknown function C can be obtained by contracting the left- and right-hand sides of this equation with the tensor $(\mu\nu qk_1)$. As a result, we obtain the following expression for C:

$$\frac{i}{2}(q^2 - s)C = \int L^{p}_{\mu\nu}(\mu\nu qk_1)\frac{d^3k}{\omega}\delta(k_2^2 - m^2). \quad (66)$$

Calculating this integral as explained above, we obtain

$$C = \frac{2\pi x}{Q^2}F^p(x, Q^2/m^2).$$

The contraction of the polarized part of the leptonic tensor (integrated over the photon phase space) and the hadronic $H_{\mu\nu}(V)$ tensor for radiative process (49) is given by

$$L^{ip}_{\mu\nu}H^{L}_{\mu\nu} = -\frac{1}{2}(2 - y)\eta V^2 C\sqrt{\rho(4\tau + \rho)}G_M^2, \quad (67)$$

$$L^{ip}_{\mu\nu}H^{T}_{\mu\nu} = -\frac{1}{4}V^2 C[2\rho(1 - y) + y^2(\rho + 2\tau)] \times \sqrt{\frac{x(4\tau + \rho)}{\tau(x - \rho - y\tau)}}G_M G. \quad (68)$$

After some algebra, we derive the following representation for the parts of the cross section that depend on the vector polarization of the recoil deuteron:

$$\frac{d\sigma^{L,T}_{H}}{dQ^2}(k_1) = \frac{\alpha}{2\pi} \times \int_{x_m}^{1-\Delta}\frac{d\sigma^{L,T}_{b}}{dQ^2}(xk_1)F^p(x, Q^2/m^2)dx. \quad (69)$$

The infrared auxiliary parameter $\Delta \ll 1$ is related to the minimal energy of the hard photon in the chosen coordinate system, and the lower integration limit is defined by its maximum value, which depends on the experimental cuts on the event selection in the experimental measurement of the observables in elastic electron–deuteron scattering:

$$\Delta = \frac{\omega_{\min}}{\varepsilon_1}, \quad x_m = 1 - \frac{\omega_{\max}}{\varepsilon_1}.$$

For example, if the lost invariant mass $M_{\max}$ (of the scattered electron and the undetected additional hard photon) is allowed,

$$(k_1 + p_1 - p_2)^2 \le M^2_{\max},$$

then

$$x_m = \frac{Q^2}{Q^2 + M^2_{\max}}.$$

On the other hand, the quantity x_m cannot be arbitrary (but must, of course, be smaller than unity) even if no experimental constraints on the event selection are used. The restriction on x_m follows from the inequality

$$x^2 - x\rho - \rho\tau > 0,$$

which reflects the obvious relation

$$-q^2 < -q^2_{\max} = \frac{x^2V^2}{xV + M^2}$$

for radiative process (49). In any case, we therefore have

$$x_m > \frac{\rho}{2}\left(1 + \sqrt{1 + \frac{4\tau}{\rho}}\right).$$

We note one interesting point regarding formula (69). It looks very similar to the corresponding result in the quasireal electron approximation [27] for the description of the collinear photon radiation ($\theta_\gamma < \theta_0$, $\theta_0 \ll 1$) by the longitudinally polarized electron, which is suitable for the leptonic variables,

$$d\sigma(k_1, k) = \frac{\alpha}{2\pi}\int(d\sigma_b(xk_1)P(x, L_0))dx, \quad k = (1 - x)k_1, \quad (70)$$
$$P(x, L_0) = \frac{1 + x^2}{1 - x}L_0 - \frac{2(1 - x + x^2)}{1 - x}, \quad L_0 = \ln\frac{\varepsilon_1^2\theta_0^2}{m^2},$$

where $d\sigma_b$ is the cross section of the radiationless process. It is not surprising that the function F^p differs from P in Eq. (70) because it also has to contain traces from the final electron radiation.

Formulas (63), (64), and (69) describe the distribution over the momentum transfer squared in reaction (49) and define the respective radiative correction due to the hard photon emission. To compute the total radiative correction, we must also add the contribution due to emission of the virtual photon and the real soft photon (with the energy less than $\Delta\varepsilon_1$). This contribution is the same for polarized and unpolarized scattering,

$$\frac{d\sigma^{S+V}}{dQ^2} = \frac{d\sigma_b}{dQ^2}\frac{\alpha}{2\pi}(\delta^V + \delta^S). \quad (71)$$

The virtual correction is standard [28],

$$\delta^V = 4(L_Q - 1)\ln\frac{\lambda}{m} - L_Q^2 + 3L_Q + \frac{\pi^2}{3} - 4, \quad L_Q = \ln\frac{Q^2}{m^2}, \quad (72)$$

where λ is the "photon mass," while the soft-photon correction has some specification in the hadronic variables

$$\delta^S = 4(L_Q - 1)\ln\frac{m\Delta}{\lambda} + 2L_Q^2 - 2L_Q - \frac{\pi^2}{3} + 2. \quad (73)$$

It can be seen that the terms proportional to L_Q^2 do not vanish in the sum $\delta^V + \delta^S$ (as they do for the leptonic variables) and the contribution of the hard photon emission has to be taken into account to cancel them (due to the terms with $\ln R_x/(1-x)$ in the functions F^p, F^T, and F^{un}).

The observed cross sections, which take the total radiative correction into account, do not depend on the auxiliary infrared parameter Δ and can be written in the form suitable for numerical integration as

$$\frac{d\sigma^{un}}{Q^2}(k_1) = \frac{d\sigma_b^{un}}{Q^2}(k_1)\left(1 + \frac{\alpha}{2\pi}\delta\right) + \frac{\alpha}{2\pi}\int_{x_m}^1\left[\frac{d\Delta\sigma_b^{un}}{dQ^2}\frac{f_\Delta}{1-x} + \frac{d\sigma_b^{un}}{dQ^2}(xk_1)\tilde{F}^{un} + \frac{\pi\alpha^2}{Q^4}\frac{W_2}{M^2}f(x,\rho,\tau)\right]dx, \quad (74)$$

$$\frac{d\sigma^{L,T}}{Q^2}(k_1) = \frac{d\sigma_b^{L,T}}{Q^2}(k_1)\left(1 + \frac{\alpha}{2\pi}\delta\right) + \frac{\alpha}{2\pi}\int_{x_m}^1\left[\frac{d\Delta\sigma_b^{L,T}}{dQ^2}\frac{f_\Delta}{1-x} + \frac{d\sigma_b^{L,T}}{dQ^2}(xk_1)\tilde{F}^{p}\right]dx. \quad (75)$$

The partial cross sections in the case of the tensor polarization of the recoil deuteron are defined by the formula

$$\frac{d\sigma^{IJ}}{Q^2}(k_1) = \frac{d\sigma_b^{IJ}}{Q^2}(k_1)\left(1 + \frac{\alpha}{2\pi}\delta\right) + \frac{\alpha}{2\pi}\int_{x_m}^1\left[\frac{d\Delta\sigma_b^{IJ}}{dQ^2}\frac{f_\Delta}{1-x} + \frac{d\sigma_b^{IJ}}{dQ^2}(xk_1)\tilde{F}^{T}\right]dx + \frac{\alpha^3}{V^2}(1+\eta)\ln x_m(G_M^2 - 2G_Q G)\delta_{IL}\delta_{JL}, \quad (76)$$

$$I, J = L, T,$$

where

$$d\Delta\sigma_b = d\sigma_b(xk_1) - d\sigma_b(k_1)$$

for both polarized and unpolarized cases and

$$\delta = \left(\frac{3}{2} + 2\ln(1-x_m)\right)L_Q - \frac{7}{2}\ln(1-x_m) - \ln^2(1-x_m) - \frac{5}{2} - \frac{\pi^2}{3},$$

$$f_\Delta = 2L_Q - 2\ln(1-x) - \frac{7}{2},$$

$$\tilde{F}^p = -(1+x)L_Q - \frac{1+x^2}{1-x}\ln x + (1+x)\ln(1-x) + 1 + 4x, \quad (77)$$

$$\tilde{F}^{un} = \tilde{F}^p + 3 - 4x,$$

$$\tilde{F}^T = \tilde{F}^p + 3 - x.$$

The singularity at $x = 1$ in the integrands of Eqs. (74)–(76) cancels by the corresponding quantity $d\Delta\sigma_b/dQ^2$. For example, in the simplest unpolarized case, we have

$$\frac{d\Delta\sigma_b^{un}}{dQ^2} = \frac{\pi\alpha^2}{Q^2V^2x^2} \times\left[2(1+x)W_1 - \frac{W_2}{\tau}(x + \tau(1+x))\right](1-x).$$

It is well known that the leading logarithmic contributions to the radiative correction of order $(\alpha L_Q)^n$, $n = 1, 2, \ldots$ are controlled by the electron structure function $D(x, L_Q)$,

$$\frac{d\sigma_{\text{lead}}}{dQ^2} = \int_{x_m}^1 D(x, L_Q)\frac{d\sigma_b}{dQ^2}(xk_1)dx, \quad (78)$$

$$D(x, L_Q) = \delta(1-x) + \frac{\alpha L_Q}{2\pi}P_1(x) + \frac{1}{2}\left(\frac{\alpha L_Q}{2\pi}\right)^2 P_2(x) + \ldots.$$

It can be verified that the leading part of the first-order correction defined by Eqs. (74)–(76) can be derived using representation (78) at

$$D(x, L_Q) = \frac{\alpha L_Q}{2\pi}P_1(x),$$

$$P_1(x) = \frac{1+x^2}{1-x}\Theta(1-x-\Delta) + \delta(1-x)\left(2\ln\Delta + \frac{3}{2}\right).$$

Thus, we can improve our result by insertion of the higher order leading contributions using Eq. (78) and the known expressions for the functions $P_n(x)$ [29]. This improvement results in modification of the quantities δ,

f_Δ, and $\tilde{F}^p$ in the right-hand sides of Eqs. (74)–(76). For example, to account for the corresponding second-order terms, we must use the substitutions

$$\delta \longrightarrow \delta + \gamma\delta_1, \quad f_\Delta \longrightarrow f_\Delta + \gamma f_{\Delta 1},$$

$$\tilde{F}^p \longrightarrow \tilde{F}^p + \gamma\tilde{F}_1^p, \quad \gamma = \frac{\alpha L_Q^2}{2\pi},$$

$$\delta_1 = \frac{9}{8} - \frac{\pi^2}{3} + 3\ln(1-x_m) + 2\ln^2(1-x_m), \qquad (79)$$

$$f_{\Delta 1} = 3 + 4\ln(1-x),$$

$$\tilde{F}_1^p = -2(1+x)\ln(1-x) - \frac{1+3x^2}{2(1-x)}\ln x - \frac{5+x}{2}.$$

On the other hand, there exists a simple method of summation of all the singularities at $x = 1$ in the total radiative correction, which goes beyond the leading logarithmic approximation (see, e.g., [30]). It consists in using the exponential form of the electron structure function and, in our case, can be introduced as

$$\delta(1-x) + \beta\left[\frac{1}{1-x}\theta(1-\Delta-x) + \delta(1-x)\ln\Delta\right] \longrightarrow \frac{\beta(1-x)^{\beta-1}}{\Gamma(1+\beta)}\exp\left[\beta\left(\frac{3}{4}-C\right)\right], \qquad (80)$$

where

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

C is the Euler constant, and $\Gamma(x)$ is the gamma function. This procedure leads to a redefinition of δ and f_Δ in Eqs. (74)–(76),

$$\delta \longrightarrow \delta^{\exp} = \frac{3}{2}(L_Q - \ln(1-x_m)) - \ln^2(1-x_m) - \frac{5}{2} - \frac{\pi^2}{3},$$

$$f_\Delta \longrightarrow f_\Delta^{\exp} = -\frac{3}{2} - 2\ln(1-x),$$

and to the appearance of an additional term

$$\int_{x_m}^1 \frac{d\sigma_b}{dQ^2}(xk_1)\beta(1-x)^{\beta-1}\frac{\exp\left[\beta\left(\frac{3}{4}-C\right)\right]}{\Gamma(1+\beta)}dx$$

that absorbs the purely Born cross section and a part of the radiative corrections. For example, partial cross sections (76) become

$$\frac{d\sigma^{IJ}}{Q^2}(k_1) = \int_{x_m}^1 \left[\frac{d\sigma_b^{IJ}}{dQ^2}(xk_1)\left(\frac{\beta(1-x)^{\beta-1}}{\Gamma(1+\beta)}\exp\left[\beta\left(\frac{3}{4}-C\right)\right] + \frac{\alpha}{2\pi}\tilde{F}^T\right) + \frac{\alpha}{2\pi}\frac{d\Delta\sigma_b^{IJ}}{dQ^2}\frac{f_\Delta^{\exp}}{1-x}\right]dx + \frac{\alpha}{2\pi}\delta^{\exp}\frac{d\sigma_b^{IJ}}{Q^2}(k_1) + \frac{\alpha^3}{V^2}(1+\eta)\ln x_m(G_M^2 - 2G_Q G)\delta_{IL}\delta_{JL}. \qquad (81)$$

4. NUMERICAL ESTIMATIONS

There are different approaches to the analysis of polarization observables. If the experimental information is extracted directly from the spin-dependent part of the cross section (see [31] for the corresponding experimental method), the radiative correction can be large due to the contribution of factored virtual and soft corrections. The nonfactored contribution to the radiative correction, caused by the hard photon emission, cannot be large in elastic scattering because the phase space of such a photon is strongly suppressed by restrictions on the event selection. The effect of the radiative correction in this case is demonstrated in Figs. 1–4 for the ratios

$$\delta_u = \frac{d\sigma_{\rm obs}^{un}}{d\sigma_b^{un}} - 1, \quad \delta_{L,T} = \frac{P_{\rm obs}^{L,T}}{P_b^{L,T}} - 1, \quad \delta_R = \frac{P_{\rm obs}^T P_b^L}{P_{\rm obs}^L P_b^T} - 1 \qquad (82)$$

in the unpolarized case and for vector polarization of the recoil deuteron, and for the ratios

$$\delta_{IJ} = \frac{d\sigma_{\rm obs}^{IJ} d\sigma_b^{un}}{d\sigma_{\rm obs}^{un} d\sigma_b^{IJ}} - 1, \quad \delta_{RQ} = \frac{d\sigma_{\rm obs}^{TT} d\sigma_b^{LT}}{d\sigma_{\rm obs}^{LT} d\sigma_b^{TT}} - 1 \qquad (83)$$

in the case of tensor polarization. We note that, if the radiative correction is ignored, all the quantities defined by Eqs. (82) and (83) are equal to zero. The quantities δ_R and δ_{RQ} are very important physical values because they can be used for an independent determination of the ratios of form factors such as G_M/G and G_M/G_Q (see Eqs. (19), (20), (31), (32), and (48)).

The observed cross sections in Eqs. (82) and (83) are defined by Eqs. (74)–(76) or their exponential modification (as in Eq. (81)). We consider two different

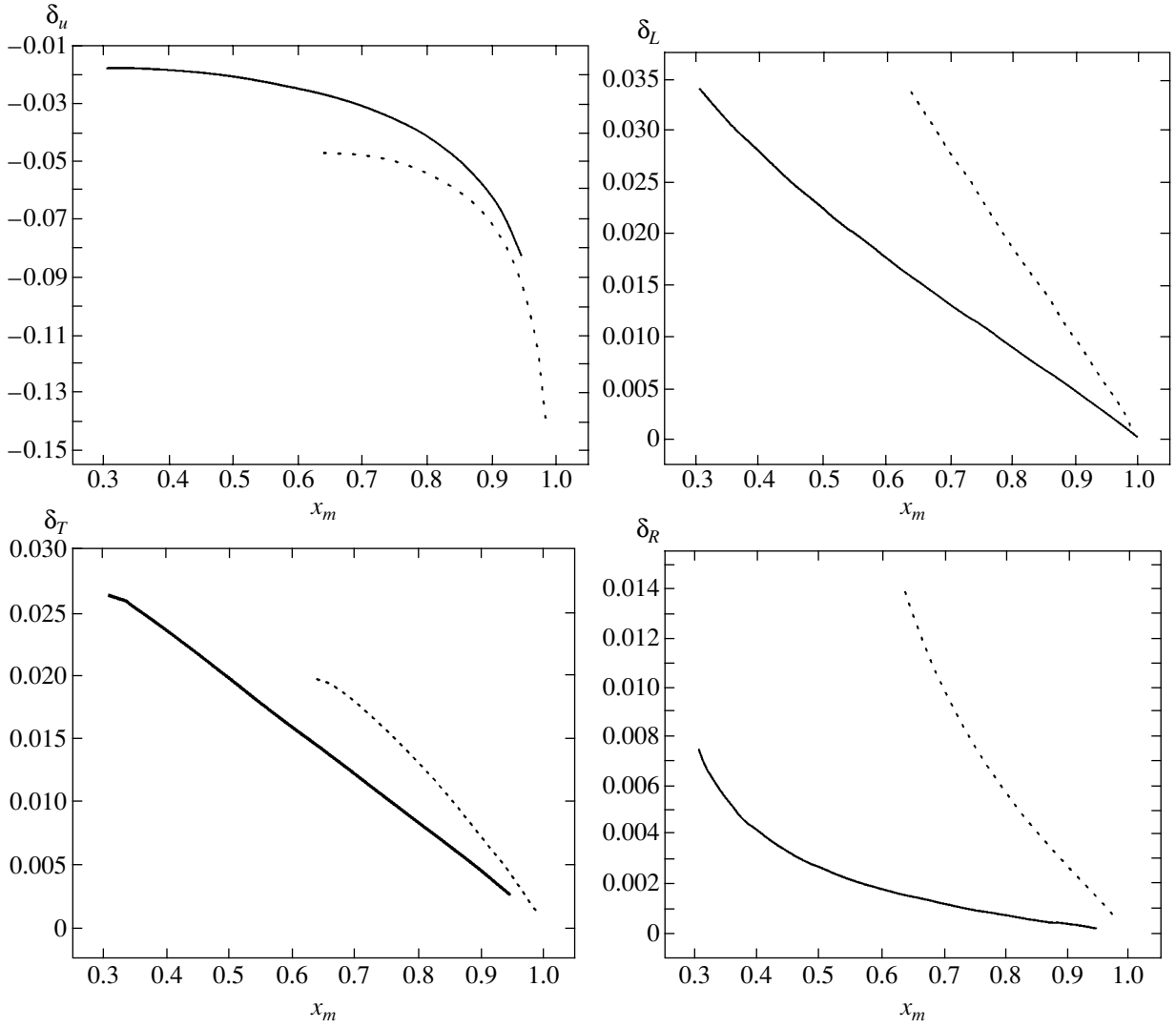


Fig. 1. The effect of the radiative correction on the unpolarized cross section and vector polarizations of the recoil deuteron given by Eqs. (74), (75) and Eq. (82) at $V = 8$ GeV^2. The solid curves correspond to $Q^2 = 1$ GeV^2, and the dashed ones, to $Q^2 = 3$ GeV^2. Parametrization I is used for the deuteron electromagnetic form factors [32].

parametrizations of the deuteron electromagnetic form factors given in [32] and label them I and II.

As can be seen from Fig. 1, the radiative corrections to the unpolarized cross section depends strongly on the value x_m that is connected with the energy of the hard photon in process (49). If x_m is close to unity ($x_m \approx 1$), the total radiative correction, being negative, can reach 10% or even more. As x_m decreases, the total radiative correction becomes much smaller. Such behavior of the radiative correction has a simple physical interpretation. If $x_m \approx 1$, the energy of the photon in process (49) is sufficiently small and the positive contribution into the radiative correction due to the hard photon emission cannot compensate the factored negative contribution caused by virtual and soft photon corrections that accompany process (1). As the hard-photon energy increases, such compensation occurs and the absolute value of the total radiative correction decreases. The same behavior is also exhibited by polarization-dependent parts of the cross section in the case of vector polarization of the recoil deuteron and by partial cross sections in the case of tensor polarization.

However, the effect of the radiative corrections is precisely the opposite for the ratios defined by Eqs. (82) and (83). At $x_m \approx 1$, the total radiative correction is defined mainly by its factored part, which is the same for the polarization-dependent and unpolarized cross sections. Therefore, the radiative correction in fact cancels in this region for such ratios. On the contrary, at smaller values of x_m, the nonfactored part of the radiative correction becomes significant and the total radiative correction increases. An unexpected fact is that the ratios δ_{IJ} in (83) are approximately one order smaller than $\delta_{L,T}$ in (82).

As our calculations show, the sensitivity of the radiative correction to two different parametrizations of the

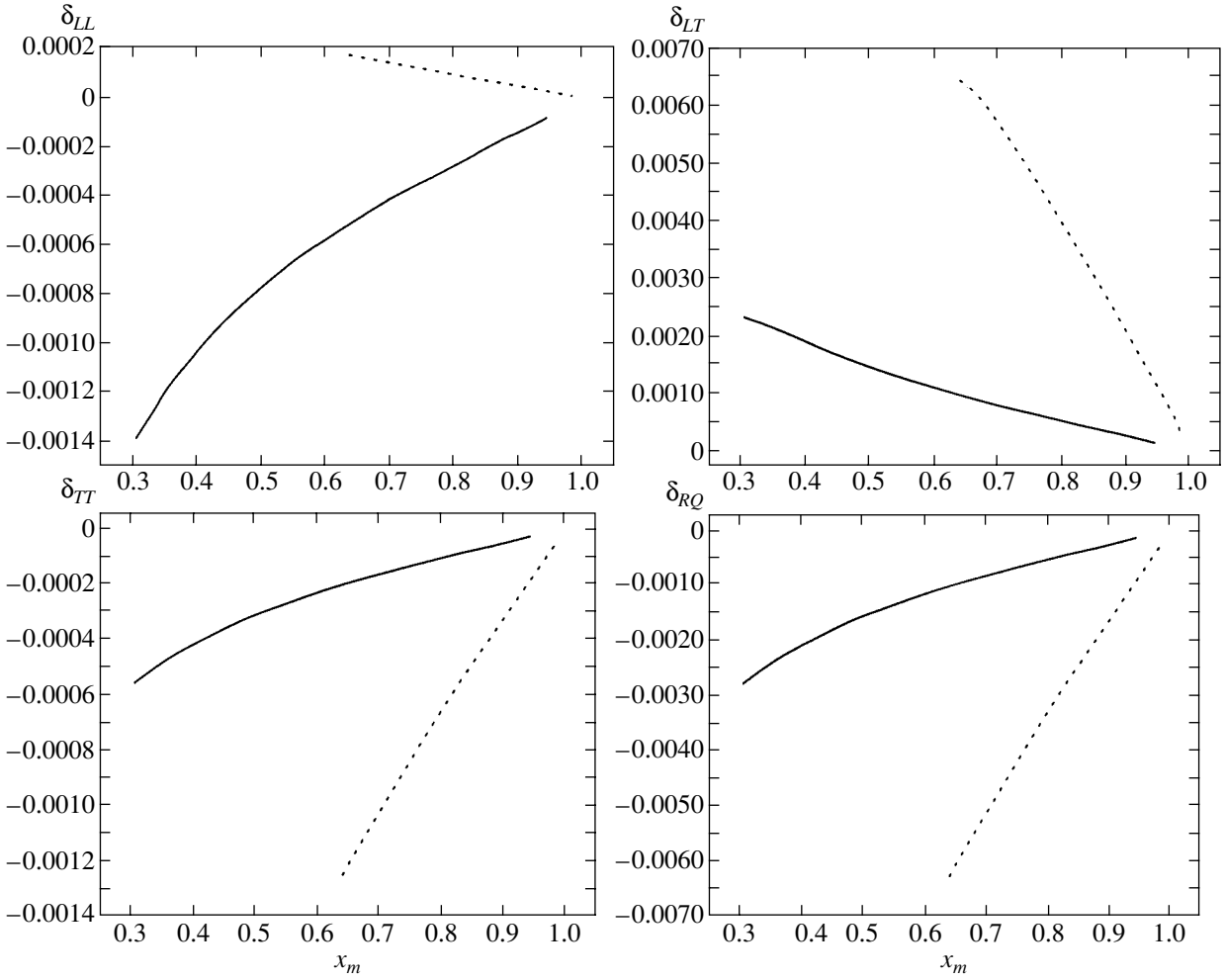


Fig. 2. The effect of the radiative correction in the case of tensor polarization of the recoil deuteron. The partial cross sections in Eq. (83) are calculated using Eq. (76). Kinematical conditions and parametrization of the form factors are the same as in Fig. 1.

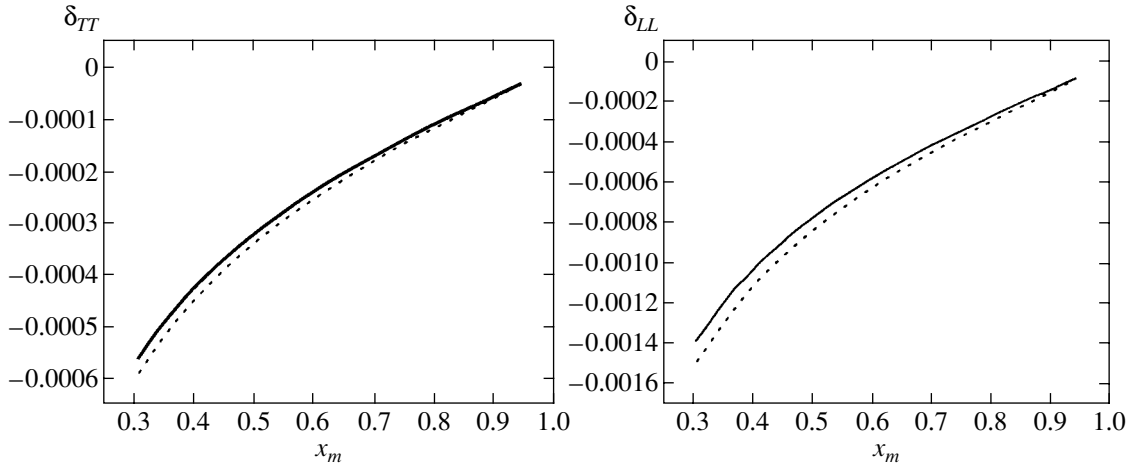


Fig. 3. The influence of different parametrizations of the deuteron electromagnetic form factors on the radiative correction. The solid (dashed) curve corresponds to parametrization I (II) [32]. The kinematical conditions are $V = 8$ GeV^2 and $Q^2 = 1$ GeV^2.

deuteron electromagnetic form factors [32] at relevant values of energies and momentum transfers is practically negligible. In fact, the respective curves coincide in the entire range of x_m (see Fig. 3).

The influence of the higher order corrections, calculated by summing the leading contributions by the exponentiation procedure, is demonstrated in Fig. 4 for the tensor polarization ratios. The corresponding curves

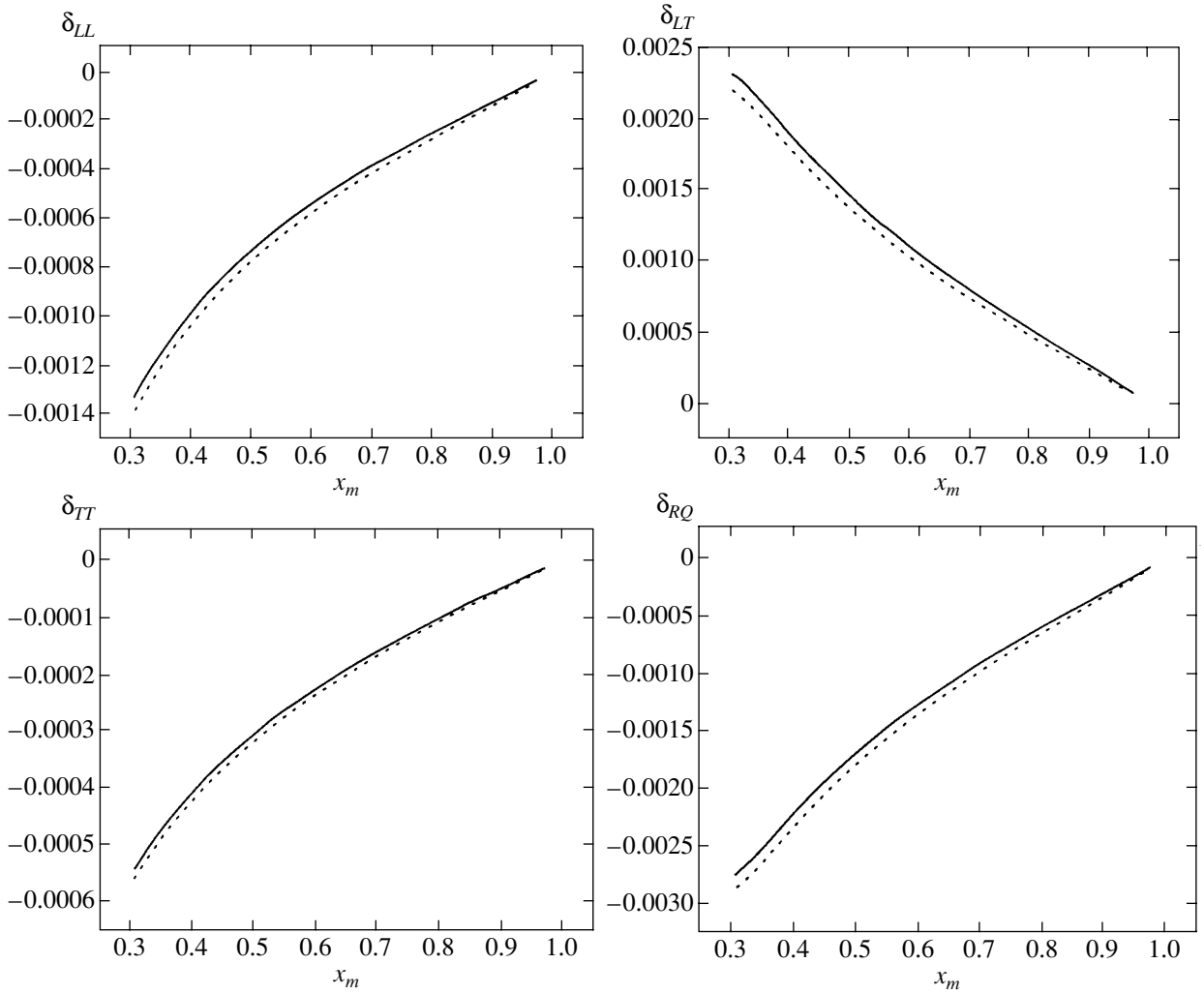


Fig. 4. Comparison of the total radiative corrections in the case of tensor polarization calculated by Eq. (76) (solid curve) and Eq. (81) (dashed curve) at $V = 8$ GeV^2 and $Q^2 = 1$ GeV^2.

for the vector ones are very similar. We see that the effect is small and cannot even exhibit itself at small x_m, where the nonfactored radiative correction contributes. As usual, the large correction factor caused by exponentiation of the higher order leading radiative corrections at $x_m \approx 1$ to the unpolarized and polarized parts of the cross section cancels in their ratios.

REFERENCES

1. M. Garcon and J. W. Van Orden, nucl-th/0102049; R. Gilman and F. Gross, J. Phys. G: Nucl. Part. Phys. **28**, R37 (2002).
2. I. Sick, nucl-ex/0208009.
3. D. M. Nikolenko, H. Arenhovel, L. M. Barkov, *et al.*, Phys. Rev. Lett. **90**, 072501 (2003).
4. G. I. Gakh and N. P. Merenkov, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 659 (2001) [JETP Lett. **73**, 579 (2001)].
5. A. V. Afanas'ev, I. V. Akushevich, G. I. Gakh, and N. P. Merenkov, Zh. Éksp. Teor. Fiz. **120**, 515 (2001) [JETP **93**, 449 (2001)].
6. D. Abbott, J. Ball, J. Ducret, *et al.*, nucl-ex/0001006; Phys. Rev. Lett. **84**, 5053 (2000).
7. R. Gilman, in *Physics and Instrumentation with 6–12 GeV Beams*, Ed. by S. Dytman, H. Fenker, and P. Roos (TJNAF, Newport News, Virginia, 1998).
8. I. V. Akushevich and N. M. Shumeiko, J. Phys. G **20**, 513 (1994).
9. I. Akushevich, A. Ilichev, N. Shumeiko, *et al.*, Comput. Phys. Commun. **104**, 201 (1997).
10. I. Akushevich, H. Boettcher, and D. Ryckbosch, hep-ph/9906408.
11. M. Gourdin and C. A. Piketty, Nuovo Cimento **32**, 1137 (1964).
12. M. Gourdin, Phys. Rep. C **11**, 29 (1974).
13. M. J. Moravcsik and P. Ghosh, Phys. Rev. Lett. **32**, 321 (1974).
14. I. Kobzarev, L. B. Okun', and M. V. Terent'ev, Pis'ma Zh. Éksp. Teor. Fiz. **2**, 466 (1965) [JETP Lett. **2**, 289

(1965)]; V. M. Dubovik, E. P. Likhtman, and A. A. Cheshkov, Zh. Éksp. Teor. Fiz. **52**, 706 (1967) [Sov. Phys. JETP **25**, 464 (1967)].

15. H. S. Song, F. L. Ridener, Jr., and R. H. Good, Jr., Phys. Rev. D **25**, 61 (1982).
16. R. G. Arnold, C. E. Carlson, and F. Gross, Phys. Rev. C **23**, 363 (1981).
17. A. I. Akhiezer and M. P. Rekalo, *Electrodynamics of Hadrons* (Naukova Dumka, Kiev, 1977), p. 216.
18. A. V. Afanasev, I. Akushevich, and N. P. Merenkov, Phys. Rev. D **65**, 013006 (2002); hep-ph/0009273.
19. I. V. Akushevich, A. V. Afanasev, and N. P. Merenkov, hep-ph/0111331.
20. B. D. Milbrath *et al.* (Bates FPP Collaboration), Phys. Rev. Lett. **80**, 452 (1998); Phys. Rev. Lett. **82**, 2221 (1999); M. K. Jones *et al.* (Jefferson Lab Hall A Collaboration), Phys. Rev. Lett. **84**, 1398 (2000).
21. A. Akhundov, D. Bardin, L. Kalinovskaya, and T. Riemann, Fortschr. Phys. **44**, 373 (1996); J. Blumlein, Phys. Lett. B **271**, 267 (1991); Z. Phys. C **65**, 293 (1995).
22. A. V. Afanasev, I. Akushevich, A. Ilyichev, and N. P. Merenkov, Phys. Lett. B **514**, 269 (2001).
23. D. Y. Bardin and N. M. Shumeiko, Nucl. Phys. B **127**, 242 (1977); D. Bardin and L. Kalinovskaya, hep-ph/9712310.
24. I. Akushevich, A. Ilyichev, and N. Shumeiko, Eur. Phys. J. C **5**, 1 (2001).
25. É. A. Kuraev, N. P. Merenkov, and V. S. Fadin, Yad. Fiz. **47**, 1593 (1988) [Sov. J. Nucl. Phys. **47**, 1009 (1988)]; Yad. Fiz. **45**, 782 (1987) [Sov. J. Nucl. Phys. **45**, 486 (1987)].
26. T. V. Kuchto and N. P. Shumeiko, Nucl. Phys. B **219**, 412 (1983).
27. V. N. Baier, V. S. Fadin, and V. A. Khoze, Nucl. Phys. B **65**, 381 (1973).
28. R. Barbieri, J. A. Mignaco, and E. Remiddi, Nuovo Cimento A **11**, 824 (1972); **11**, 865 (1972); P. Mastrolia and E. Remiddi, hep-ph/0302162.
29. S. Jadach, M. Skrzypek, and B. F. L. Ward, Phys. Rev. D **47**, 3733 (1993).
30. É. A. Kuraev and V. S. Fadin, Yad. Fiz. **41**, 733 (1985) [Sov. J. Nucl. Phys. **41**, 466 (1985)]; F. A. Berends, W. L. van Neerven, and G. J. H. Burgers, Nucl. Phys. B **297**, 429 (1988).
31. N. D. Gagunashvili, Nucl. Instrum. Methods Phys. Res. A **343**, 606 (1994).
32. www-dapnia.cea.fr/SphnT20

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 870–881.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 999–1011.

ATOMS, SPECTRA, RADIATION

Wedge Refraction of Electromagnetic Waves in Absorbing Crystals

V. I. Alshits and V. N. Lyubimov

Shubnikov Institute of Crystallography, Russian Academy of Sciences, Leninskiĭ pr. 59, Moscow, 117333 Russia

e-mail: alshits@ns.crys.ras.ru

Received March 14, 2003

Abstract—Topological features of the self-intersection of wave surfaces near singular optical axes of an absorbing crystal are investigated. Distributions of complex polarization fields in the neighborhood of singular directions are obtained. It is shown that, when the wave normal **m** circumvents an optical axis, the corresponding rotation of polarization ellipses is characterized by the Poincaré index $n = 1/4$. Using the example of an orthorhombic crystal, a wedge refraction of electromagnetic waves on the intersection line of the sheets of the surface of refractive indices is predicted and theoretically investigated. It is shown that the directions of the mean energy fluxes $\bar{\mathbf{P}}_\pm$ are close to the direction of normals $\mathbf{n}_\pm$ to the refraction surface only in the central region of a wedge, i.e., only in the domain where the polarization is almost linear and the group velocity of waves is well defined. When **m** moves to singular axes, the ellipticity of the polarization increases at the ends of the edge of the wedge and the orientations of the vectors $\bar{\mathbf{P}}_\pm$ and $\mathbf{n}_\pm$ gradually diverge, yet remain in the same plane that is orthogonal to the edge. The angle between $\bar{\mathbf{P}}_+$ and $\bar{\mathbf{P}}_-$ monotonically decreases, and $\bar{\mathbf{P}}_+ \parallel \bar{\mathbf{P}}_-$ for the propagation along singular axes; in this case, the angle between $\mathbf{n}_+$ and $\mathbf{n}_-$ increases, and they have a plane-fan-type orientational singularity along the optical axes. When **m i**s scanned along the edge of the wedge, the unaveraged vectors $\mathbf{P}_\pm$ describe per period the same conical surface that coincides with the refraction cone of a transparent crystal, while the endpoints of the vectors $\mathbf{P}_\pm$ run over elliptic orbits whose shape and slope depend on **m**. The possibilities of observing a wedge refraction are analyzed. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The conical refraction of electromagnetic waves in transparent crystals is a well-known phenomenon [1–3]. It arises in biaxial crystals when waves propagate along any of the two existing optical axes; henceforth, the directions of these axes will be denoted by $\mathbf{m}^{(1)}$ and $\mathbf{m}^{(2)}$. When the wave normal **m** coincides with the directions $\mathbf{m}^{(1,2)}$, the phase velocities $\mathbf{v} = v\mathbf{m}$ of the waves of independent polarizations degenerate (here, $v = c/n$, where n is the refractive index and c is the speed of light in vacuum). In this case, allowed orientations of polarization vectors form a whole continuum along the optical axis. Each polarization corresponds to its own energy-flux vector **P**, the Poynting vector, oriented along an appropriate generator of the refraction cone. Therefore, for example, a circularly polarized wave directed along $\mathbf{m}^{(j)}$ ($j = 1, 2$) should distribute energy over the cone. It is well known that the vector **P** in transparent crystals is collinear to the group velocity $\mathbf{u} = u\mathbf{n}$. Here, **n** is a unit normal to the surface of refractive indices $n(\mathbf{m})$. This surface has a conical singularity at the point $\mathbf{m} = \mathbf{m}^{(j)}$. The normals **n** defined in the immediate vicinity of the direction $\mathbf{m}^{(j)}$ coincide with the directions **P** and form a classical refraction cone in the limit; geometrically, this cone is exactly the same as the cone formed by the vectors **P**. Such cones were experimentally observed as early as the 19th century [2].

A similar phenomenon exists in crystal acoustics, where the conical refraction was experimentally and theoretically studied much later [2]. Interestingly, there is another type of refraction in crystal acoustics, the so-called wedge refraction, which was theoretically predicted in [4] and experimentally observed in [5]. This phenomenon is observed in hexagonal crystals in which two slowness surfaces of acoustic waves intersect. Near the intersection lines of the sheets of the slowness surface, the normals to these surfaces form a specific wedge of directions of group velocities.

In contrast to acoustics, where the wedge refraction occurs even in the absence of absorption, the wedge refraction in the optics of transparent crystals is impossible. However, in absorbing crystals, the situation changes. When the absorption is "switched on," each optical axis of a transparent biaxial crystal splits into two axes, $\mathbf{m}^{(j)} \longrightarrow \mathbf{m}_\pm^{(j)}$, which are called singular axes [1, 6]. The crystal becomes optically tetraaxial. The sheets of the double-sheeted surface of refractive indices intersect along the lines that connect the directions of split axes. The surface of refractive indices has wedgelike singularities at which the intersection line of

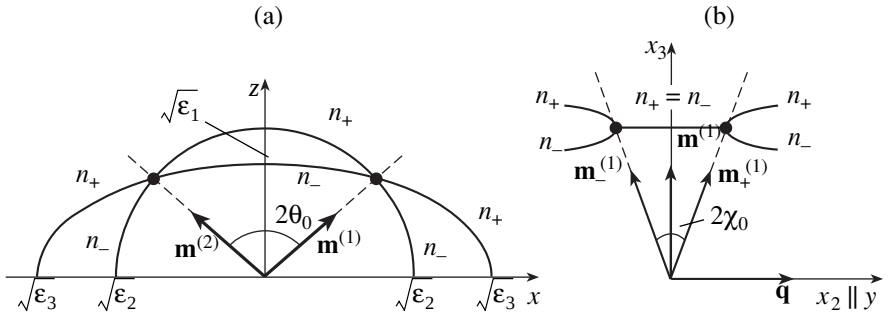


Fig. 1. Sections of the surfaces $n_\pm(\mathbf{m})$ by the planes containing optical axes; (a) transparent crystal and (b) absorbing crystal.

the sheets represents the edge of the wedge. It can naturally be expected that the switching on of absorption modifies the classical conical refraction of a transparent crystal, turning it into a qualitatively new phenomenon, a wedge refraction.

The present paper is devoted to the analysis of this very problem. We will thoroughly investigate the complex geometry of the self-intersection of the surface of refractive indices and analyze the distribution of energy fluxes of eigenwaves in the self-intersection region. We will study the behavior of the Poynting vectors on the intersection line of sheets and the behavior of geometric normals to the surface of refractive indices, which determine the directions of group velocities in transparent crystals. We will see that these two vector characteristics, $\mathbf{P}(\mathbf{m})$ and $\mathbf{n}(\mathbf{m})$, exhibit radically different types of behavior near the singular axes. This fact is associated with the strong ellipticity of the wave fields in this region and makes the introduction of such a characteristic of wave fields as the group velocity impossible [7]. On the other hand, there always exists a line between singular axes on the unit sphere of directions $\mathbf{m} \cdot \mathbf{m} = 1$ on which the ellipticity is equal to zero and the group velocity is rigorously defined. We will see below that the directions of the energy-flux $\mathbf{P}$ and the normal $\mathbf{n}$ naturally coincide on this line, and the character of wedge refraction near this line is similar to the related phenomenon in the acoustics of nonabsorbing crystals. When moving away from this region along the edge of the wedge, the ellipticity of polarization increases, while the energy flux $\mathbf{P}(\mathbf{m})$ rotates along different sections of the universal cone, which coincides with the refraction cone of a transparent crystal. At the ends of the wedge, these rotations occur in opposite directions along circular sections of the same cone. Thus, the behavior of the energy flux along singular axes actually coincides with the conical refraction in a transparent crystal for a circularly polarized wave. Below, we will show that these absorption-induced features, which are characterized by a smooth transition from one refraction cone to another at the ends of the wedge through a narrow area of pure wedge refraction, permit experimental observation even in the case of small absorption.

Analysis of topological anomalies in the distribution of complex vector fields of polarization in the neighborhood of the singular axes presents a separate nontrivial problem. According to [8], these features can be characterized by "topological charge," the Poincaré index $n = 1/4$. This important result, a brief account of which was earlier published by one of the authors of this paper in a nearly inaccessible publication [8], has unfortunately been overlooked by specialists in the field. Therefore, we also touch upon this problem below.

2. STATEMENT OF THE PROBLEM AND THE BASIC EQUATIONS

First, we consider a transparent triclinic crystal characterized by a real permittivity tensor $\boldsymbol{\varepsilon}$ or by its inverse $\boldsymbol{\eta} = \boldsymbol{\varepsilon}^{-1}$. If we assume, for definiteness, that $\varepsilon_1 < \varepsilon_2 < \varepsilon_3$, then, in a standard crystallophysical system of coordinates in which the tenors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ are diagonal, the optical axes will lie in the plane xz. In each singular direction $\mathbf{m}^{(j)}$, the sheets of the surface of refractive indices $n_\pm(\mathbf{m})$ have a conical contact point at $\mathbf{m}^{(1,2)}$: $n_+(\mathbf{m}^{(j)}) = n_-(\mathbf{m}^{(j)})$. Here, the subscripts $(\pm)$ denote two different sheets of the surface of refractive indices, the external (+) and the internal (–) sheets, which correspond to independent isonormal electromagnetic waves. The orientations of the directions $\mathbf{m}^{(j)}$ are defined by the angle θ_0 (Fig. 1a), which is easily determined when one considers the cross section of the surface of refractive indices by the xz plane:

$$\tan\theta_0 = \sqrt{d_{12}/d_{23}}, \quad d_{ij} = \eta_i - \eta_j. \tag{1}$$

Thus, in the chosen crystallophysical system of coordinates $\{x, y, z\}$, the orientations of optical axes are defined by the directions

$$\mathbf{m}^{(1,2)} = (\pm\sin\theta_0, \cos\theta_0). \tag{2}$$

When the absorption is switched on, the material tensors $\boldsymbol{\varepsilon}$ and $\boldsymbol{\eta}$ are formally complemented by imaginary components: $\boldsymbol{\varepsilon} \longrightarrow \boldsymbol{\varepsilon} + i\boldsymbol{\varepsilon}'$ and $\boldsymbol{\eta} \longrightarrow \boldsymbol{\eta} - i\boldsymbol{\eta}'$. Accordingly, it is convenient to combine refractive indices $n_\pm$ and absorption indices $n'_\pm$ by introducing

complex refractive indices $N_\pm = n_\pm + in'_\pm$. In terms of these indices, isonormal electromagnetic waves of frequency ω are expressed in a form similar to the case of a transparent crystal:

$$\begin{pmatrix} \mathbf{E}_\pm(\mathbf{r}, t) \\ \mathbf{H}_\pm(\mathbf{r}, t) \end{pmatrix} = \begin{pmatrix} \mathbf{E}^0_\pm \\ \mathbf{H}^0_\pm \end{pmatrix} \exp\left[i\omega\left(\frac{N_\pm}{c}\mathbf{m}\cdot\mathbf{r} - t \right)\right]. \quad (3)$$

The relation between the wave fields (3) and the complex refractive indices $N_\pm$, on the one hand, and the direction $\mathbf{m}$ and the material constants, on the other, is well known [1–3, 6]. As applied to our problem, it is convenient to make use of the Fedorov invariant equations [6], which explicitly depend on $\mathbf{m}$:

$$\begin{aligned} \mathbf{E}_\pm &= (\boldsymbol{\eta} - i\boldsymbol{\eta}')[\mathbf{H}_\pm \times \mathbf{m}], \\ \mathbf{H}_{(\pm)} &\parallel \mathbf{p}_1\sqrt{\mathbf{p}_2\cdot\mathbf{p}_2} \pm \mathbf{p}_2\sqrt{\mathbf{p}_1\cdot\mathbf{p}_1}, \end{aligned} \quad (4)$$

$$N_{(\pm)} = \{a + b[\mathbf{p}_1\cdot\mathbf{p}_2 \pm \sqrt{(\mathbf{p}_1\cdot\mathbf{p}_1)(\mathbf{p}_2\cdot\mathbf{p}_2)}]\}^{-1/2}, \quad (5)$$

where $\mathbf{p}_i = \mathbf{m}\times\mathbf{c}^{(i)}$ and the parameters a, b, and $\mathbf{c}^{(i)}$ define a dyadic form of the complex tensor

$$\boldsymbol{\eta} - i\boldsymbol{\eta}' = a\mathbf{I} + b(\mathbf{c}^{(1)}\otimes\mathbf{c}^{(2)} + \mathbf{c}^{(2)}\otimes\mathbf{c}^{(1)}). \quad (6)$$

Note that the transition from $\mathbf{H}_{(\pm)}$ to $\mathbf{H}_\pm$ and from $N_{(\pm)} = n_{(\pm)} + in'_{(\pm)}$ to $N_\pm = n_\pm + in'_\pm$ in (4), (5) requires additional sorting of sheets into external and internal ones. In (6), $\mathbf{I}$ is the identity tensor. A specific relation between the parameters a, b, and $\mathbf{c}^{(i)}$ entering in (6) and the components of the usual matrix form of the tensor $\boldsymbol{\eta} - i\boldsymbol{\eta}'$ in the crystallophysical system of coordinates was considered in [9]. Here, we present these parameters in the explicit form only for a particular case of crystals of orthorhombic symmetry, which guarantees the diagonalization of the real and imaginary parts of the tensors $\boldsymbol{\varepsilon} + i\boldsymbol{\varepsilon}'$ and $\boldsymbol{\eta} - i\boldsymbol{\eta}'$ in the same system of coordinates. In the latter case,

$$\begin{aligned} a &= \eta_2 - i\eta'_2, \quad b = -d_{13} + id'_{13}, \\ \mathbf{c}^{(1)} &= (c_1, 0, c_3), \quad \mathbf{c}^{(2)} = (-c_1, 0, c_3), \end{aligned} \quad (7)$$

where, by analogy with (1), we denoted $d'_{ij} = \eta'_i - \eta'_j$ and introduced the parameters

$$c_1 = \sqrt{\frac{d_{12} - id'_{12}}{d_{13} - id'_{13}}}, \quad c_3 = \sqrt{\frac{d_{23} - id'_{23}}{d_{13} - id'_{13}}}. \quad (8)$$

The condition under which the complex refractive indices coincide, $N_+ = N_- = N_0$, corresponds to the directions of optical axes and is satisfied when $\mathbf{p}_1\cdot\mathbf{p}_1 = 0$ and $\mathbf{p}_2\cdot\mathbf{p}_2 = 0$. Each of these two complex equations determines a pair of singular directions, so that, in general, there exist four optical (singular) axes in absorbing crystals [1, 6].

Here, we should make one important remark. Although no special restrictions are imposed on the absorption level in the above formulas, in the class of problems under consideration, we will focus on a sufficiently low level of absorption such that an electromagnetic wave preserves its wave structure as it propagates over distances much greater than the wavelength. We are dealing with the cases when each optical axis of a transparent biaxial crystal splits, under the influence of weak anisotropic absorption, into two new axes: $\mathbf{m}^{(j)} \longrightarrow \mathbf{m}^{(j)}_\pm = \mathbf{m}^{(j)} + \Delta\mathbf{m}^{(j)}_\pm$. Of course, under weak absorption, such splitting must be small.[1] For example, for an orthorhombic crystal, the equations $\mathbf{p}_1\cdot\mathbf{p}_1 = 0$ and $\mathbf{p}_2\cdot\mathbf{p}_2 = 0$ with regard to (7) and (8) yield $\mathbf{m}^{(j)} \parallel \mathrm{Re}\,\mathbf{c}^{(j)}$, which, as can easily be verified, corresponds to (2), and

$$\begin{aligned} \Delta\mathbf{m}^{(j)}_\pm &\equiv \Delta\mathbf{m}_\pm = (0, \pm\chi_0, 0), \\ \chi_0 &= \frac{\sqrt{d_{12}d_{23}}}{2d_{13}}\left|\frac{d'_{12}}{d_{12}} - \frac{d'_{23}}{d_{23}}\right|. \end{aligned} \quad (9)$$

This symmetric splitting, shown in Fig. 1b, is quite natural because the original axes $\mathbf{m}^{(1,2)}$ lie in the xz plane, which is a symmetry plane.

According to (4), the complex polarization vectors of isonormal waves of magnetic field are orthogonal: $\mathbf{H}_+\cdot\mathbf{H}_- = 0$. This means that the corresponding polarization ellipses are orthogonal and the directions of their circumvention are identical. The eccentricities of isonormal ellipses are also identical. Let $a_\pm$ and $b_\pm$ be the lengths of the semiaxes of the polarization ellipses of isonormal waves. The orientations of these semiaxes are given by the formula

$$\mathbf{a}_\pm + i\mathbf{b}_\pm \parallel \mathbf{H}_\pm/\sqrt{\mathbf{H}_\pm\cdot\mathbf{H}_\pm}. \quad (10)$$

When a wave propagates along singular axes, when $\mathbf{p}_1\cdot\mathbf{p}_1 = 0$ or $\mathbf{p}_2\cdot\mathbf{p}_2 = 0$ and, accordingly, the fields $\mathbf{H}_\pm$ are parallel to $\mathbf{p}_1$ and $\mathbf{p}_2$ (up to a complex scalar factor, see (4)), the eccentricity of the polarization ellipses of the wave of magnetic field vanishes, i.e., the polarization is circular. Therefore, singular axes are sometimes called circular axes. However, according to (4), the electric component of the wave in this cases remains elliptic.

3. WAVE CHARACTERISTICS IN THE NEIGHBORHOOD OF A SINGULAR AXIS

Consider in greater detail the wave characteristics near a singular optical axis in a crystal with arbitrary anisotropy. Choose any of the four axes $\mathbf{m}^{(1,2)}_\pm$ and denote its direction by $\mathbf{m}_0$ for short. Assume, for definiteness, that $\mathbf{m}_0$ is a solution to the equation $\mathbf{p}_1\cdot\mathbf{p}_1 =$

[1] However, as is shown in [10, 11], under a small anisotropy comparable to the absorption level, the splitting is determined by the ratio of small parameters and may not be small. Below, we will not consider such special cases.

0. It can easily be shown that the vector $\mathbf{c}^{(1)}$ in this case can be represented as $\mathbf{c}^{(1)} = \mathbf{J} + g\mathbf{m}_0$, where $\mathbf{J}$ is a circular vector orthogonal to $\mathbf{m}_0$ ($\mathbf{J} \cdot \mathbf{J} = 0$ and $\mathbf{J} \cdot \mathbf{m}_0 = 0$). In a local system of coordinates $\{x_1, x_2, x_3\}$ with the axis x_3 directed along $\mathbf{m}_0$, the vector $\mathbf{J}$ has the following components: $\mathbf{J} = (1, i, 0)$. By an appropriate choice of the parameter b in (6), one can reduce the vector $\mathbf{c}^{(2)}$ to the form $\mathbf{c}^{(2)} = (1, \tilde{c}_2, \tilde{c}_3)$. Since the vectors $\mathbf{c}^{(1)}$ and $\mathbf{c}^{(2)}$ are fully determined by the material tensor $\boldsymbol{\eta} - i\boldsymbol{\eta}'$ (6), the parameters g, $\tilde{c}_2$, and $\tilde{c}_3$ can be considered known. In this notation, the complex refractive index for the singular direction $\mathbf{m}_0$ is given by

$$N_0 = n_0 + in_0' = [a + b(1 + i\tilde{c}_2)]^{-1/2}. \quad (11)$$

Our primary interest, however, lies in the behavior of the wave characteristics in the neighborhood of a singular direction where

$$\mathbf{m} = \mathbf{m}_0 + \theta(\cos\varphi, \sin\varphi, 0), \quad 0 \le \theta \ll 1. \quad (12)$$

Here, θ is the angle of deviation of the wave normal from the singular axis, and the polar angle φ is measured from the axis x_1. Calculations based on general expression (5) yield the following expressions:

$$n_{(\pm)} = n_0 \pm \sqrt{\theta}|P|\cos\frac{\varphi - \varphi_0}{2},$$
$$n'_{(\pm)} = n'_0 \pm \sqrt{\theta}|P|\sin\frac{\varphi - \varphi_0}{2}, \quad (13)$$

where $|P|$ and φ_0 are given by

$$P \equiv |P|\exp\left(-\frac{i\varphi_0}{2}\right) = ib\left(g\frac{1 + \tilde{c}_2^2}{2}\right)^{1/2} N_0^3. \quad (14)$$

In a similar way, applying (4), we obtain the following expressions for the complex amplitudes of the magnetic component of the wave:

$$\mathbf{H}^0_{(\pm)} = C_\pm[\mathbf{J} \pm \sqrt{\theta}\mathbf{R}\exp(i\varphi/2)], \quad (15)$$

$$\mathbf{R} = A(-\tilde{c}_2, 1, \theta), \quad A = \sqrt{2g/(1 + \tilde{c}_2^2)}. \quad (16)$$

In (15), $C_\pm$ are amplitude coefficients. As is clear from (15), when the propagation direction deviates from the singular axis ($\theta \neq 0$), the wave polarization ceases to be circular; ellipses arise whose eccentricities are the smaller, the smaller θ is,

$$e^2 = 1 - b_\pm^2/a_\pm^2 = \sqrt{\theta}|Q|, \quad (17)$$

and with the orientations of the semiaxes given by (10),

$$\mathbf{a}_{(+)} \parallel \left(\cos\frac{1}{4}(\varphi - \varphi^0), \sin\frac{1}{4}(\varphi - \varphi^0), 0\right),$$
$$\mathbf{a}_{(-)} \parallel \left(-\sin\frac{1}{4}(\varphi - \varphi^0), \cos\frac{1}{4}(\varphi - \varphi^0), 0\right). \quad (18)$$

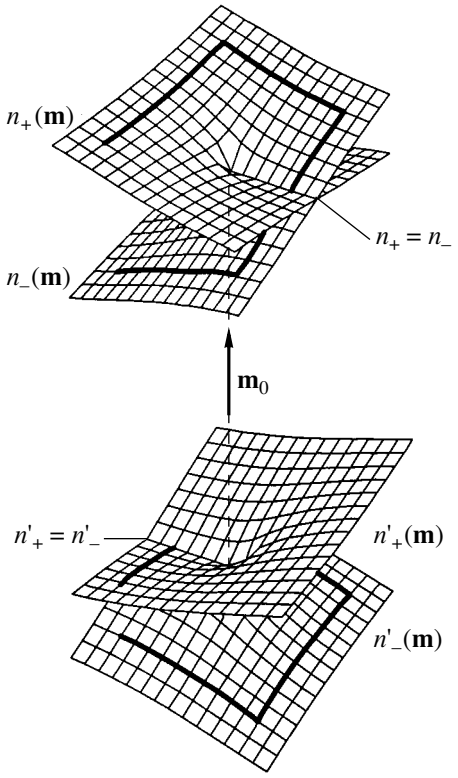


Fig. 2. Configuration of the surfaces $n_\pm(\mathbf{m})$ and $n'_\pm(\mathbf{m})$ in the neighborhood of a singular axis. The heavy lines show a continuous transition between external and internal sheets, $n_+ + in'_+ \longleftrightarrow n_- + in'_-$, under a complete circumvention of a singular axis.

In these formulas, $|Q|$ and φ^0 are defined by

$$Q \equiv |Q|\exp(-i\varphi^0/2) = 2A(i - \tilde{c}_2). \quad (19)$$

Formulas (13) completely describe the wave surfaces $n(\mathbf{m})$ and $n'(\mathbf{m})$ in the neighborhood of a singular axis. Locally, these surfaces are completely similar and differ only in the scale and the rotation through angle π around $\mathbf{m}_0$. Indeed, the difference $n_{(\pm)} - n_0$ coincides with $n'_{(\pm)} - n'_0$ when φ is replaced by $\varphi + \pi$. Each of the surfaces $n(\mathbf{m})$ and $n'(\mathbf{m})$ has a self-intersection line that emanates from a singular axis. The refractive indices degenerate on the line $\varphi = \varphi_0 + \pi$ ($n_+ = n_-$), while the absorption indices degenerate on the line $\varphi = \varphi_0$ ($n'_+ = n'_-$) (see Fig. 2). On the other hand, a full rotation around $\mathbf{m}_0$, i.e., the replacement of φ by $\varphi + 2\pi$, reverses the signs $\pm$ in (13), which means a simultaneous change of the branches: $n_+ + in'_+ \longleftrightarrow n_- + in'_-$ (see the relevant contours in Fig. 2). In this case, the physical equivalence of the position rotated through 2π is guaranteed by the fact that, according to (18), the polarization ellipses of isonormal waves also change places, $\mathbf{a}_+ + i\mathbf{b}_+ \longleftrightarrow \mathbf{a}_- + i\mathbf{b}_-$, so that the mutually orthogonal pairs of ellipses coincide after the rotation.

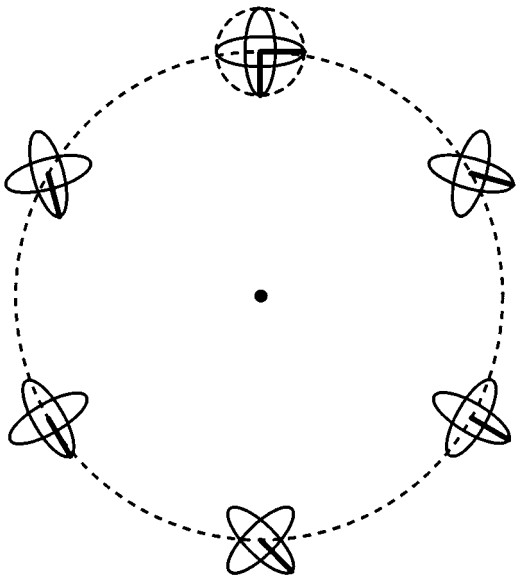


Fig. 3. Rotation of polarization ellipses under a complete circumvention of a singular axis; the small upper circle indicates the beginning and the end of circumvention.

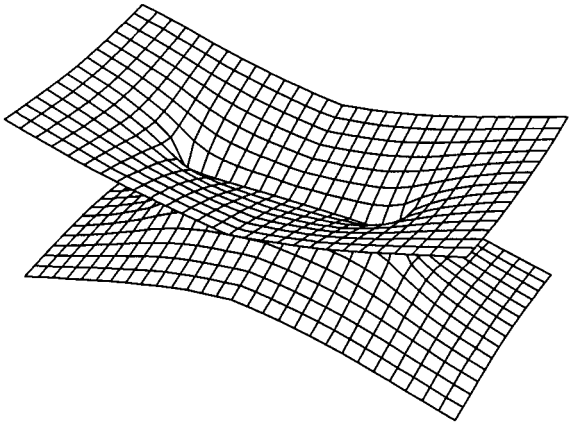


Fig. 4. Self-intersection of the surface $n_{\pm}(\mathbf{m})$.

As a result, after a full rotation around $\mathbf{m}_0$, the above pairs rotate in the same direction through a resulting angle of $\pi/2$, which corresponds to the Poincaré index 1/4 mentioned in the Introduction (see Fig. 3).

In [12], we showed that a similar topological scenario takes place in the description of the propagation of elastic waves in absorbing crystals. However, in this case, the Poincaré index of the appropriate distributions of polarization fields may be of either sign: $n = \pm 1/4$.

4. GEOMETRY OF THE SURFACE OF REFRACTIVE INDICES IN ORTHORHOMBIC CRYSTALS

Now, we proceed to the description of wedge refraction in the neighborhood of the self-intersection line of the surface of refractive indices. The general configuration of the surface $n_{\pm}(\mathbf{m})$ in this neighborhood follows from the previous description and is shown in Fig. 4. However, further analysis requires more detailed information about the geometry of this surface. This analysis is beyond the applicability of formula (13), which is only valid near one of the split axes. However, for small absorption and, accordingly, small splitting, one can describe at once the whole neighborhood of the pair of axes $\mathbf{m}_{\pm}^{(j)}$. Below, we do so for the neighborhood of the pair $\mathbf{m}_{\pm}^{(1)}$ in an orthorhombic crystal.

Arrange the Cartesian system of coordinates $\{x_1, x_2, x_3\}$ introduced above so that $x_3 \parallel \mathbf{m}^{(1)}$ and assume that the x_2 axis is parallel to the y axis of the original crystallophysical system of coordinates, as is shown in Fig. 1b. In other words, we choose the vectors $\mathbf{p}$, $\mathbf{q}$, and $\mathbf{m}^{(1)}$ as unit vectors of the system $\{x_1, x_2, x_3\}$; in the original system of coordinates, these vectors are given by

$$\begin{aligned} \mathbf{p} &= (\cos\theta_0, 0, -\sin\theta_0), \\ \mathbf{q} &= (0, 1, 0), \\ \mathbf{m}^{(1)} &= (\sin\theta_0, 0, \cos\theta_0). \end{aligned} \tag{20}$$

It is convenient to represent the wave normal $\mathbf{m}$ in the new system of coordinates as

$$\begin{aligned} \mathbf{m} &= \mathbf{m}^{(1)} + \Delta\mathbf{m}, \quad \Delta\mathbf{m} = (m_1, m_2, 0), \\ & m_1^2, m_2^2 \ll 1. \end{aligned} \tag{21}$$

Combining (5) and (7) and taking into account that the absorption is weak, after rather tedious calculations we obtain

$$\begin{aligned} n_{(\pm)} &= n_0 \pm D\sqrt{\varepsilon_2}\,\mathrm{Re}\,r, \quad n_0 = \sqrt{\varepsilon_2}(1 + Dm_1), \\ n'_{(\pm)} &= n'_0 \pm (\mathrm{Im}\,r)D\sqrt{\varepsilon_2}, \\ n'_0 &= (\eta'_2/2\eta_2 + \chi_1 D)\sqrt{\varepsilon_2}. \end{aligned} \tag{22}$$

Here,

$$\begin{aligned} D &= \sqrt{d_{12}d_{23}}/2\eta_2, \\ r &= \sqrt{m_1^2 + m_2^2 - \chi_1^2 + 2i\chi_1 m_1}. \end{aligned} \tag{23}$$

In (22) and (23), χ_1 is defined by (9) in which the sign of modulus should be changed by mere brackets. The condition $\mathrm{Re}\,r = 0$, which reduces to the relations

$$m_1 = 0, \quad -\chi_0 \le m_2 \le \chi_0, \tag{24}$$

determines the self-intersection line of the surface $n_{\pm}(\mathbf{m})$ on which $n_+ = n_-$ (Fig. 4). On the other hand, the condition $\mathrm{Im}\,r = 0$ yields the two relations

$$m_1 = 0, \quad m_2 \le -\chi_0 \text{ and } m_1 = 0, \quad m_2 \ge \chi_0, \tag{25}$$

which determine two self-intersection lines of the sur-

face $n'_{\pm}(\mathbf{m})$ such that $n'_{+} = n'_{-}$. On the unit sphere of directions $\mathbf{m}\cdot\mathbf{m} = 1$, the lines (25) continue the line (24). At the endpoints of the segment (24) (at the points $m_2 = \pm\chi_0$), both the refractive indices and the absorption indices coincide. These points define the directions of singular axes. The absorption splits the optical axis $\mathbf{m}^{(1)}$ of a transparent crystal, thus giving rise to a pair of singular axes $\mathbf{m}^{(1)}_{+}$ and $\mathbf{m}^{(1)}_{-}$ that make an angle of $2\chi_0$ (in radians) (see (9) and Fig. 1b). The system of coordinates used in this work has proved to be convenient for describing the local geometry of the contact of sheets because the x_2 axis is parallel to the self-intersection lines.

The line (24) represents a wedge singularity of the surface of refractive indices of absorbing crystals. The geometric normals $\mathbf{n}_{\pm}$ can be defined for each sheet $n_{\pm}(\mathbf{m})$:

$$\mathbf{n}_{\pm} \parallel \frac{\partial n_{\pm}}{\partial \mathbf{m}}. \tag{26}$$

We will focus on the orientation of $\mathbf{n}_{\pm}$ in the immediate vicinity of the self-intersection lines. When speaking of such normals on the line (24), we will mean that they are defined in the limit of infinitely close vicinity of this line. Here, it is obvious that $\mathbf{n}_{\pm}(m_1 > 0) \parallel \mathbf{n}_{\mp}(m_1 < 0)$; therefore, a single isonormal pair $\mathbf{n}_{\pm}$ characterizes both faces of the wedge surface, on the upper (+) and the lower (–) sheets. Assume, for definiteness, that $\mathbf{n}_{\pm} \equiv \mathbf{n}_{\pm}$ $(m_1 > 0)$.

In a particular case of transparent crystals ($\chi_0 = 0$), when the line (24) shrinks to a point $m_1 = m_2 = 0$, we obtain the following expressions for the refractive indices $n_{\pm}$ near this point and the cone of normals $\mathbf{n}^0(\mathbf{m})$ to the surface:

$$\begin{aligned} n_{\pm} &= \sqrt{\varepsilon_2}[1 + \chi D(\cos\psi \pm 1)], \\ \mathbf{n}^0 &\parallel D(\cos\psi - 1)\mathbf{p} + \mathbf{q}D\sin\psi + \mathbf{m}^{(1)}. \end{aligned} \tag{27}$$

Here, we used the polar coordinates (χ, ψ): $m_1 = \chi\cos\psi$ and $m_2 = \chi\sin\psi$. It is this cone of normals $\mathbf{n}^0(\psi)$ that is shown in Fig. 5. In the case of zero absorption considered here, this cone is simultaneously a cone of group velocities and Poynting vectors; i.e., it represents a refraction cone. In the plane x_1x_2 perpendicular to the direction of $\mathbf{m}^{(1)}$, this cone has a circular cross section. The angular dimensions α_0 and β_0 of this cone in two perpendicular cross sections passing through its geometric axis (see Fig. 5) are given by

$$\tan\alpha_0 = 2D, \quad \tan\frac{\beta_0}{2} = \sqrt{\frac{1}{2}\tan\frac{\alpha_0}{2}\tan\alpha_0}, \tag{28}$$

where D is defined by (23). It is obvious that $\beta_0 > \alpha_0$

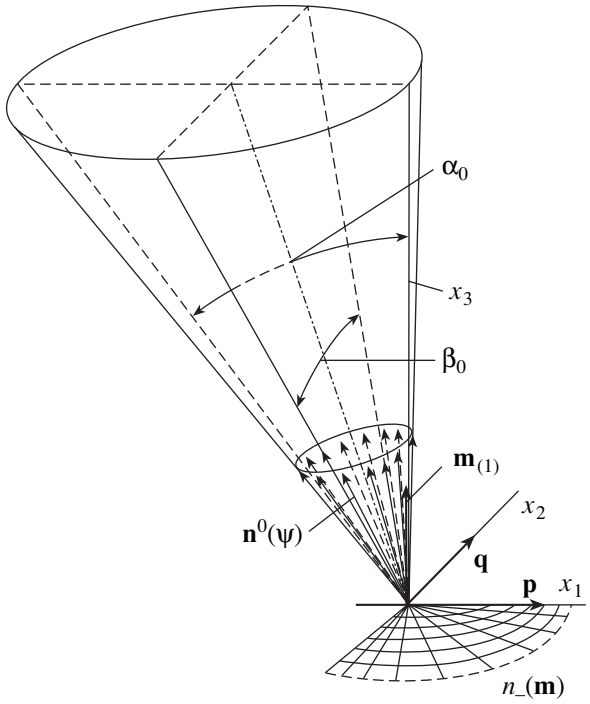


Fig. 5. Refraction cone of a transparent crystal (27). The upper circle that bounds the conical surface and belongs to the plane x_1x_2 represents the trajectory of the endpoint of the energy-flux vector of a circularly polarized wave propagating along the optical axis $\mathbf{m}^{(1)}$.

always; i.e., the cone is not circular. However, the less D is, the closer the cone is to a circular cone with small opening: $\beta_0 \approx \alpha_0 \approx 2D \ll 1$. In practice, the latter condition is usually fulfilled because, as a rule, $2D$ in most crystals is no greater than 0.1; therefore, the opening α_0 of the refraction cone is very small even for record-breaking crystals. For instance, $\alpha_0 = 4°51'$ in ammonium oxalate and $\alpha_0 = 3°18'$ in iodic acid [3].

In absorbing crystals near the line (24), when $|m_1| \ll 1$, the relief of the surface of refractive indices (22) and the field of geometric normals are given by

$$\begin{aligned} \frac{n_{\pm}}{\sqrt{\varepsilon_2}} &= \begin{cases} 1 + \gamma_{\pm}m_1, & m_1 > 0, \\ 1 + \gamma_{\mp}m_1, & m_1 < 0, \end{cases} \\ \mathbf{n}_{\pm} &\parallel \mathbf{m}^{(1)} - D\gamma_{\pm}\mathbf{p}. \end{aligned} \tag{29}$$

Here, the following parameters are introduced:

$$\gamma_{\pm} = 1 \pm 1/\sqrt{1 - m_2^2/\chi_0^2}. \tag{30}$$

When deriving these formulas, we assumed that $|m_1| \ll (\chi_0^2 - m_2^2)/2\chi_0$. When m_1 changes its sign, a transition $n_+ \longleftrightarrow n_-$ occurs, which means that there is a kink of the surface $n_{\pm}(\mathbf{m})$ on the line (24); in this case, the vectors $\mathbf{n}_{\pm}$ change their direction stepwise, thus forming

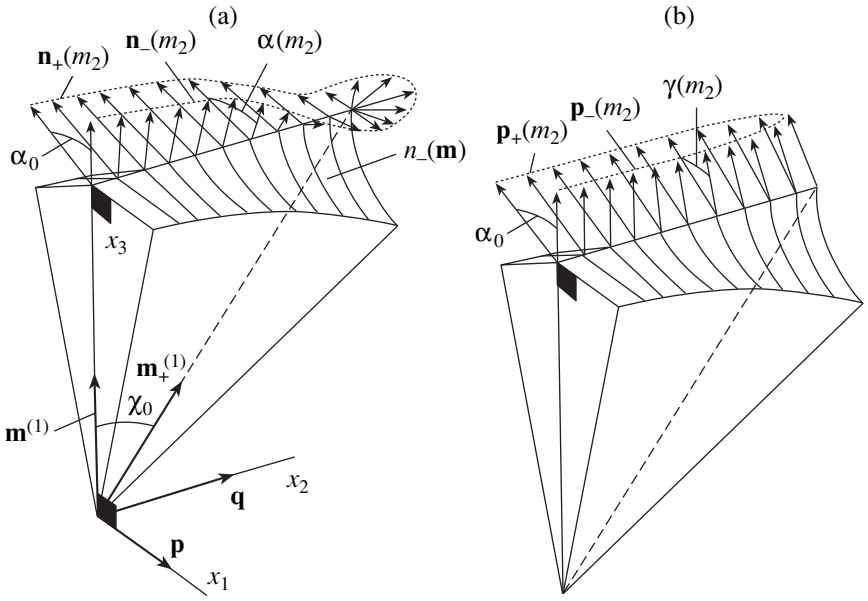


Fig. 6. Schematic distribution of the vector characteristics of wedge refraction; (a) normals $\mathbf{n}_\pm$ on the intersection line of sheets and (b) normalized Poynting vectors $\mathbf{p}_\pm = \bar{\mathbf{P}}_\pm/|\bar{\mathbf{P}}_\pm|$ along the edge of the refraction wedge on the wave sheet $n_-(\mathbf{m})$ (dark quadrangles denote right angles).

a wedge of orientations (see Fig. 6a). The opening angle α of this wedge is defined by

$$\tan\alpha = \frac{D(\gamma_+ - \gamma_-)}{1 + D^2\gamma_+\gamma_-} = \frac{2D\sqrt{1 - m_2^2/\chi_0^2}}{1 - (1 + D^2)m_2^2/\chi_0^2} \qquad (31)$$

as a function of the coordinate. At the center of the intersection line of the sheets, where $m_2 = 0$, the angle α (31) coincides with α_0 (28) and proves to be the minimal opening angle of the wedge of geometric normals (see Fig. 6a). Away from the center, the angle between the normals increases, while the kink of the sheets $n_\pm(\mathbf{m})$ becomes still sharper. When $m_2^2/\chi_0^2 \ll 1$, we have

$$\mathbf{n}_\pm \parallel \begin{cases} D\left(-2 - \dfrac{m_2^2}{2\chi_0^2}\right)\mathbf{p} + \mathbf{m}^{(1)}, \\ \left(\dfrac{m_2^2}{2\chi_0^2}\right)D\mathbf{p} + \mathbf{m}^{(1)}. \end{cases} \qquad (32)$$

Below, we will see that, under a small deviation from the center of the intersection line of the sheets, the wave polarization is slightly different from a linear polarization; therefore, formulas (32) for $\mathbf{n}_\pm$ approximately characterize the direction of group velocities $\mathbf{u}_\pm$ as well.

A sufficiently smooth increase in angle α with m_2 at the center of the wedge changes into a sharp increase in the vicinity of the direction $\mathbf{m}_+^{(1)}$. Let us evaluate the region of the anomalous growth of α. The angle α increases by a factor of two from α_0 to $2\alpha_0$ for $m_2/\chi_0 \approx \sqrt{3}/2 \approx 0.86$. In this region, $\chi_0 - m_2 \ll \chi_0$, and it is convenient to use the notation $m_2 = \chi_0 - \mu$, $0 \le \mu \ll \chi_0$. Then, $\alpha = \pi/2$ for $\mu/\chi_0 \approx \alpha_0^2/8$. For $\alpha_0 = 0.1$, this equality holds at $\mu/\chi_0 \sim 10^{-3}$. A further decrease in μ/χ_0 to zero leads to a still sharper increase in α from $\pi/2$ to π.

A neighborhood of a specific direction of the singular axis requires a separate consideration. Here, it is convenient to introduce polar coordinates μ, $\tilde{\psi}$ with the origin corresponding to the direction $\mathbf{m}_+^{(1)}$: $m_2 = \chi_0 + \mu\cos\tilde{\psi}$, $m_1 = \mu\sin\tilde{\psi}$. In this case, we obtain the following expressions from (22) and (26):

$$n_\pm \approx \sqrt{\varepsilon_2} \pm \delta\cos\frac{\tilde{\psi}}{2}, \quad \delta = (2\mu\chi_0\varepsilon_2)^{1/2}D,$$
$$\mathbf{n}_+ \approx \mathbf{n}_- \approx \mathbf{n}_0 = \mathbf{p}\sin\frac{\tilde{\psi}}{2} + \mathbf{q}\cos\frac{\tilde{\psi}}{2}. \qquad (33)$$

Expression (33) shows that, for $m_2 = \chi_0$ (propagation along the singular axis $\mathbf{m}_+^{(1)}$) the surface $n_\pm(\mathbf{m})$ has a pointed-tip-type singularity (see Fig. 6a). In this case, in contrast to the pronounced asymmetry at the center of the edge, the shape of the wedge at the ends of the edge proves to be symmetric with respect to the plane x_2x_3 that passes through the edge of the wedge and

locally coincides with both faces of the wedge at the singular points $m_1 = 0$ and $m_2 = \pm\chi_0$. Accordingly, the geometric normals at the ends of the edge of the wedge have a fan-type singularity that is parallel to the x_1x_2 plane and symmetric with respect to the x_2x_3 plane (Fig. 6a).

5. WEDGE REFRACTION OF ELECTROMAGNETIC WAVES ON THE SELF-INTERSECTION LINE OF THE SURFACE $n_\pm(\mathbf{m})$

Below, we will see that the polarization of waves near the self-intersection line of the surface $n_\pm(\mathbf{m})$ is linear only on the central x_1x_3 plane, whereas, in other places, it is elliptic. Therefore, the group velocity of the waves can be defined only in a relatively narrow area surrounding the x_1x_3 plane. Hence, it is clear that the field of normals $\mathbf{n}_\pm(\mathbf{m})$ determined above can hardly serve as a characteristic of the wedge refraction of waves considered here. To describe the latter, one has to determine a field of Poynting vectors

$$\mathbf{P}_\pm = \mathrm{Re}\mathbf{E}_\pm \times \mathrm{Re}\mathbf{H}_\pm. \tag{34}$$

For complex vector amplitudes of electric and magnetic fields near the intersection line of the sheets $\mathbf{n}_\pm(\mathbf{m})$, we obtain the following expressions from (4) and (7), which complement formulas (22):

$$\begin{aligned}\mathbf{E}_\pm^0 &\parallel m_2\mathbf{q} + (m_1 + i\chi_1 \pm r)\mathbf{G},\\ \mathbf{H}_\pm^0 &\parallel m_2\mathbf{p} - (m_1 + i\chi_1 \pm r)\mathbf{q}.\end{aligned} \tag{35}$$

Here, $\mathbf{G} = \mathbf{p} + 2D\mathbf{m}^{(1)}$. On the intersection line of the sheets (24), when $m_1 = 0$, we have the following expressions for ac electric and magnetic fields:

$$\begin{aligned}\mathrm{Re}\mathbf{E}_\pm &\parallel g_\pm\mathbf{G}\sin\tilde{\varphi}_\pm - \frac{m_2}{\chi_1}\mathbf{q}\cos\tilde{\varphi}_\pm,\\ \mathrm{Re}\mathbf{H}_\pm &\parallel g_\pm\mathbf{q}\sin\tilde{\varphi}_\pm + \frac{m_2}{\chi_1}\mathbf{p}\cos\tilde{\varphi}_\pm.\end{aligned} \tag{36}$$

Here,

$$g_\pm = 1 \pm \sqrt{1 - (m_2/\chi_0)^2}, \tag{37}$$

$\tilde{\varphi}_\pm$ is a phase that depends on time t and the coordinate $\mathbf{r}$: $\tilde{\varphi}_\pm = \tilde{\varphi} - \varphi_\pm^0$, $\tilde{\varphi} = \mathbf{k}\cdot\mathbf{r} - \omega t$, and $\varphi_\pm^0$ is the reference point of the phase. Formulas (36) describe an elliptic polarization of isonormal waves; in the present case, they have identical phase velocities but different absorption indices. The coefficients of $\sin\tilde{\varphi}_\pm$ and $\cos\tilde{\varphi}_\pm$ in (36) define the vectors $\mathbf{a}_\pm$ and $\mathbf{b}_\pm$ of semiaxes of appropriate ellipses.

Near the midpoint of the intersection line of the sheets, when $|m_2| \ll \chi_0 \equiv |\chi_1|$, from (36) and (37) we obtain

$$\mathrm{Re}\mathbf{E}_\pm \parallel \begin{cases} 2\mathbf{G}\sin\tilde{\varphi}_+ - \dfrac{m_2}{\chi_1}\mathbf{q}\cos\tilde{\varphi}_+, \\ 2\mathbf{q}\cos\tilde{\varphi}_- - \dfrac{m_2}{\chi_1}\mathbf{G}\sin\tilde{\varphi}_-, \end{cases}$$

$$\mathrm{Re}\mathbf{H}_\pm \parallel \begin{cases} 2\mathbf{q}\sin\tilde{\varphi}_+ + \dfrac{m_2}{\chi_1}\mathbf{p}\cos\tilde{\varphi}_+, \\ 2\mathbf{p}\cos\tilde{\varphi}_- + \dfrac{m_2}{\chi_1}\mathbf{q}\sin\tilde{\varphi}_-. \end{cases} \tag{38}$$

These formulas show that, at the center of the wedge, where $m_2 = 0$, the wave polarization is linear, whereas, a little ways from the center, weak ellipticity arises. Finally, when a wave propagates strictly along a singular axis, the electric and magnetic fields are polarized elliptically and circularly, respectively:

$$\begin{aligned}\mathrm{Re}\mathbf{E}_\pm &\parallel \mathbf{G}\sin\tilde{\varphi}_\pm - \mathbf{q}\,\mathrm{sgn}\,\chi_1\cos\tilde{\varphi}_\pm,\\ \mathrm{Re}\mathbf{H}_\pm &\parallel \mathbf{q}\sin\tilde{\varphi}_\pm + \mathbf{p}\,\mathrm{sgn}\,\chi_1\cos\tilde{\varphi}_\pm.\end{aligned} \tag{39}$$

The Poynting vector $\mathbf{P}_\pm$ (34) can be represented as a sum of time-independent (averaged over a period) energy flux $\bar{\mathbf{P}}_\pm$ and a time-dependent flux $\tilde{\mathbf{P}}_\pm$:

$$\begin{aligned}\mathbf{P}_\pm &= \bar{\mathbf{P}}_\pm + \tilde{\mathbf{P}}_\pm,\\ \tilde{\mathbf{P}}_\pm &\parallel \mathbf{c}_\pm\cos 2\tilde{\varphi}_\pm + \mathbf{d}_\pm\sin 2\tilde{\varphi}_\pm.\end{aligned} \tag{40}$$

Here, $\mathbf{c}_\pm$ and $\mathbf{d}_\pm$ are the semiaxes of the corresponding ellipses. According to (34) and (36), on the intersection line of the sheets (24), these vectors are defined by

$$\begin{aligned}\bar{\mathbf{P}}_\pm &\parallel -2Dg_\pm^2\mathbf{p} + \left(\frac{m_2^2}{\chi_0^2} + g_\pm^2\right)\mathbf{m}^{(1)},\\ \mathbf{c}_\pm &\parallel 2Dg_\pm^2\mathbf{p} + \left(\frac{m_2^2}{\chi_0^2} - g_\pm^2\right)\mathbf{m}^{(1)},\\ \mathbf{d}_\pm &\parallel 2Dg_\pm\left(\frac{m_2}{\chi_0}\right)\mathbf{q}.\end{aligned} \tag{41}$$

We omitted identical factors in the expressions for $\bar{\mathbf{P}}_\pm$ and $\tilde{\mathbf{P}}_\pm$. One can see from (41) that the two time-averaged middle vectors $\bar{\mathbf{P}}_+$ and $\bar{\mathbf{P}}_-$, whose orientations can also be represented as

$$\bar{\mathbf{P}}_\pm \parallel \mathbf{m}^{(1)} - Dg_\pm\mathbf{p}, \tag{42}$$

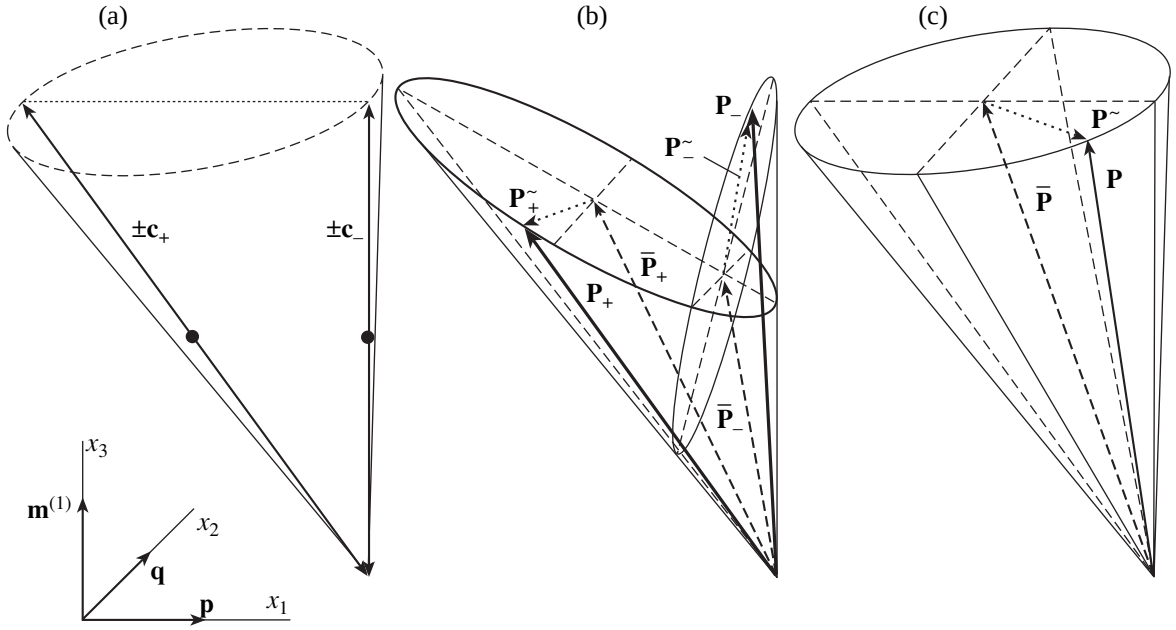


Fig. 7. Trajectories of the endpoints of the Poynting vectors $\mathbf{P}_\pm$ of isonormal waves on the universal refraction cone (Fig. 5) for three directions of the wave normal on the self-intersection line of the surface $n_\pm(\mathbf{m})$: $a - \mathbf{m} \parallel \mathbf{m}^{(1)}$ ($m_2 = 0$), heavy dots indicate the positions of the endpoints of $\bar{\mathbf{P}}_\pm$; (b) intermediate orientation, $0 < m_2 < \chi_0$, two elliptic sections of the cone are parallel to the vector $\mathbf{q}$ (to the edge of the wedge); (c) $\mathbf{m} \parallel \mathbf{m}_+^{(1)}$ ($m_2 = \chi_0$), the horizontal circle in the section corresponds to the merging of two ellipses for a singular direction.

always lie in the plane x_1x_3, which is orthogonal to the edge of the wedge, while the orientations of these vectors smoothly vary as the coordinate m_2 is varied, making an angle of γ (Fig. 6b):

$$\tan\gamma = \frac{D(g_+ - g_-)}{1 + D^2 g_+ g_-} = \frac{2D\sqrt{1 - m_2^2/\chi_0^2}}{1 + D^2 m_2^2/\chi_0^2}. \qquad (43)$$

It is interesting to note that expressions (42) and (43) are similar to (29) and (31).

The ends of the vectors $\bar{\mathbf{P}}_\pm$ correspond to the centers of elliptic trajectories described by the ends of the vectors $\tilde{\mathbf{P}}_+$ and $\tilde{\mathbf{P}}_-$. The minor semiaxes of the ellipses are parallel to the edge of the wedge, $\mathbf{d}_\pm \parallel \mathbf{q}$, while their eccentricity monotonically decreases from unity to zero as m_2 increases from 0 to χ_0. Simultaneously, the slope of the major semiaxes $\mathbf{c}_\pm$ with respect to the horizontal plane decreases to zero: for $m_2 = \chi_0$, we have $\mathbf{c}_\pm \parallel \mathbf{p}$. A rather tedious analysis leads to a very interesting picture: all elliptic orbits can be represented as lines that arise in the plane sections of a universal cone that coincides with the cone of refraction of a transparent crystal (see formulas (28) and Fig. 5). A variation in m_2 changes only the orientations of the secant planes, which remain parallel to the edge of the wedge, i.e., to the vector $\mathbf{q}$. In this case, the full energy-flux vectors $\mathbf{P}_\pm$ traverse all the generators of the universal cone, running over the corresponding elliptic orbits (see Fig. 7).

Consider the behavior of the characteristics of interest near the midpoint and at the ends of the intersection line of the sheets in greater detail. When $|m_2| \ll \chi_0$, the energy fluxes of isonormal waves are given by the relations

$$\begin{cases} \bar{\mathbf{P}}_+ \parallel D\left(-2 + \dfrac{m_2^2}{2\chi_0^2}\right)\mathbf{p} + \mathbf{m}^{(1)}, \\ \mathbf{c}_+ \parallel D\left(2 - \dfrac{m_2^2}{2\chi_0^2}\right)\mathbf{p} - \left(1 - \dfrac{m_2^2}{2\chi_0^2}\right)\mathbf{m}^{(1)}, \\ \mathbf{d}_+ \parallel \left(\dfrac{m_2}{\chi_0}\right)D\mathbf{q}; \end{cases} \qquad (44)$$

$$\begin{cases} \bar{\mathbf{P}}_- \parallel -\left(\dfrac{m_2^2}{2\chi_0^2}\right)D\mathbf{p} + \mathbf{m}^{(1)}, \\ \mathbf{c}_- \parallel \left(\dfrac{m_2^2}{2\chi_0^2}\right)D\mathbf{p} + \left(1 - \dfrac{m_2^2}{4\chi_0^2}\right)\mathbf{m}^{(1)}, \\ \mathbf{d}_- \parallel \left(\dfrac{m_2}{\chi_0}\right)D\mathbf{q}. \end{cases} \qquad (45)$$

These formulas show that the time-averaged vectors $\bar{\mathbf{P}}_+$ and $\bar{\mathbf{P}}_-$ of two isonormal waves start to converge as they move away from the wedge center. Moreover, as we expected, the initial orientations of these vectors for $m_2 = 0$ coincide with the corresponding directions of the normals $\mathbf{n}_\pm$ (see (32)); however, the latter, unlike the fluxes $\bar{\mathbf{P}}_\pm$, diverge as m_2 increases, rather than converge (Fig. 6).

Near the singular axis, when $m_2 = \chi_0 - \mu$, $0 \le \mu \ll \chi_0$, the energy fluxes are characterized by the expressions

$$\begin{aligned}
\bar{\mathbf{P}}_\pm &\parallel \bar{\mathbf{P}}_0 \pm \sqrt{\frac{2\mu}{\chi_0}}(-2D\mathbf{p} + \mathbf{m}^{(1)}),\\
\bar{\mathbf{P}}_0 &= -D\mathbf{p} + \mathbf{m}^{(1)},\\
\bar{\mathbf{c}}_\pm &\parallel D\mathbf{p} \pm \sqrt{\frac{2\mu}{\chi_0}}(2D\mathbf{p} - \mathbf{m}^{(1)}),\\
\mathbf{d}_\pm &\parallel D\left(1 \pm \sqrt{\frac{2\mu}{\chi_0}}\right)\mathbf{q}.
\end{aligned} \tag{46}$$

These relations show that, for a wave propagating strictly along a singular axis, when $\mu = 0$, the directions of the mean fluxes of isonormal waves coincide, $\bar{\mathbf{P}}_+ \parallel \bar{\mathbf{P}}_- \parallel \bar{\mathbf{P}}_0$, while their endpoints represent the centers of appropriate circles that the vectors $\mathbf{P}_+^{\sim}$ and $\mathbf{P}_-^{\sim}$ describe in the x_1x_2 plane (Fig. 7c). The full vectors $\mathbf{P} = \bar{\mathbf{P}} + \mathbf{P}^{\sim}$ behave in exactly the same way as the flux vectors in transparent crystals when a circularly polarized wave propagates along the optical axis (Fig. 5). Formulas (46) show that the average flux vector $\bar{\mathbf{P}}_0$ retains its asymmetric orientation in the x_1x_3 plane.

Away from the singular axis, when the waves lose their circular polarization, the vectors $\bar{\mathbf{P}}_+$ and $\bar{\mathbf{P}}_-$ diverge more and more as μ increases (Fig. 6b). In this case, the circles described by the vectors $\mathbf{P}_+^{\sim}$ and $\mathbf{P}_-^{\sim}$ for $\mu = 0$ turn into gradually elongating ellipses. The angle γ (43) between the vectors $\bar{\mathbf{P}}_+$ and $\bar{\mathbf{P}}_-$ smoothly increases when moving toward the wedge center: it attains its maximum equal to α_0 at the center; i.e., as we have already pointed out above, it coincides with the angle between the normals $\mathbf{n}_\pm$ (see Fig. 6).

6. DISCUSSION

Thus, absorption gives rise to quite nontrivial topological features both in the geometry of a contact between degenerate wave surfaces (Figs. 2 and 4) and in the distribution of complex vector polarization fields that are characterized by the Poincaré index $n = 1/4$ in the vicinity of singular axes (Fig. 3). In the propagation directions that are close to the self-intersection lines of the surfaces of refractive indices, internal wedge refraction is predicted. Analysis has shown that the character of refraction is analogous to a similar phenomenon in the acoustics of nonabsorbing crystals only in the central region of a wedge, where the polarization is close to a linear polarization. Away from this region, the ellipticity of polarization increases, while the direction of the Poynting vector averaged over a period, $\bar{\mathbf{P}}_\pm$, deviates more and more from the corresponding normal $\mathbf{n}_\pm$ toward the surface of refractive indices. In this case, the nonaveraged Poynting vectors $\mathbf{P}_\pm$ on the entire edge of the wedge rotate along the same cone of directions, which coincides with the refraction cone of a transparent crystal, whereas the endpoints of these vectors run over different plane (elliptic) sections of the cone that depend on the position of the wave normal on the edge of the wedge. Along singular axes (at the ends of the edge of the wedge), these sections become circular, while the refraction pattern fully reproduces the conical refraction of circularly polarized waves in a transparent crystal (Fig. 5).

In the last sections of this paper, when analyzing the geometry of the self-intersection of the surface of refractive indices and discussing the specific features of the wedge refraction of electromagnetic waves, we restricted ourselves to the analysis of orthorhombic crystals. In this case, the x_1x_3 plane passing through the center of the wedge represents a symmetry plane, and the pictures on either side of this plane are equivalent. The center of the wedge corresponds to a linear polarization of waves, which is preserved on the whole this plane, while the signs of rotation of elliptically polarized waves on different sides of the symmetry plane are opposite. Obviously, the signs of rotation of circularly polarized waves that propagate along the crystallographically equivalent singular axes $\mathbf{m}_+^{(j)}$ and $\mathbf{m}_-^{(j)}$ are also opposite in this case. Naturally, in monoclinic and triclinic crystals that have no such symmetry plane connecting the singular directions, these directions are not equivalent, and the general picture is somewhat distorted. Nevertheless, according to analysis, all the main features of the phenomenon are qualitatively the same even in the absence of symmetry. In particular, a triclinic perturbation of symmetry does not remove the line of orientations for the propagation of linearly polarized waves on the unit sphere of directions $\mathbf{m} \cdot \mathbf{m} = 1$ but distorts it. Such lines are well known in absorbing crystals of arbitrary anisotropy [13, 14] and do not disappear even under the additional switching on of optical activity [15].

Let us touch upon the possibility of experimental observation of wedge refraction in absorbing crystals. In the experiment, the physical parameter with which one should compare the splitting angle of optical axes is the divergence of the optical beam. In particular,

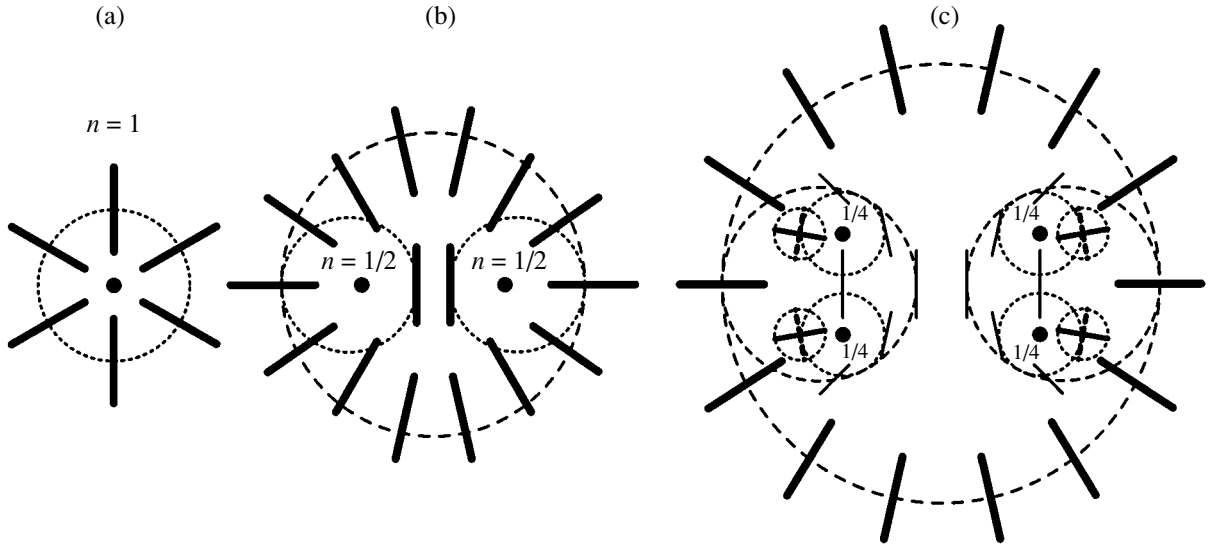


Fig. 8. Vector polarization fields in one of isonormal branches around (a) a tangency point of degeneracy in a uniaxial crystal, (b) a pair of conical points in a biaxial transparent crystal, and (c) a quadruple of singular axes in an absorbing crystal (the rotation of the major semiaxes of polarization ellipses are shown near singular points).

when a light beam is transmitted along the direction $\mathbf{m}^{(1)}$ (or $\mathbf{m}^{(2)}$) with the divergence $\Delta\vartheta$ much greater than the angular splitting $2\chi_0$ of optical axes, the pattern observed differs only slightly from the ordinary conical refraction: the fine details associated with the absorption do not manifest themselves. However, if we choose a laser beam with a divergence much less than the angular splitting of axes, then it becomes possible to observe a wedge refraction as well as the anomalies described above near new positions of optical axes. Since the divergence of a laser beam may be very small, $\Delta\vartheta \sim 1'$, the corresponding estimate of admissible relations between the material parameters, $2\chi_0 \gg \Delta\vartheta$, yields the inequality $\eta'/d \gg 3 \times 10^{-4}$, which leads to inessential restrictions on the choice of the objects of investigation.

Let us dwell on an interesting topological aspect of the problem. It is well known that, in optically uniaxial crystals, the degeneracy corresponding to the tangency of the wave sheets and characterized by the Poincaré index $n = 1$ (Fig. 8a) is unstable with respect to any real perturbation $\delta\boldsymbol{\varepsilon}$ of the "uniaxial" symmetry of the tensor $\boldsymbol{\varepsilon}$. Any such perturbation (due to external actions on a crystal or a phase transformation) splits the tangent degeneracy into two conical ones each of which has topological "charge" $n = 1/2$, so that the total Poincaré index remains equal to the original one (Fig. 8b). Conical degeneracies are stable with respect to the variations of anisotropy, which only shift the degeneracy points but do not change their index *n*. However, as we have seen, this stability does not apply to imaginary perturbations $\delta\boldsymbol{\varepsilon} = i\delta\boldsymbol{\varepsilon}'$, which correspond to absorption. According to the results obtained above, the switching on of absorption also splits conical degeneracies into two pairs of singular points with indices $n = 1/4$ (Fig. 8c), so that the conservation of the original index again holds: $\frac{1}{4} \times 4 = 1$. It is this fact that guarantees a continual switching off of the two types of perturbation considered above; as a result, the regions with split axes shown in Figs. 8b and 8c (in the circles drawn by dotted lines) shrink to points, and we return to the original pattern of a transparent biaxial, and then also uniaxial, crystal.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 03-02-16871) and by the Polish–Japanese Institute of Information Technologies (project no. PJ/MKT/02/2003).

REFERENCES

1. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media*, 3rd ed. (Nauka, Moscow, 1992; Pergamon, New York, 1984).
2. Yu. I. Sirotin and M. P. Shaskolskaya, *Fundamentals of Crystal Physics* (Nauka, Moscow, 1975; Mir, Moscow, 1982).
3. F. I. Fedorov and V. V. Filippov, *Reflection and Refraction of Light by Transparent Crystals* (Nauka i Tekhnika, Minsk, 1976).
4. A. G. Khatkevich, Kristallografiya **7**, 742 (1962) [Sov. Phys. Crystallogr. **7**, 601 (1962)]; Kristallografiya **7**, 916 (1962) [Sov. Phys. Crystallogr. **7**, 742 (1962)].
5. E. Henneke and R. J. Green, J. Acoust. Soc. Am. **45**, 1367 (1969).

6. F. I. Fedorov, *Theory of Gyrotropy* (Nauka i Tekhnika, Minsk, 1976).

7. V. M. Agranovich and V. L. Ginzburg, *Crystal Optics with Spatial Dispersion, and Excitons*, 2nd ed. (Nauka, Moscow, 1979; Springer, New York, 1984).

8. V. N. Lyubimov, in *Proceedings of International Workshop on Dissipation in Physical Systems, Borkow, 1995*, Ed. by A. Radowicz (Zeszyty Naukowe Politechniki Swetokrzyskiej, Kielce, Poland, 1995); Mechanika **59**, 77 (1995).

9. V. I. Alshits and V. N. Lyubimov, Kristallografiya **45**, 731 (2000) [Crystallogr. Rep. **45**, 670 (2000)].

10. V. I. Alshits, V. N. Lyubimov, A. Radowicz, and L. A. Shuvalov, Kristallografiya **46**, 519 (2001) [Crystallogr. Rep. **46**, 468 (2001)].

11. V. I. Alshits, V. N. Lyubimov, and L. A. Shuvalov, Kristallografiya **46**, 742 (2001) [Crystallogr. Rep. **46**, 673 (2001)].

12. V. I. Alshits and V. N. Lyubimov, in *Proceedings of 2nd International Workshop on Dissipation in Physical Systems, Borkow, 1997*, Ed. by A. Radowicz (Zeszyty Naukowe Politechniki Swetokrzyskiej, Kielce, Poland, 1997); Mechanika **66**, 15 (1998).

13. G. Sziwessy, *Kristalloptik*, Handb. Phys. **20** (Springer, Berlin, 1928).

14. A. M. Goncharenko, Dokl. Akad. Nauk BSSR **2**, 457 (1958).

15. M. V. Berry and M. R. Dennis, Pros. R. Soc. London, Ser. A **459**, 1261 (2003).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 882–887.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1012–1017.

ATOMS, SPECTRA, RADIATION

A Particle in the Field of Point Potentials

A. S. Chikhachev

State Research Center All-Russia Electrical Engineering Institute, Moscow, 111250 Russia
e-mail: churchev@mail.ru

Received November 1, 2003

Abstract—Bound states of a particle in the field of two pointlike δ centers are considered in the 3D problem. The exact solution to the Schrödinger equation is obtained for a system of scattered centers. © *2004 MAIK "Nauka/Interperiodica".*

Point (zero-range) potentials were used for the first time by Fermi [1] in the 1D problem to study the shift of spectral lines. The application of the δ function as the potential in the Schrödinger equation makes it possible to obtain a compact analytic solution. This facilitated subsequent wide application of point potentials. In addition, actual forces (e.g., nuclear forces or the forces emerging during the screening of the Coulomb potential) can often be treated as short-range forces.

The method of zero-range potential has been successfully used in various branches of physics. In nuclear physics, this method was employed for studying the scattering of particles (including that by two stationary centers). Point interactions were used for solving three-body problems and for studying nucleon tunneling. The method is also extensively used in atomic physics. It was employed for describing molecular systems, in the theory of atomic collisions (e.g., in studying charge exchange and neutralization of atomic particles), and in solid-state theory.

Extensive literature on this problem is presently available. Steady-state solvable point models were considered in monograph [2] in 1D, 2D, and 3D cases. One-dimensional steady-state equations with δ and δ' potentials were studied in [3]. A number of questions associated with 3D point potentials were studied in detail in monographs [4, 5].

The models with pointlike interactions are distinguished, among other things, by their applicability for constructing exact solutions of nonstationary problems. Breit [6] was apparently the first to study nucleon tunneling with the help of a nonstationary 1D model. Nucleon tunneling was also analyzed in [7, 8].

An analogous nonstationary model equation with scattered δ potentials was studied in [9–15]. The exact solution describing a "bound" state (i.e., the state described by a rapidly (exponentially) decreasing function of coordinates in the case of identical constants characterizing the depth of the levels of pointlike centers) was obtained for the first time in [9]. In addition, an important feature of the problem on scattering of pointlike centers with a constant velocity was also considered in [9]. If the Schrödinger equation is used in integral form, an equation with a difference kernel appears under a certain substitution of variables. This makes it possible to obtain exact solutions for a wide class of problems. A relation for the charge-exchange probability was obtained in [10] for an arbitrary relation between the relative velocity v and parameter α characterizing the level depth. A compact expression for the recharging probability amplitude for δ wells characterized by different depths α and β is given in [11]. This expression is valid for any relation between quantities α, β, and v. A solution with an oscillatory asymptotic form was obtained (i.e., "free" states were determined) in [12] for scattering of identical wells. The Cauchy problem for different wells and for scattering of δ centers from different points was also solved in [12]. The "bound" state of dispersing centers characterized by different depths of a single bound level was determined in [13]. In [14], the propagator in the field of two dispersing centers of various depths was determined for the first time. The same problem was solved later in [15].

This communication is devoted to analysis of pointlike 3D systems. In Section 1, the form of the potential operator for a moving 3D well is determined. In Sections 2 and 3, the bound states for two stationary wells (including those with different depths) are determined. The exact solution to the Schrödinger equation for identical dispersing wells is obtained in Sections 4 and 5, where the "bound" state (i.e., the state described by a rapidly decreasing ψ function of coordinates) is determined. The solution in the case of dispersal from different points is also obtained. The results are discussed in Section 6.

1. In the spherically symmetric case, the bound state of a δ center resting at the origin can be described by the Schrödinger equation of the form

$$i\frac{\partial\psi}{\partial t} + \frac{1}{2}\Delta\psi = \frac{2\pi}{\kappa}\delta(\mathbf{r})(\psi + \mathbf{r}\nabla\psi), \qquad (1)$$

1063-7761/04/9805-0882$26.00 © 2004 MAIK "Nauka/Interperiodica"

where $\psi(\mathbf{r}, t)$ is the psi function; in the system of units used here, $\hbar = m = e = 1$, $\hbar$ is Planck's constant, m is the mass, and e is the charge.

The solution to Eq. (1) describing a bound state has the form

$$\psi = \frac{\text{const}}{r}\exp\left(-\kappa r + \frac{i\kappa^2}{2}t\right). \tag{2}$$

The validity of relation (2) can be verified, for example, by substituting this solution into Eq. (1) written in integral form:

$$\psi = \frac{2\pi}{\kappa}\int_{-\infty}^{t} dt \int d\mathbf{r}' G^{(+)}(\mathbf{r}-\mathbf{r}', t-t') \times \delta(\mathbf{r}')[\psi(\mathbf{r}', t') + \mathbf{r}'\nabla\psi(\mathbf{r}', t')], \tag{3}$$

where $G^{(+)}$ is the retarded Green function (see [4, 5]),

$$G^{(+)} = -\frac{i\sigma(t-t')}{[2\pi i(t-t')]^{3/2}}\exp\left[i\frac{(\mathbf{r}-\mathbf{r}')^2}{2(t-t')}\right]. \tag{4}$$

Here, $\sigma(x) = 0$ if $x < 0$ and $\sigma(x) = 1$ for $x \geq 0$.

If we have a δ center moving with velocity $\mathbf{v}$, we must modify the right-hand side of Eq. (1) and write the equation as follows:

$$i\frac{\partial\psi}{\partial t} + \frac{1}{2}\Delta\psi = \frac{2\pi}{\kappa}\delta(\mathbf{r}-\mathbf{v}t)\{\psi[1 - i\mathbf{v}(\mathbf{r}-\mathbf{v}t)] + (\mathbf{r}-\mathbf{v}t)\nabla\psi\}. \tag{5}$$

This equation has the solution

$$\psi = \frac{\text{const}}{|\mathbf{r}-\mathbf{v}t|} \times \exp\left\{-\kappa|\mathbf{r}-\mathbf{v}t| + \frac{it}{2}\kappa^2 + i\mathbf{v}(\mathbf{r}-\mathbf{v}t) + \frac{iv^2t}{2}\right\}. \tag{6}$$

2. Let us consider the case of two stationary centers located at points $\mathbf{r} = \mathbf{r}_0$ and $\mathbf{r} = -\mathbf{r}_0$. The corresponding Schrödinger equation has the form

$$i\frac{\partial\psi}{\partial t} + \frac{1}{2}\Delta\psi = \frac{2\pi}{\kappa}[\delta(\mathbf{r}-\mathbf{r}_0)(\psi + (\mathbf{r}-\mathbf{r}_0)\nabla\psi) + \delta(\mathbf{r}+\mathbf{r}_0)(\psi + (\mathbf{r}+\mathbf{r}_0)\nabla\psi)]. \tag{7}$$

Here, the quantities κ and r_0 determine the energy level κ_0. Representing the equation in integral form with the help of the retarded Green function, we obtain

$$\psi = \frac{1}{(2\pi i)^{1/2}\kappa}\int_{-\infty}^{t}\frac{dt'\kappa_0 C(t')}{(t-t')^{3/2}} \times\left\{\exp\left[\frac{i(\mathbf{r}-\mathbf{r}_0)^2}{2(t-t')}\right] + \exp\left[\frac{i(\mathbf{r}+\mathbf{r}_0)}{2(t-t')}\right]\right\}. \tag{8}$$

While deriving this equation, we took into account the boundary condition

$$[\psi + (\mathbf{r}-\mathbf{r}_0)\nabla\psi]_{\mathbf{r}\to\mathbf{r}_0} = [\psi + (\mathbf{r}+\mathbf{r}_0)\nabla\psi]_{\mathbf{r}\to-\mathbf{r}_0} = -\kappa_0 C(t).$$

Applying the operator $[\psi + (\mathbf{r}-\mathbf{r}_0)\nabla\psi]_{\mathbf{r}\to\mathbf{r}_0}$ to expression (8), we obtain the following equation for $C(t)$:

$$-\kappa_0 C(t) = \frac{1}{\kappa\sqrt{2\pi i}}\int_{-\infty}^{t}\frac{dt'C(t')}{(t-t')^{3/2}}\kappa_0 \times\exp\left[\frac{i(\mathbf{r}-\mathbf{r}_0)^2}{2(t-t')}\right]\left(1 + \frac{i(\mathbf{r}-\mathbf{r}_0)}{t-t'}\right)_{\mathbf{r}\to\mathbf{r}_0} + \frac{1}{\kappa\sqrt{2\pi i}}\int_{-\infty}^{t}\frac{dt'C(t')}{(t-t')^{3/2}}\kappa_0\exp\left(\frac{2ir_0^2}{t-t'}\right).$$

Before we set $\mathbf{r} = \mathbf{r}_0$ in the first term, we must carry out integration by parts, assuming that the substitution

$$\left.\frac{C(t')\exp\left[\frac{i(\mathbf{r}-\mathbf{r}_0)^2}{2(t-t')}\right]}{\sqrt{t-t'}}\right|_{t'=-\infty}^{t'=t}$$

is equal to zero. For $t' \longrightarrow t$, this equality can be ensured by the presence of the exponential factor. Representing the exponent $i(\mathbf{r}-\mathbf{r}_0)^2/2(t-t')$ in dimensional variables, we obtain $im(\mathbf{r}-\mathbf{r}_0)^2/2\hbar(t-t')$ (see [5]). In our calculation, we must carry out the substitution $\hbar \longrightarrow (1 - i\varepsilon)$, where the positive quantity $\varepsilon \longrightarrow 0$; before proceeding to the limit $\varepsilon \longrightarrow 0$, we must ensure that $t' \longrightarrow t$. Quantity ε does not appear subsequently in the final expressions. This operation is essentially similar to the operation "subtraction of diverging terms" proposed in [4]. For $C(t)$, we have the equation

$$C(t) = \frac{2}{\kappa\sqrt{2\pi i}}\int_{-\infty}^{t}\frac{dt'C(t')}{\sqrt{t-t'}} - \frac{1}{\kappa\sqrt{2\pi i}}\int_{-\infty}^{t}\frac{dt'C(t')\exp\left(\frac{2ir_0^2}{t-t'}\right)}{(t-t')^{3/2}}. \tag{9}$$

If we set in this equation

$$C(t) = \exp\left(\frac{i\kappa_0^2 t}{2}\right),$$

we obtain the relation

$$\kappa = \kappa_0 - \frac{\exp(-2\kappa_0 r_0)}{2r_0}. \quad (10)$$

The depth of the bound level κ_0 depends not only on κ, but also on the distance $2r_0$ between the centers. For $r_0 \longrightarrow 0$, a smooth transition of solution (7) to the solution for a single δ center does not exist. This distinguishes the problem considered here from the 1D case, in which such a smooth transition is possible (see [15]). Solution (8) has the form

$$\psi = \text{const}\left\{\frac{\exp(-\kappa_0|\mathbf{r}-\mathbf{r}_0|)}{|\mathbf{r}-\mathbf{r}_0|} + \frac{\exp(-\kappa_0|\mathbf{r}+\mathbf{r}_0|)}{|\mathbf{r}+\mathbf{r}_0|}\right\} \times \exp\left(\frac{i\kappa_0^2 t}{2}\right). \quad (11)$$

If we seek the solution to Eq. (7) in form (11), we can readily obtain relation (10) also.

3. In the case of two stationary δ centers, bound states with two different values of parameters κ_1 and κ_2 characterizing the level depth can exist.

We assume that the solution for the ψ function has the form of the sum of four terms:

$$\psi = \alpha_1\frac{\exp\left(\frac{i\kappa_1^2 t}{2} - \kappa_1|\mathbf{r}-\mathbf{r}_0|\right)}{|\mathbf{r}-\mathbf{r}_0|} + \alpha_2\frac{\exp\left(\frac{i\kappa_2^2 t}{2} - \kappa_2|\mathbf{r}-\mathbf{r}_0|\right)}{|\mathbf{r}-\mathbf{r}_0|} + \beta_1\frac{\exp\left(\frac{i\kappa_1^2 t}{2} - \kappa_1|\mathbf{r}+\mathbf{r}_0|\right)}{|\mathbf{r}+\mathbf{r}_0|} + \beta_2\frac{\exp\left(\frac{i\kappa_2^2 t}{2} - \kappa_2|\mathbf{r}+\mathbf{r}_0|\right)}{|\mathbf{r}+\mathbf{r}_0|}. \quad (12)$$

This gives

$$i\frac{\partial\psi}{\partial t} + \frac{1}{2}\Delta\psi = -2\pi\left\{\left[\alpha_1\exp\left(\frac{i\kappa_1^2 t}{2}\right) + \alpha_2\exp\left(\frac{i\kappa_2^2 t}{2}\right)\right]\delta(\mathbf{r}-\mathbf{r}_0) + \left[\beta_1\exp\left(\frac{i\kappa_1^2 t}{2}\right) + \beta_2\exp\left(\frac{i\kappa_2^2 t}{2}\right)\right]\delta(\mathbf{r}+\mathbf{r}_0)\right\}.$$

The right-hand side of this equation can be written in the form

$$C_1\delta(\mathbf{r}-\mathbf{r}_0)\left[\psi + (\mathbf{r}-\mathbf{r}_0)\cdot\frac{\partial\psi}{\partial\mathbf{r}}\right] + C_2\delta(\mathbf{r}+\mathbf{r}_0)\left[\psi + (\mathbf{r}+\mathbf{r}_0)\cdot\frac{\partial\psi}{\partial\mathbf{r}}\right].$$

The substitution of expression (12) into this relation ultimately leads to the system of four equations,

$$\begin{aligned}
-2\pi\alpha_1 &= \left(-\kappa_1\alpha_1 + \beta_1\frac{\exp(-2\kappa_1 r_0)}{2r_0}\right)C_1,\\
-2\pi\alpha_2 &= \left(-\kappa_2\alpha_2 + \beta_2\frac{\exp(-2\kappa_2 r_0)}{2r_0}\right)C_1,\\
-2\pi\beta_1 &= \left(-\kappa_1\beta_1 + \alpha_1\frac{\exp(-2\kappa_1 r_0)}{2r_0}\right)C_2,\\
-2\pi\beta_2 &= \left(-\kappa_2\beta_2 + \alpha_2\frac{\exp(-2\kappa_2 r_0)}{2r_0}\right)C_2.
\end{aligned} \quad (13)$$

From the condition of existence of nonzero solution of system (13), we can find two equations for determining constants C_1 and C_2:

$$\left(1-\frac{\kappa_1 C_1}{2\pi}\right)\left(1-\frac{\kappa_1 C_2}{2\pi}\right) = C_1 C_2\frac{\exp(-4\kappa_1 r_0)}{4r_0^2}, \quad (14)$$

$$\left(1-\frac{\kappa_2 C_1}{2\pi}\right)\left(1-\frac{\kappa_2 C_2}{2\pi}\right) = C_1 C_2\frac{\exp(-4\kappa_2 r_0)}{4r_0^2}. \quad (15)$$

These equations define the values of C_1 and C_2 in terms of κ_1, κ_2, and r_0. It should be noted that the equality $C_1 = C_2$ is possible only for $\kappa_1 = \kappa_2$.

4. Let us now consider a more complex case of moving δ centers. We assume that the centers are scattered from point $\mathbf{r} = 0$ with equal and opposite velocities. The Schrödinger equation can be written in the form

$$i\frac{\partial\psi}{\partial t} + \frac{1}{2}\Delta\psi = \frac{2\pi}{\kappa}\{\delta(\mathrm{r}-\mathbf{v}t)[\psi(1-i\mathbf{v}\cdot(\mathbf{r}-\mathbf{v}t)) + (\mathbf{r}-\mathbf{v}t)\nabla\psi] + \delta(\mathbf{r}+\mathbf{v}t) \times [\psi(1+i\mathbf{v}\cdot(\mathbf{r}+\mathbf{v}t)) + (\mathbf{r}+\mathbf{v}t)\nabla\psi]\}. \quad (16)$$

Using the advanced Green function

$$G^{(-)} = -\frac{i\sigma(t'-t)}{(2\pi i(t-t'))^{3/2}}\exp\left[-\frac{i(\mathbf{r}-\mathbf{r}')^2}{2(t'-t)}\right]$$

(see [5]) and the boundary conditions at points $\mathbf{r} = \pm\mathbf{v}t$,

$$\psi \approx C(t)\left(\frac{1}{|\mathbf{r} \pm \mathbf{v}t|} - \kappa\right),$$

we can write Eq. (16) in integral form:

$$\psi = \frac{i}{(-2\pi i)^{3/2}}\frac{2\pi}{\kappa}\int_t^\infty \frac{dt' C(t')}{(t'-t)^{3/2}} \times\left\{-\kappa\exp\left[-\frac{i(\mathbf{r}-\mathbf{v}t')^2}{2(t'-t)}\right] - \kappa\exp\left[-\frac{i(\mathbf{r}+\mathbf{v}t)^2}{2(t'-t)}\right]\right\}. \quad (17)$$

We denote

$$C(t)\exp\left(-\frac{iv^2 t}{2}\right) = g(t).$$

Analogously to the procedure used in Section 2, we can obtain the following expression for $g(t)$ from Eq. (17):

$$g(t) = -\frac{2i}{\kappa\sqrt{2\pi i}}\int_t^\infty \frac{\dot g(t')dt'}{\sqrt{t'-t}} - \frac{i}{\kappa\sqrt{2\pi i}}\int_t^\infty \frac{g(t')dt'}{(t'-t)^{3/2}}\exp\left(-\frac{2iv^2 tt'}{t'-t}\right). \quad (18)$$

While deriving this expression, we also integrated by parts and assumed that

$$\frac{2}{\sqrt{t'-t}}\exp\left\{\frac{iv^2 t'}{2} - \frac{(\mathbf{r}-\mathbf{r}')^2}{2(t'-t)}\right\}\Bigg|_{t'=t}^{t'=\infty} = 0,$$

which is possible under the substitution $\hbar \longrightarrow \hbar(1+i\varepsilon)$ and $\varepsilon \longrightarrow 0$. We set

$$h(\tau) = \frac{g\left(\frac{1}{\tau}\right)}{\sqrt{\tau}}, \quad \tau = \frac{1}{t}, \quad \tau' = \frac{1}{t'}.$$

Equation (18) leads to

$$h(\tau) = \frac{i}{\kappa\sqrt{2\pi i}}\int_0^\tau \frac{d\tau'(2h(\tau')\tau' + h(\tau'))}{\sqrt{\tau-\tau'}} - \frac{\tau}{\kappa\sqrt{2\pi i}}\int_0^\tau \frac{d\tau' h(\tau')}{(\tau-\tau')^{3/2}}\exp\left(-\frac{2iv^2}{\tau-\tau'}\right). \quad (19)$$

This equation can be solved using the Laplace transformation in the same way as in the case of the 1D nonstationary problem (see [9, 11, 12]). Assuming that

$$H(p) = \int_0^\infty h(\tau)\exp(-p\tau)d\tau,$$

we can obtain from Eq. (19)

$$\kappa H(p) = -\frac{i}{\sqrt{2\pi i}}\sqrt{\frac{\pi}{p}}\left[H(p) + 2p\frac{dH}{dp}\right] - \frac{i}{\sqrt{2\pi i}}\sqrt{\frac{\pi}{2iv^2}}\frac{d}{dp}[\exp(-2v\sqrt{2ip})H(p)]. \quad (20)$$

If we set $p = q^2/2i$, we can obtain from expression (20) the solution for H in the form

$$H(q) = \frac{\text{const}}{a(q)}\exp\left\{i\kappa\int_0^q \frac{q'dq'}{a(q')}\right\}, \quad (21)$$

where

$$a(q) = q - \frac{\exp(-2vq)}{2v}.$$

For $g(t)$, we obtain the following expression:

$$g(t) = \frac{\text{const}}{\sqrt{t}}\int_L qdqH(q)\exp\left(-\frac{iq^2}{2t}\right). \quad (22)$$

Path L may have the form of two rays located in the second and fourth quadrants of complex plane q and will be defined more exactly below. Since

$$C(t) = g(t)\exp\left(\frac{iv^2 t}{2}\right),$$

the solution to the Schrödinger equation is given by

$$\psi(\mathbf{r}, t) = \text{const}\int_t^\infty \frac{dt'\exp\left(\frac{iv^2 t'}{2}\right)}{(t'-t)^{3/2}\sqrt{t'}} \times\int_L qdq\exp\left(-\frac{iq^2}{2t'}\right)\frac{1}{a(q)}\exp\left\{i\kappa\int_0^q \frac{q'dq'}{a(q')}\right\} \times\left[\exp\left\{\frac{i(\mathbf{r}-\mathbf{v}t')^2}{2(t'-t)}\right\} + \exp\left\{-\frac{i(\mathbf{r}+\mathbf{v}t')^2}{2(t'-t)}\right\}\right] = \text{const}\{\Psi_- + \Psi_+\}. \quad (23)$$

The integrals defining $\Psi_\pm$ can be transformed to the simpler integrals

$$\Psi_\pm = \frac{\exp\left(\frac{iv^2 t}{2} \mp i\mathbf{v}\cdot\mathbf{r}_\pm\right)}{r_\pm\sqrt{t}} \times\int_L \frac{qdq}{a(q)}\exp\left\{-\frac{i(q-ir_\pm)^2}{2t} + i\kappa\int_0^q \frac{q'dq'}{a(q')}\right\}. \quad (24)$$

Here, $r_\pm = r \pm \mathbf{v}t$, $r_\pm = |r_\pm|$, and

$$a(q) = q - \frac{\exp(-2\nu q)}{2\nu}.$$

The sum $\Psi_+ + \Psi_-$ satisfies the initial equation (16); this can be verified by direct substitution.

Let us now determine more exactly the path of integration L. Let us suppose that $q = q_1 + iq_2$, where q_1 and q_2 are real numbers. Path L consists of the imaginary axis ($q_1 = 0$), where q_2 varies from ∞ to 0, and the real axis ($q_2 = 0$), where q_1 varies from 0 to ∞. On the imaginary axis, the integrand function has no singularities and exponentially decreases for $q_2 \longrightarrow \infty$; on the real axis, a pole exists for $2\nu q_1 = \exp(-2\nu q_1)$ and we assume that contour L bypasses this pole from above. The integrand function in expression (24) decreases exponentially for $q_1 \longrightarrow \infty$ on the real axis as well. It is important for the next section to note that contour L for $q_1 \longrightarrow \infty$ can be displaced to region $q_2 < 0$, which ensures the convergence of the integral.

5. After certain modifications, relations (23) and (24) give the solution of a more complex problem, viz., scattering from different points. We assume that the δ centers are located at points $\mathbf{r} = \pm\mathbf{r}_0$ at $t = 0$. The velocities of the centers are directed along $+\mathbf{r}_0$ and $-\mathbf{r}_0$, respectively. In this case, we must replace $\mathbf{r} - \mathbf{v}t$ by $\mathbf{r} - \mathbf{v}t - \mathbf{r}_0 = \mathbf{r}_-$ and $\mathbf{r} + \mathbf{v}t$ by $\mathbf{r} + \mathbf{v}t + \mathbf{r}_0 = \mathbf{r}_+$ in Eq. (16). We denote

$$Q_q^\pm(\mathbf{r}, t) = \frac{\exp\left(\frac{i\nu^2 t}{2} \mp i\mathbf{v}\cdot\mathbf{r}_\pm\right)}{|\mathbf{r}_\pm|\sqrt{t + \frac{r_0}{\nu}}} \exp\left\{-\frac{i(q - ir_\pm)^2}{2\left(t + \frac{r_0}{\nu}\right)}\right\}.$$

This gives

$$i\frac{\partial Q^\pm}{\partial t} + \frac{1}{2}\Delta Q^\pm = -2\pi\delta(\mathbf{r}_\pm) \times \exp\left\{\frac{i\nu^2 t}{2} - \frac{iq^2}{2\left(t + \frac{r_0}{\nu}\right)}\right\}\frac{1}{\sqrt{t + \frac{r_0}{\nu}}}. \quad (25)$$

We represent the Ψ function in the form

$$\Psi = \int_L (Q_q^+ + Q_q^-)S(q)\,dq. \quad (26)$$

Substituting this expression into Eq. (16) modified in accordance with the above arguments, we obtain the following equation for $S(q)$:

$$\int_L dq \exp\left\{-\frac{iq^2}{2\left(t + \frac{r_0}{\nu}\right)}\right\}S(q) = \frac{1}{\kappa\left(t + \frac{r_0}{\nu}\right)} \times \int_L dq \exp\left\{-\frac{iq^2}{2\left(t + \frac{r_0}{\nu}\right)}\right\}a(q)S(q). \quad (27)$$

This equation has the solution

$$S(q) = \frac{q}{a(q)}C_* \exp\left\{i\kappa\int_0^q \frac{q'dq'}{a(q')}\right\}. \quad (28)$$

Substitution of expression (28) into Eq. (27) and integration of the right-hand side of the resultant equation by parts lead to an identity. The form of Eqs. (26), (27), and (28) implies that relations (23) and (24) give the solution of the problem of δ centers scattered from different points ($\mathbf{r} = \pm\mathbf{r}_0$) if we carry out the following substitution in these relations:

$$t \longrightarrow t + \frac{r_0}{\nu}, \quad r_\pm \longrightarrow |\mathbf{r} \pm (\mathbf{v}t + \mathbf{r}_0)|.$$

Constant C_* depends on ν, κ, and r_0, the complex form of the dependence complicating the passage to the steady-state limit $\nu \longrightarrow 0$. To determine the solution for $\nu \longrightarrow 0$, we will seek $S(q)$ in the form $S(q) = \exp(i\kappa_0 q)$. In this case, we can obtain from Eq. (27)

$$\exp\left[\frac{i\kappa_0}{2}\left(t + \frac{r_0}{\nu}\right)\right] = \frac{\kappa_0}{\kappa}\exp\left[\frac{i\kappa_0^2}{2}\left(t + \frac{r_0}{\nu}\right)\right] - \frac{\kappa_0}{\kappa}\frac{1}{2\nu\left(t + \frac{r_0}{\nu}\right)}\exp\left[\frac{i(\kappa_0 + 2i\nu)^2}{2}\left(t + \frac{r_0}{\nu}\right)\right]. \quad (29)$$

Proceeding to the limit $\nu \longrightarrow 0$, we obtain the equality

$$\kappa = \kappa_0 - \frac{\exp(-2\kappa_0 r_0)}{2r_0},$$

coinciding with relation (10). In accordance with this equality, the steady-state solution is defined by formula (11). It should be noted, however, that the form of function $S(q)$ used here does not provide a solution for $\nu \neq 0$.

6. Solution (23), (24) is basically analogous to the solution obtained in [9] for the bound state of scattered centers in the 1D problem. In both cases, the solution is the sum of two expressions corresponding to the direction of motion of the center and affecting each other. However, the 3D problem exhibits a number of pecu-

liarities. The bound state in this case cannot be represented in the form of a discrete series of "metastable" decaying levels. In the case of scattering from the same point in the 1D problem, in the limit $v \longrightarrow 0$, the state continuously transforms into the state of a single δ center, while no such transition can be made in the 3D problem. This is due to the fact that there is no continuous transition to the state of one center in the steady-state two-point problem upon a limiting transition to zero of the separation between the centers. In monograph [4], the existence of relative motion is simulated by the presence of bound levels with a time-dependent depth. Apparently, it is possible to construct more realistic solvable models of various nonstationary quantum-mechanical processes such as, for example, the tunneling of nucleons or charge exchange of atomic particles. In this connection, nonstationary problems with 3D pointlike δ centers are of special significance.

REFERENCES

1. E. Fermi, Nuovo Cimento **34**, 157 (1934).
2. S. Albeverio, F. Gesztesy, R. Hoegh-Krohn, and H. Holden, *Solvable Models in Quantum Mechanics* (Springer, New York, 1988; Mir, Moscow, 1991).
3. J. M. Roman and R. Tarrach, J. Phys. A: Math. Gen. **29**, 6073 (1996).
4. Yu. N. Demkov and V. N. Ostrovskiĭ, *Zero-Range Potentials and Their Applications in Atomic Physics* (Leningr. Gos. Univ., Leningrad, 1975; Plenum, New York, 1988).
5. A. I. Baz', Ya. B. Zel'dovich, and A. M. Perelomov, *Scattering, Reactions and Decays in Nonrelativistic Quantum Mechanics*, 2nd ed. (Nauka, Moscow, 1971; Israel Program for Scientific Translations, Jerusalem, 1966).
6. G. Breit, Ann. Phys. (N.Y.) **34**, 377 (1965).
7. G. H. Herlihg and Y. Nishida, Ann. Phys. (N.Y.) **34**, 400 (1965).
8. Y. Nishida, Ann. Phys. (N.Y.) **34**, 415 (1965).
9. S. K. Zhdanov and A. S. Chikhachev, Dokl. Akad. Nauk SSSR **218**, 1323 (1974) [Sov. Phys. Dokl. **19**, 696 (1974)].
10. W. Dappen, J. Phys. B **10**, 2399 (1977).
11. S. K. Zhdanov and A. S. Chikhachev, Available from VINITI, No. 2221-74 (Moscow, 1974).
12. A. S. Chikhachev, Zh. Éksp. Teor. Fiz. **107**, 1153 (1995) [JETP **80**, 641 (1995)].
13. V. I. Man'ko and A. S. Chikhachev, Zh. Éksp. Teor. Fiz. **113**, 606 (1998) [JETP **86**, 335 (1998)].
14. G. Scheitler and M. Kleber, Phys. Rev. A **42**, 55 (1990).
15. V. I. Man'ko and A. S. Chikhachev, Yad. Fiz. **64**, 1533 (2001) [Phys. At. Nucl. **64**, 1457 (2001)].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 888–895.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1018–1026.

ATOMS, SPECTRA, RADIATION

Three-Dimensional Rectification of a Gradient Force in a Strong Nonmonochromatic Field

I. V. Krasnov

Institute of Computational Modeling, Siberian Division, Russian Academy of Sciences, Krasnoyarsk, 660036 Russia

e-mail: krasn@icm.krasn.ru

Received October 24, 2003

Abstract—A new three-dimensional scheme for rectifying a gradient force is proposed and analyzed. The scheme is based on the use of a strong, partially coherent optical field involving a component with a fluctuating phase. It is shown that the rectification of a gradient force acting on atoms with a nondegenerate ground state is a second-order effect with respect to field strength in this scheme, whereas an analogous effect is third-order in coherent bichromatic fields. Conditions for three-dimensional confinement of atoms are obtained by using the velocity dependence of the rectified radiative force. For a large class of atoms, such as even-even isotopes of ytterbium and alkaline-earth elements, these conditions can be implemented at a relatively high effective temperature (of the particle ensemble) of about 10 K. This finding can be used to widen substantially the range of energies of atoms amenable to effective three-dimensional optical manipulation. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Resonant atoms placed in a strong nonmonochromatic nonuniform optical field (as a standing wave) are subjected to a strong rectified gradient force (RGF) [1–5]. Its sign remains invariant over a distance much larger than the optical wavelength λ. The force does not saturate with increasing field intensity (in contrast to spontaneous radiation pressure).

Various theoretical aspects and implementation schemes for rectifying radiative forces were considered in numerous studies (e.g., see [6–11] and references cited therein). In particular, their results suggest that RGF can be used to create extremely deep potential wells [12] and provide dissipative optical traps for confining relatively "hot" atoms with energies well above typical lower limits for laser cooling. Their practical implementation can substantially widen the range of energies of atoms amenable to effective three-dimensional optical manipulation. However, optimization of necessary physical conditions must rely on an analysis of three-dimensional models of rectification that allow for polarization phenomena in mechanical effects of light [4].

In this paper, we propose and analyze a new three-dimensional scheme for rectifying a gradient force in a strong nonmonochromatic field involving a component with a fluctuating phase. The analysis is performed for atoms with $J = 0 \longrightarrow J = 1$ transitions (as even–even Yb and alkaline-earth isotopes), which are deemed promising for new experiments on laser cooling (e.g., see 13–16] and references cited therein). In this scheme, the effects due to the RGF and the delayed gradient force (radiative friction) are only of sixth order (!) in the amplitude of the acting field in the limit case of weakly saturated population of excited levels when a coherent field is used [1–4]. For this reason, analysis is complicated and the radiative force has to be modified. The scheme differs from those with atoms with degenerate ground states [8, 17].

We show that rectification of a gradient force in a strong, partially coherent field is a fourth-order effect with respect to the field amplitude (i.e., a second-order one in intensity). We derive expressions for RGF and delayed gradient force (DGF) in a 3D nonmonochromatic field and use them to determine conditions for stable 3D confinement of resonant particles with an effective temperature T of at least several kelvins (much higher than the known lower limits for laser cooling in similar problems).

We note that the opposite limit case of weak coherent bichromatic field and particles with $T \ll 1$ K was considered in previous studies [18, 19] (also devoted to three-dimensional rectification of radiative forces for atoms with strong singlet–singlet transitions and weak $J = 0 \longrightarrow J = 1$ transitions).

2. MODEL

Consider an atom of mass m moving with velocity $\mathbf{v}$ in an electromagnetic field

$$E(\mathbf{r}, t)e^{-i\omega_0 t} + \text{c.c.}$$

with carrier frequency ω_0 tuned to resonance with the $|J_g = 0, M_g = 0\rangle \longrightarrow |J_e = 1, M_e = 0, \pm 1\rangle$ atomic transi-

tion, where J_α is the total angular moment and M_α denotes its projections in the ground ($\alpha = g$) and excited ($\alpha = e$) states.

The field is the superposition of coherent quasi-resonant components with three different frequencies polarized in mutually perpendicular directions and a partially coherent (fluctuating) resonant field $\mathbf{E}'$ with a bandwidth Γ:

$$\mathbf{E}(\mathbf{r}, t) = \sum_{j = x, y, z} E_{j1}\mathbf{e}_j \exp[-i\Delta_j t] + \mathbf{E}'(\mathbf{r}, t), \quad (2.1)$$

where $\mathbf{e}_j$ denotes the unit basis vectors of a Cartesian coordinate system and Δ_j is the detuning from the resonant frequency ω_0.

In accordance with the original concept of gradient-force rectification [1], assume the following hierarchy of characteristic frequencies:

$$|\Delta_j|, \quad |\Delta_j - \Delta_l| \gg |V_{j1}|,$$
$$\Gamma \gg |U_j|, \quad \frac{|V_{j1}|^2}{\Delta_j}, \quad |\delta|, \quad (2.2)$$

$$\frac{|V_{j1}|^2}{\Delta_j} \gg \gamma, \quad kv, \quad \frac{|U_j|^2}{\Gamma}, \quad (2.3)$$

where l and $j \neq l$ denote indices x, y, or z; $V_{j1}(\mathbf{r})$ and $U_j(\mathbf{r}, t)$ are the Rabi frequencies defined as

$$V_{j1} = \frac{dE_{j1}}{\hbar}, \quad U_j = \frac{d(\mathbf{e}_j\mathbf{E}'(\mathbf{r}, t))}{\hbar},$$

with $d = \|d\|/\sqrt{3}$ ($\|d\|$ is the reduced dipole transition matrix element); $k = \omega_0/c$ is the wave number; γ is the decay rate for the excited state; and δ is the fluctuating-component detuning from the resonant frequency. Inequality (2.2) implies that the coherent components of $\mathbf{E}_1$ are "quasi-resonant," i.e., give rise to a spatially nonuniform Stark shift, and the fluctuating component is "resonant," i.e., ensures excitation of the atom.[1] Condition (2.3) means that the coherent field $\mathbf{E}_1$ is sufficiently strong to ensure that the light-induced Stark shifts exceed the optical resonance width. The opposite limit of a weak coherent bichromatic field was considered in [18, 19]. Note that superposition (2.1) a fortiori admits a 3D acting-field configuration (cf. [1–3]).

An atom placed in field (2.1) is driven by the force [4, 20]

$$\mathbf{F} = \hbar\sum_j (\rho_j \nabla \hat{V}_j^* + \text{c.c.}), \quad (2.4)$$

where

$$\hat{V}_j(\mathbf{r}, t) = V_{j1}\exp(-i\Delta_j t) + U_j,$$

and ρ_j denotes the projections of the induced dipole moment measured in d, which are determined by solving the optical Bloch equations written for a prescribed unperturbed classical trajectory $\mathbf{r} = \mathbf{v}t$. In the "Cartesian" representation adapted to the present problem [4, 19, 21], these equations and the expression for the force are averaged over oscillations of frequency Δ_j (cf. [22]) to obtain (using the same notation for averaged quantities)

$$i\left(\frac{d}{dt} + \gamma_\perp - i\hat{\Delta}_i(\mathbf{r})\right)\rho_i = \sum_j q_{ij}U_j,$$
$$j, i = x, y, z,$$
$$i\left(\frac{d}{dt} + \gamma - i\hat{\Delta}_{ij}(\mathbf{r})\right)q_{ij} = -i\gamma\delta_{ij} \quad (2.5)$$
$$+ (\rho_i U_j^* - U_i\rho_j^*) - \delta_{ij}\sum_{l = x, y, z} (\rho_l^* U_l - \text{c.c.}),$$

$$\mathbf{F} = \hbar\sum_j (\rho_j \nabla U_j^* + \text{c.c.}) + \hbar\sum_j q_{jj}\frac{\nabla|V_{j1}|^2}{\Delta_j}, \quad (2.6)$$

where q_{ij} is the population difference between the excited and ground states, q_{ij} (with $i \neq j$) characterize the coherence of excited atomic states,

$$\hat{\Delta}_i(\mathbf{r}) = \frac{2|V_{i1}(\mathbf{r})|^2}{\Delta_i} + \sum_{l \neq i}\frac{|V_{l1}(\mathbf{r})|^2}{\Delta_l}$$

are the effective spatially nonuniform detunings due to light-induced Stark shifts, and

$$\hat{\Delta}_{ij}(\mathbf{r}) = \hat{\Delta}_i(\mathbf{r}) - \hat{\Delta}_j(\mathbf{r}).$$

Next, Eqs. (2.5) and (2.6) are averaged over fluctuations of $\mathbf{E}'$.[2] Bloch equations (2.5) constitute a system of multiplicative linear equations, and the averaging over the ensemble of random processes U_j conditioned on the right-hand inequality in (2.2) can be performed by using the expansions of their solution in terms of $\zeta \ll 1$, which is proportional to the autocorrelation time $\tau_c \sim \Gamma^{-1}$ [25]:

$$|U_j|\tau_c, \quad |\hat{\Delta}_j|\tau_c, \quad kv\tau_c, \quad \gamma\tau_c \leq \zeta \ll 1.$$

By assuming that $\langle\langle U_j\rangle\rangle = 0$ and the $\mathbf{E}'$ components with different polarizations fluctuate independently, i.e.,

$$\langle\langle U_j(\mathbf{r}, t)U_i(\mathbf{r}, t + \tau)\rangle\rangle = \langle\langle U_j(\mathbf{r}, t)U_i^*(\mathbf{r}, t + \tau)\rangle\rangle = 0,$$
$$i \neq j.$$

[1] In the scheme considered in [1], this is achieved by using a "controlling" coherent field component with a small detuning.

[2] In the theory of resonant radiation pressure, radiative forces due to fluctuating fields with finite bandwidths were originally considered in [23, 24].

Eqs. (2.5) and (2.6) are reduced to

$$\bar{\mathbf{F}} = \langle\langle \mathbf{F} \rangle\rangle = \mathbf{F}_g + \mathbf{F}_s,$$
$$\mathbf{F}_g = \sum_i \frac{\nabla |V_{i1}|^2}{\Delta_i} Q_i, \tag{2.7}$$
$$\mathbf{F}_s = -\hbar i$$
$$\times \sum_i Q_i \left(\int_{-\infty}^{0} d\tau \langle\langle \nabla U_i^*(\mathbf{r}, t) U_i(\mathbf{r}, t+\tau) \rangle\rangle - \text{c.c.} \right), \tag{2.8}$$

$$\left[\frac{d}{dt} + \gamma + 2R_i(\mathbf{r}) \right] Q_i + \sum_{l \neq i} R_e(\mathbf{r}) Q_e = -\gamma, \tag{2.9}$$

where $Q_i = \langle\langle q_{ii} \rangle\rangle$ and the rate constants for transitions between the ground and excited atomic states induced by the field $\mathbf{E}'$ are determined by the correlators

$$R_j(\mathbf{r}) = 2\text{Re} \int_{-\infty}^{0} \langle\langle U_j(\mathbf{r}, t) U_j^*(\mathbf{r}, t+\tau) \rangle\rangle d\tau. \tag{2.10}$$

Note also that U_j is treated as a stationary random process and only first-order terms in $\zeta \ll 1$ are retained in the reduced equations.

Equations (2.7)–(2.9) show that, under condition (2.3) of strong quasi-resonant field,

$$|V_{j1}^2/\Delta_{j1}| \gg |U_e|^2/\Gamma,$$

the radiative force $\mathbf{F}_s$ exerted on the atom by the fluctuating field is weak as compared to the gradient force $\mathbf{F}_g$, which is proportional to the sum of the population differences multiplied by the gradients of E_{j1} components:

$$|\mathbf{F}_s| \ll |\mathbf{F}_g|.$$

Accordingly, Eqs. (2.7)–(2.9) expose the roles played by the fields $\mathbf{E}_1$ and $\mathbf{E}'$ in the present model. The fluctuating field $\mathbf{E}'$ is responsible for incoherent mixing of atomic states, and the quasi-resonant coherent field $\mathbf{E}_1$ induces the effective potentials that determine the motion of the atom: the excited atom moves in the field with

$$|V_{i1}(\mathbf{r})|^2/\Delta_{i1}, \quad i = x, y, z;$$

the unexcited one, in the field with

$$-\sum_i |V_{i1}(\mathbf{r})|^2/\Delta_{i1}, \quad i = x, y, z.$$

An analogous model (in the basis of adiabatic states) describes a two-level atom moving in a coherent bichromatic field [3, 22]. It is obvious that a rectified force

$$\mathbf{F}_R = \langle \mathbf{F}_g \rangle = \sum_i \left\langle Q_i \frac{\nabla |V_{i1}|^2}{\Delta_i} \right\rangle \neq 0 \tag{2.11}$$

exists when $R_i = R_i(\mathbf{r})$ (transition rates are spatially modulated), which is possible only if the coherent field has mutually interfering components. (Hereinafter, angle brackets denote averaging over oscillations with periods comparable to the optical wavelength.)

Finally, note that Eqs. (2.9), where the effect of $\mathbf{E}_1$ on transition saturation is ignored, are derived under conditions (2.2) and (2.3) supplemented with a refined quasi-resonance condition for $\mathbf{E}_1$:

$$\left| \frac{V_{j1}}{\Delta_j} \right|^2 \sim g_j^2 \ll \frac{R_j}{\gamma}.$$

This makes it possible to restrict analysis to the first approximation (i.e., Eqs. (2.5)) in averaging the original Bloch equations over oscillations with frequencies comparable to Δ_j (higher order approximations for a related problem were discussed in [22]).

3. RECTIFIED GRADIENT FORCE AND THREE-DIMENSIONAL CONFINEMENT

To obtain expressions for the RGF, we specify the fields $\mathbf{E}'$ and $\mathbf{E}_1$ as superpositions of plane waves ($j = x, y, z$):

$$U_j(\mathbf{r}, t) = \frac{U}{2}$$
$$\times \{ \exp(i\phi_j(t)) [\exp(i\mathbf{k}_{j1} \cdot \mathbf{r}) + \exp(i\mathbf{k}_{j2} \cdot \mathbf{r})] \tag{3.1}$$
$$+ \exp(i\psi_j(t)) [\exp(-i\mathbf{k}_{j1} \cdot \mathbf{r}) + \exp(-i\mathbf{k}_{j2} \cdot \mathbf{r})] \},$$

$$V_{j1}(\mathbf{r}) = \frac{V_j}{2} [\exp(i(\mathbf{q}_{j1} \cdot \mathbf{r}) + \eta_{j1}) + \exp(i(\mathbf{q}_{j2} \cdot \mathbf{r} + \eta_{j2}))], \tag{3.2}$$

where V_j and $\eta_{j\alpha}$ are the amplitudes and initial phases of the coherent field components, and $\phi_j(t)$ and $\psi_j(t)$ are independent fluctuating phases (with delta-correlated zero-mean derivatives), which determine the correlators of $\mathbf{E}'$ components by the relations

$$\langle\langle \exp i[\phi_j(t) - \phi_j(t+\tau)] \rangle\rangle$$
$$= \langle\langle \exp i[\psi_j(t) - \psi_j(\tau+\tau)] \rangle\rangle = \exp(-\Gamma|\tau|), \tag{3.3}$$
$$\langle\langle \exp i[\psi_j(t) - \phi_j(t+\tau)] \rangle\rangle = 0$$

in a model of radiation with phase diffusion [23, 26].

Thus, each Cartesian component of $\mathbf{E}'$ consists of two independent fluctuating components. Their structure implies that $\mathbf{F}_s = 0$ (in approximation (2.8)), and the field $\mathbf{E}'$ has a Lorentzian spectral profile with bandwidth Γ:

$$I(\omega) \propto \frac{2\Gamma}{(\omega - \omega_0)^2 + \Gamma^2}.$$

Note that representation (3.1) in the region occupied

by atoms is valid only if the coherence length

$$l_c = c\tau_c = c/\Gamma$$

is much greater than the diameter of the region:

$$l_c \gg b$$

(see [23]). Moreover, if correlated light beams with wave vectors $\mathbf{k}_{j1}$ and $\mathbf{k}_{j2}$ are obtained from the same source by using an appropriate optical system, then the optical path difference between them must also be much smaller than l_c.

The vectors $\mathbf{k}_{j\alpha}$ and $\mathbf{q}_{j\alpha}$ in (3.1) and (3.2), with the magnitudes

$$|\mathbf{k}_{j\alpha}| = k = \omega_0/c, \quad |\mathbf{q}_{j\alpha}| = q_j = |\omega_0 + \Delta_j|/c$$

lie in the planes perpendicular to the corresponding basis vectors of the Cartesian coordinate system:

$$\mathbf{k}_{j\alpha}\cdot\mathbf{e}_j = \mathbf{q}_{j\alpha}\cdot\mathbf{e}_j = 0.$$

To be specific, suppose that

$$\begin{aligned}\Delta\mathbf{q}_x &= \Delta q\mathbf{e}_y, \quad \Delta\mathbf{q}_y = \Delta q\mathbf{e}_z, \quad \Delta\mathbf{q}_z = \Delta q\mathbf{e}_x,\\ \Delta\mathbf{k}_x &= \Delta k\mathbf{e}_y, \quad \Delta\mathbf{k}_y = \Delta k\mathbf{e}_z, \quad \Delta k_z = \Delta k\mathbf{e}_x,\end{aligned} \tag{3.4}$$

$$|\delta k| \ll |\Delta k|, \quad |\Delta q|, \tag{3.5}$$

where

$$\delta k = \Delta q - \Delta k, \quad \Delta\mathbf{q}_j = [\mathbf{q}_{j2} - \mathbf{q}_{j1}]/2,$$

$$\Delta\mathbf{k}_j = [\mathbf{k}_{j2} - \mathbf{k}_{j1}]/2,$$

and the values of Δ_q, Δk, and δk are determined by prescribing the angles β and β_j between the wave propagation directions in (3.1) and (3.2), i.e., between the pairs $\{\mathbf{k}_{j2}, \mathbf{k}_{j1}\}$ and $\{\mathbf{q}_{j2}, \mathbf{q}_{j1}\}$:

$$\Delta k = k\sin(\beta/2), \quad \Delta q_j = q_j\sin(\beta_j/2).$$

Consequently, the "microscopic" and "macroscopic" length scales, λ_M and Λ ($\lambda_M \ll \Lambda$, see [3, 22]), are estimated as $\lambda_M = \pi/\Delta q \sim \lambda$ and $\Lambda = \pi/\delta k$ in this problem and are parameters that can be adjusted by choosing values β and β_j. The optical field configuration is schematized in the figure.

Expressions for the transition rates $R_j(\mathbf{r})$ and the effective potentials $|V_{j1}(\mathbf{r})|^2/\Delta_j$ are obtained by combining (2.10), (3.1), and (3.2):

$$\begin{aligned}R_j(\mathbf{r}) &= R\cos^2(\Delta\mathbf{k}_j\cdot\mathbf{r}), \quad R = 4|U|^2/\Gamma,\\ \frac{|V_{j1}(\mathbf{r})|^2}{\Delta_j} &= \frac{|V_j|^2}{\Delta_j}\cos^2[\Delta\mathbf{q}_j\cdot\mathbf{r} + \xi_j], \quad \xi_j = [\eta_{j2} - \eta_{j1}]/2.\end{aligned} \tag{3.6}$$

When the transition is not saturated, i.e.,

$$4R/\gamma < 1,$$

the steady-state solution to Eqs. (2.9) (at $t > \gamma^{-1}$) can be represented as a convergent series in powers of the **E'** wave intensity:

$$Q_j = \sum_{j=0}^{\infty} Q_j^{(n)}, \quad Q_j^{(0)} = -1, \tag{3.7}$$

where $Q_j^{(n)}(\mathbf{r})$ are defined by the recursive relations

$$\begin{aligned}Q_j^{(n)}(\mathbf{r}) = -\int_{-\infty}^{0}\Bigg[&2R_j(\mathbf{r}+\mathbf{v}\tau)Q_j^{(n-1)}(\mathbf{r}+\mathbf{v}\tau)\\ &+\sum_{l\neq j}R_l(\mathbf{r}+\mathbf{v}\tau)Q_l^{(n-1)}(\mathbf{r}+\mathbf{v}\tau)\Bigg]e^{\gamma\tau}d\tau.\end{aligned} \tag{3.8}$$

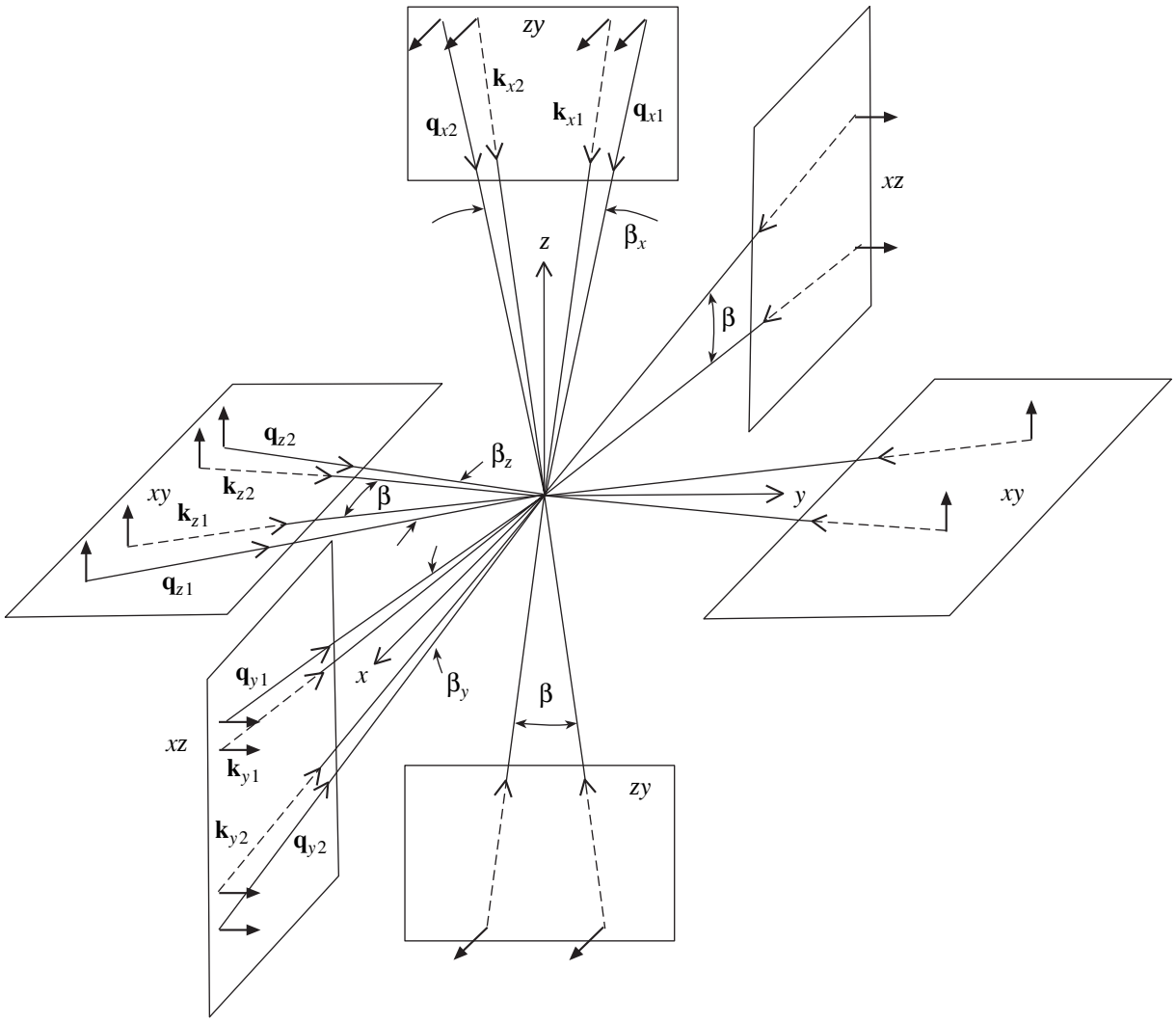
In the linear approximation with respect to the **E'** wave intensity, (3.7) and (3.8) yield an expression for the population difference,

$$\begin{aligned}Q_j(\mathbf{r}) &\approx \frac{\gamma^2 R}{(\gamma^2 + 4(\Delta\mathbf{k}_j\cdot\mathbf{v})^2)}\\ &\times\left[\frac{1}{\gamma}\cos(2\Delta\mathbf{k}_j\cdot\mathbf{r}) + \frac{2\Delta\mathbf{k}_j\cdot\mathbf{v}}{\gamma^2}\sin(2\Delta\mathbf{k}_j\cdot\mathbf{r})\right]\\ &+\sum_{l\neq j}\frac{\gamma^2 R}{2(\gamma^2 + 4(\Delta\mathbf{k}_l\cdot\mathbf{v})^2)}\\ &\times\left[\frac{1}{\gamma}\cos(2\Delta\mathbf{k}_l\cdot\mathbf{r}) + \frac{2\Delta\mathbf{k}_l\cdot\mathbf{v}}{\gamma^2}\sin(2\Delta\mathbf{k}_l\cdot\mathbf{r})\right],\end{aligned}$$

which can be combined with (2.7), (3.4), and (3.5) to find the rectified radiative force (after averaging $\bar{\mathbf{F}}$ over spatial oscillations of period λ_M):

$$\begin{aligned}\mathbf{F}_R &= \langle\bar{\mathbf{F}}\rangle = \sum_i (F_{0i} + F_{1i})\mathbf{e}_i,\\ F_{0i} &= -\frac{\hbar\Delta k\Delta_j}{1 + (v_i/v_c)^2}\frac{R}{2\gamma}g_j^2\sin\Phi_i,\\ \Phi_i &= 2\delta k r_i + 2\xi_j,\\ F_{1i} &= -\frac{m\chi_i v_i}{1 + (v_i/v_c)^2}, \quad \chi_i = \kappa_i\cos\Phi_i,\\ \kappa_i &= \frac{\hbar\Delta k^2 g_j^2 R\Delta_j}{m\gamma^2}.\end{aligned} \tag{3.9}$$

In accordance with (3.4), the pairs of indices (i, j) are



Three-dimensional optical field configuration corresponding to superposition (3.1), (3.2) and satisfying conditions (3.4): long dashed and solid arrows indicate the directions of propagation of partially coherent and coherent waves (with wave vectors $\pm\mathbf{k}_{j1}$, $\pm\mathbf{k}_{j2}$ and $\mathbf{q}_{j1}$, $\mathbf{q}_{j2}$), respectively; short arrows, polarization directions; β and β_j are angular widths.

(x, z), (y, x), or (z, y),

$$g_j^2 = |V_j/\Delta_j|^2,$$

$$r_i = \mathbf{e}_i \cdot \mathbf{r}, \quad r_x = x, \quad r_y = y, \quad r_z = z,$$

$$v_i = \mathbf{e}_i \cdot \mathbf{v}, \quad v_c = \gamma/2\Delta k,$$

$\mathbf{F}_0$ is the rectified gradient force, and $\mathbf{F}_1$ is the delayed gradient force (radiative friction) (by the terminology of [4]).

It is clear from (3.9) that both RGF and DGF are second-order quantities with respect to field strength here, whereas third-order analogous quantities are obtained in coherent bichromatic fields [3, 4]. The velocity dependence of RGF has a Lorentzian profile with a width determined by the "microscopic" length scale:

$$v_c = \gamma\lambda_M/2\pi.$$

When $v_i \gg v_c$, the RGF scales with the inverse square of particle velocity; when $v_i \ll v_c$, it is virtually independent of the velocity. In the latter case, macroscopic potential wells are created, with depths greater than the characteristic depth $\hbar|V_{j1}|^2/\Delta_j$ of microscopic potential wells. Note also that the DGF is a nonlinear function of both velocity and coordinate of the atom.

It is remarkable that the "macroscopic" motions of particles along the axes of the Cartesian coordinate system induced by RGF and DGF are mutually independent. When

$$mv_i^2/2 = T_i \gg \hbar|V_{j1}|^2/\Delta_j,$$

they are governed by the equations

$$m\left[1+\frac{v_i^2}{v_c^2}\right]\frac{dv_i}{dt} = -\frac{\partial\Pi(r_i)}{\partial r_i} - m\kappa v_i\cos\Phi_i, \qquad \frac{dr_i}{dt} = v_i, \quad i = x, y, z, \tag{3.10}$$

i.e., by the Newton equations with a "renormalized" (velocity-dependent) mass, where

$$\Pi(r_i) = \Pi_0(1-\cos\Phi_i), \quad \Pi_0 = \hbar\omega_0\frac{R\Lambda\Delta k}{4\pi ck}\sqrt{\frac{Ig^2}{I_s}},$$

$I_s = \hbar\omega_0\gamma k^2/6\pi$ is the wave intensity that saturates the atomic transition, $I = I_x$, $g^2 = g_x^2$, I_i denotes the intensity of a plane-wave component in superposition (3.2) polarized along $\mathbf{e}_j$, and $\Delta k > 0$. Furthermore, all detunings $\Delta_j > 0$ induced by the RGF also supposed to be similarly distributed along each Cartesian coordinate axis:

$$|V_l|^2/\Delta_l = |V_i|^2/\Delta_i.$$

Therefore,

$$I_l g_l^2 = I_i g_i^2 = Ig^2$$

for every pair of indices l and i, and

$$\kappa_i \equiv \kappa.$$

Thus, the model of three-dimensional confinement is reformulated as a nonlinear model of one-dimensional motion. Under the conditions $\Delta_j > 0$, the minima of the potential $\Pi(r)$ are found by solving the equation

$$\cos\Phi_i(r_m) = 1.$$

It is clear that each point A_m with phase-space coordinates ($r_i = r_m$, $v_i = 0$) is a stable stationary point (attractor) of system (3.10). However, a particle moving in the vicinity of the RGF node located at r_m is confined in its region of attraction G_m only if its kinetic energy $mv_i^2/2 = T_i$ does not exceed a certain critical T_k determined not only by the potential-barrier height $2\Pi_0$, but also by the profile width of the RGF as a function of velocity. If $mv_c^2/2 = T_c \ll \Pi_0$, then $T_k \ll \Pi_0$ (since the RGF rapidly decreases at $v_i \gg v_c$). If $T_c \gg \Pi_0$, then T_k is comparable to Π_0, but is substantially lower than T_c. Both G_m and T_k are difficult to determine because the sign of the friction coefficient depends on the particle's location.

Let us find sufficient conditions for three-dimensional confinement of atoms and estimate T_k, using the fact that DGF plays the role of friction only in the regions Ω_m where

$$\kappa\cos\Phi_i > 0, \tag{3.11}$$

i.e., when

$$r_i \in (r_m - \pi/4\delta k, r_m + \pi/4\delta k).$$

Define the generalized energy

$$\mathscr{E}(r_i, v_i) = T_i(1 + T_i/2T_c) + \Pi(r_i) \tag{3.12}$$

(when $\kappa = 0$, $\mathscr{E}$ is an integral of motion). Alternatively, $\mathscr{E}$ is interpreted as the Lyapunov function of system (3.10) in the phase-space domain N_m bounded by the closed contour $\mathscr{E}(r_i, v_i) = \Pi_0$ encompassing the attractor A_m. Indeed, (3.10) implies that its derivative along the trajectory $r_i = r_i(t)$, $v_i = v_i(t)$ almost everywhere in N_m (except for A_m) satisfies the differential inequality

$$\frac{d\mathscr{E}(r_i, v_i)}{dt} = -2\kappa T_i\cos\Phi_i < 0, \tag{3.13}$$

because the condition for particle confinement inside N_m,

$$\mathscr{E}(r_i, v_i) < \Pi_0, \tag{3.14}$$

entails (3.11) and, therefore, $r_i \in \Omega_m$. Note that the function $\mathscr{E}$ is positive definite everywhere in N_m except for A_m (where $\mathscr{E} = 0$).

Thus, every trajectory passing through N_m asymptotically approaches the point A_m as $t \longrightarrow \infty$, crossing the closed contours of constant $\mathscr{E}$ inwards, and inequality (3.14) is a sufficient condition for confinement of atoms at the nodes of RGF. Note that, even though $N_m \in G_m$ (i.e., N_m is just a subregion of the region of attraction of A_m, as shown numerically), the availability of analytical representation (3.14) facilitates analysis of the confinement conditions.

Condition (3.14) entails a constraint on the kinetic energy of particles and an estimate for T_k:

$$T_i < T_M = \frac{2\Pi_0}{1+\sqrt{1+2\Pi_0/T_c}} < T_k.$$

This means that an atom that passes through the RGF node located at r_m and has an energy not higher than T_M will be trapped in its vicinity. On the other hand, an atom with energy $T_i \ll T_M$ confined in a small neighborhood of an RGF node cannot be released from the region of attraction by a sudden perturbation (e.g., by a single collision with a "hot" particle) if the resulting increase in its energy is not greater than T_M.

Since

$$\Pi_0 \propto \Delta k/k, \quad \Pi_0/T_c \propto (\Delta k/k)^3,$$

the quantity $T_M = T_M(\Delta k)$ as a function of the parameter $\Delta k = \pi/\lambda_M$ reaches a maximum value T_M', which can be

expressed in a form suitable for estimation:

$$\max_{\Delta k} T_M = T'_M \approx \frac{T_0}{3.2}\left[\frac{\hbar\omega_0 Rg}{2\pi T_0 c}\right]^{2/3}\left[\frac{I\Lambda^2}{I_s}\right]^{1/3}, \quad (3.15)$$

where $T_0 = m\gamma^2/2k^2$.

The maximum is reached when

$$\Delta k/k = \sqrt{T_0/2T'_M} < 1$$

and is associated with a specific relation between the profile width of the RGF as a function of velocity and its magnitude:

$$\Pi_0 = 4T_c.$$

By solving Eq. (3.10) numerically, a simple relation is found:

$$T_k/T'_M = \eta \approx 2.07.$$

Note that the value of T_k is lower than $\max\Pi(r_i) = 2\Pi_0$ approximately by half in the optimal regime considered here.

Thus, under an optimal choice of the field configuration, the RGF can be used to confine particles in three-dimensional traps of size smaller than Λ if their effective temperature satisfies the condition

$$T_{\text{eff}} < T = 2\eta T'_M/k_B,$$

where k_B is the Boltzmann constant, T'_M is defined by (3.15), and $\eta \sim 2$.

It is important that T increases with the coherent-field intensity even when both g and R/γ are held constant. As an example consider an ytterbium atom with the $^1S_0$–$^1P_1$ singlet–singlet transition ($\lambda = 398.8$ nm, $\gamma = 1.8 \times 10^8$ s^{-1}). If $R/\gamma \approx 0.2$ and $g^2 \approx 0.05$ are taken as estimated values, then (3.15) yields a simple expression for the limit temperature (in kelvins) for atoms confined by means of the RGF:

$$T \approx 2\left[\frac{I\Lambda^2}{I_s}\right]^{1/3},$$

where I and I_s are measured in W/cm^2 and Λ in centimeters. In particular, if $\Lambda \approx 0.5$ cm, $I/I_s = 10^3$, and $\Lambda^2 I \approx 25$ W, then $T \approx 12$ K. In this case, $\Delta k/k \sim 0.38$, $\Delta_j \gtrsim 2 \times 10^{10}$ s^{-1}, and all starting conditions of the problem are satisfied if the fluctuating-field intensity is $I' \approx 5I_s$ and its bandwidth is $\Gamma \approx 5 \times 10^9$ s^{-1}. For comparison, note that $T_0 \approx 1.5$ K in the example considered here, whereas the lower temperature limit for confined atoms corresponding to quantum fluctuations of radiative forces does not exceed

$$T_1 \approx \hbar|V_{i1}|^2/\Delta_i \approx 0.01 \text{ K}.$$

4. CONCLUSIONS

The scheme of the ponderomotive effect of a strong, partially coherent field on atoms with a $J = 0 \longrightarrow J = 1$ transition analyzed here is remarkable in two respects. First, both rectified gradient force and friction force are second-order quantities with respect to the field intensity. Second, the light-induced motion of a particle is (on a macroscopic scale) a superposition of independent one-dimensional motions along three mutually orthogonal axes. Each of these motions is controlled only by field components having a certain polarization in the plane perpendicular to the direction of motion.[3] This finding can be used to simplify optical control of three-dimensional particle motion by independently varying the parameters and geometry of field components with mutually orthogonal polarizations.

In principle, the proposed scheme for rectifying the gradient force makes it possible to implement three-dimensional confinement of relatively "hot" particles with temperatures as high as several kelvins under an optimal choice of the optical field geometry and parameters. In particular, deep traps of this kind may help to solve the challenging problem of optical trapping of an ultracold electron–ion plasma with ions in resonance with laser light, because its electron subsystem may have a relatively high temperature of 1 to 10 K (even when its density is low) [27].

ACKNOWLEDGMENTS

This work was supported by the Krasnoyarsk Science Foundation, grant no. 11F0036C.

REFERENCES

1. A. P. Kazantsev and I. V. Krasnov, Pis'ma Zh. Éksp. Teor. Fiz. **46**, 264 (1987) [JETP Lett. **46**, 332 (1987)].
2. A. P. Kazantsev and I. V. Krasnov, Zh. Éksp. Teor. Fiz. **95**, 104 (1989) [Sov. Phys. JETP **68**, 59 (1989)].
3. A. P. Kazantsev and I. V. Krasnov, J. Opt. Soc. Am. B **6**, 2140 (1989).
4. A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *The Mechanical Action of Light on Atoms* (Nauka, Moscow, 1991).
5. R. Grimm, Y. B. Ovchinnikov, A. I. Sidorov, and V. S. Letokchov, Phys. Rev. Lett. **65**, 1415 (1990).
6. V. S. Voĭtsekhovich, M. V. Danileĭko, A. M. Negriĭko, *et al.*, Zh. Éksp. Teor. Fiz. **99**, 393 (1991) [Sov. Phys. JETP **72**, 219 (1991)].
7. V. S. Letokhov, M. A. Ol'shanii, and Yu. B. Ovchinnikov, Quantum Semiclassic. Opt. **7**, 5 (1995).
8. R. Grimm, J. Soding, and Y. B. Ovchinnikov, Pis'ma Zh. Éksp. Teor. Fiz. **61**, 362 (1995) [JETP Lett. **61**, 367 (1995)].
9. I. V. Krasnov, Zh. Éksp. Teor. Fiz. **107**, 1135 (1995) [JETP **80**, 632 (1995)].

[3] This behavior has never been observed in coherent bichromatic fields [18, 19].

10. R. Grimm, G. Wasik, J. Soding, and Yu. B. Ovchinnikov, in *Proceedings of International School of Physics "Enrico Fermi"* (IOS Press, Amsterdam, 1996), Course CXXXI, p. 481.
11. A. Gorlitz, T. Kinoshita, T. W. Hansh, and A. Hemmerich, Phys. Rev. A **64**, 011401(R) (2001).
12. A. P. Kazantsev and I. V. Krasnov, Phys. Lett. A **127**, 33 (1988).
13. K. Honda, Y. Takahasi, T. Kuwamoto, *et al.*, Phys. Rev. A **59**, R934 (1999).
14. T. Kuwamoto, K. Honda, Y. Takahasi, and T. Yabuzaki, Phys. Rev. A **60**, R745 (1999).
15. J. Grunert and A. Hemmerich, Phys. Rev. A **65**, 041401(R) (2002).
16. E. A. Curtis, C. W. Oates, and L. Holberg, Phys. Rev. A **64**, 031403(R) (2001).
17. A. I. Sidorov, R. Grimm, and V. S. Letokhov, J. Phys. B: At. Mol. Opt. Phys. **24**, 3733 (1991).
18. I. V. Krasnov and S. P. Polyutov, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 328 (2002) [JETP Lett. **76**, 270 (2002)].
19. S. A. Gavrilyuk, I. V. Krasnov, and S. P. Polyutov, Zh. Éksp. Teor. Fiz. **120**, 1135 (2001) [JETP **93**, 985 (2001)].
20. V. G. Minogin and V. S. Letokhov, *The Pressure of Laser Radiation on Atoms* (Nauka, Moscow, 1986).
21. A. P. Kazantsev, V. S. Smirnov, G. I. Surdutovich, *et al.*, J. Opt. Soc. Am. B **2**, 1731 (1985).
22. I. V. Krasnov, Laser Phys. **4**, 906 (1994).
23. R. G. Cook, Phys. Rev. A **21**, 268 (1980).
24. V. G. Minogin, Opt. Spektrosk. **53**, 125 (1982) [Opt. Spectrosc. **53**, 73 (1982)].
25. N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 2nd ed. (North-Holland, Amsterdam, 1992; Vysshaya Shkola, Moscow, 1990).
26. S. Stenholm, *Foundations of Laser Spectroscopy* (Wiley, New York, 1984; Mir, Moscow, 1987).
27. A. P. Gavrilyuk, I. V. Krasnov, and N. Ya. Shaparev, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 497 (2002) [JETP Lett. **76**, 423 (2002)].

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 896–907.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1027–1040.

ATOMS, SPECTRA, RADIATION

Theory of Periodic States for Feedback-Controlled Photorefractive Nonlinear Systems

E. V. Podivilov[a], B. I. Sturman[a], and M. V. Gorkunov[b,1]

[a] *International Institute for Nonlinear Studies, Institute of Automation and Electrometry, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia*
[b] *Universität Osnabrück, D-49069 Osnabrück, Germany*
e-mail: sturman@iae.nsk.su
Received October 24, 2003

Abstract—Certain feedback loops can be used in photorefractive optical schemes to implement periodic states and create spatial gratings characterized by extremely high or low diffraction efficiencies. This highly nonlinear phenomenon is studied both experimentally and numerically. An analytical method is developed for analyzing periodic states with the use of symmetries of time-dependent diffraction equations and fast feedback response. The method is applied to describe the properties of periodic states, including their spatial structure, diffraction-efficiency oscillation period and amplitude, and characteristics of feedback-controlled strong phase modulation. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Photorefractive beam coupling has been the subject of intensive theoretical and experimental studies over the last twenty years [1–3]. This is explained by the diversity and strength of photorefractive nonlinear effects and by their potential applicability in various optical devices [2, 4].

Photorefractive nonlinearity is caused by photoinduced charge separation and change in optical permittivity due to the linear electro-optic effect [3]. Gain saturation (at 10^1–10^2 cm^{-1}) is frequently observed even for intensities of about a milliwatt, when the photoinduced contribution begins to play a dominant role in conductivity. The nonlinear response time decreases with increasing intensity, varying between 10^2 and 10^{-3} s in experiments with continuous radiation.

The most widely known photorefractive nonlinear effects include optical gain [5], phase conjugation [6], optical emission [7, 8], soliton propagation [9], recording and fixing of volume gratings [10], and nonlinear scattering [11]. In recent years, particular interest is taken in combined effects of optical and material nonlinearities, including spatial subharmonic generation [12], critical enhancement of nonlinear response [13], and space-charge singularities [14].

Control of nonlinear optical properties by means of electronic feedback has also become increasingly important in the past few years [15–20]. A basic experiment is schematized in Fig. 1. The phase φ_s of the input signal wave is coupled to the corresponding output intensity by a certain relation (see details below). This leads to a drastic change in characteristics of two-beam coupling.

The historical background is as follows. Initially, feedback was introduced to eliminate effects due to fluctuations of input phase difference, i.e., fluctuations of the optical interference pattern in a nonlinear medium. However, experiments demonstrated that feedback effects by far exceed all expectations. In particular, feedback resulted in 100% diffraction efficiency of the light-induced refractive-index grating and a drastic change in the energy exchange between the beams [15–17]. The first attempts to explain the observed effects by assuming that the induced refractive-index

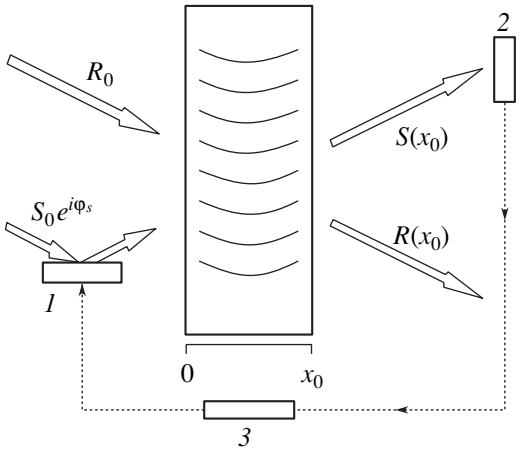


Fig. 1. Experiment on feedback-controlled two-beam coupling: *1*—piezomounted mirror, *2*—photodetector, *3*—feedback control system. Curved line segments represent grating fringes.

[1] Permanent address: Institute of Crystallography, Russian Academy of Sciences, Moscow, 117333 Russia.

grating is spatially uniform have failed. However, these attempts involved the correct conjecture that feedback tends to induce a phase shift of $\pm\pi/2$ between the diffracted and transmitted components of the output signal (S) wave.

Equations that adequately describe beam coupling controlled by $\pm\pi/2$ feedback, including nonlocal boundary conditions, were formulated in [18], where a direct numerical simulation was also reported. It was found that the perfect $\pm\pi/2$ feedback did produce a clear-cut maximizing effect. Under reasonable assumptions about the parameters of a nonlinear crystal, the feedback induced a refractive-index grating characterized by either 100% or zero diffraction efficiency. However, perfect feedback is not beneficial when these extreme states are obtained. In other words, the formulated ideal model remained correct only for a limited time interval beyond which the formulated equations failed to describe a continuously operating real device.

Subsequently, it was established that feedback delay is responsible for continuous operation and is substantial only near extreme states [19]. Numerical simulations showed that, because of feedback delay, a nonlinear system reaches a periodic state (attractor) in which the phase $\varphi_s(t)$ of the input signal wave exhibits fast large-amplitude oscillations combined with a slow drift [19, 20]. This behavior carries over to output beam characteristics. Moreover, numerical simulations revealed that variations of crystal thickness and input-wave intensity ratio induce transitions between periodic states with different characteristics and periods of $\varphi_s(t)$ oscillations. Specially designed experiments confirmed that multiple periodic states exist [19].

It should be noted that, while the nonlinear effects discussed here are obviously of importance for experiment and applications, they present a nontrivial problem. As far as we know, no close analogy can be found in nonlinear physics. This is special because nonlinear evolution equations for waves inside a crystal are supplemented with a nonlocal nonlinear feedback condition. The results of direct numerical simulations stimulate theoretical analysis rather than provide alternative solutions.

In this paper, we develop an analytical method to describe beam coupling controlled by $\pm\pi/2$ feedback and demonstrate that it can be effectively used to obtain various results. The key point in the theory is the averaging over fast phase oscillations [21]. This procedure makes use of the only available small parameter, viz., the ratio of the feedback response time to the nonlinear response time, and some general symmetries of the nonlinear wave equations.

The theory developed here is restricted to an analysis of local nonlinear response, which is both easiest to examine and most important for experiment. However, this restriction is not essential. The results obtained can be extended to other photorefractive nonlinearities of practical interest (e.g., see [1–3]). From a more general perspective, these results are of interest for analysis of media characterized by nonlinear response times, such as liquid crystals [22]. Owing to nonlinear response, feedback loops can be used to perform various optical functions. Maximization and minimization of diffraction efficiency of spatial gratings can be viewed as examples of such functions.

2. BASIC RELATIONS

2.1. Two-Beam Coupling Equations

The starting system of equations for the wave amplitudes S and R and the amplitude E of the field generated by space charge (see Fig. 1) is written in dimensionless form as

$$R_x = iES, \tag{1}$$

$$S_x = iE^*R, \tag{2}$$

$$E_t + E = RS^*. \tag{3}$$

The subscripts x and t denote derivatives with respect to dimensionless coordinate and time, respectively. The coordinate x is normalized to the nonlinearity length scale (independent of light intensity), and t is normalized to the nonlinear response time (inversely proportional to the total wave intensity).

Equations (1) and (2) are derived from Maxwell's equations to describe Bragg diffraction of the R- and S-waves by a refractive-index grating. They do not contain time derivatives because of slow nonlinear response (light intensity is fully controlled by variation of $E(t)$). The total wave intensity is conserved: $(|S|^2 + |R|^2)_x = 0$. Since the input intensities (at $x = 0$) are independent of time in the problem considered here, the total intensity is independent of both x and t. Therefore, R and S can be normalized so that

$$|R(x, t)|^2 + |S(x, t)|^2 = 1.$$

Equation (3) describes the slow development of a light-induced grating and reflects the material properties of the crystal. The absolute value of the product RS^* on the right-hand side of (3) is equal to half the contrast of the interference pattern. Note that the material equation used here is obtained in various microscopic models of photorefractive response [3, 18].

According to (3), $E = RS^*$ under steady-state conditions; i.e., the maxima of the static refractive-index grating coincide with the maxima of the interference pattern. This property is the locality of nonlinear response assumed here. In the general case, these periodic distributions differ by a shift in space [3]. However, the theory becomes too cumbersome when it is taken into account, and allowance for the shift is not essential for actual experiments.

2.2. Fundamental Solutions

The spatial distributions of the amplitudes R and S are determined by their input values $R(0, t)$ and $S(0, t)$ and by the grating-amplitude profile $E(x, t)$. These factors characterize readout and recording conditions, respectively. In other words, the same grating recorded by an instant t can be read by different methods when the input values are rapidly varied. This property reflects the structure of Eqs. (1)–(3): the first two equations contain time as a parameter in $E(x, t)$, and the last one ensures weak variation of E over $t \ll 1$.

It is important for further analysis that the recording and readout processes are separated as much as possible. To do this, we first consider Eqs. (1) and (2). These equations are linear and homogeneous in R and S. Their common solution (for prescribed $E(x)$) can be represented as a linear combination of two independent particular solutions constituting a basis.

One of the basis vectors is defined as the pair $R_F(x)$, $S_F(x)$ corresponding to the input amplitudes $R_F(0) = 1$ and $S_F(0) = 0$. This fundamental solution corresponds to the process of reading the grating with an R-beam of unit amplitude. Since the grating amplitude varies with time, the fundamental amplitudes are parameterized by t. Next, it is readily verified that the pair $R = -S_F^*(x)$, $S = R_F^*(x)$ also solves Eqs. (1) and (2). The latter particular solution corresponds to the process of reading the grating with an S-beam of unit amplitude. It is linearly independent of the former one and can be chosen as the second basis vector. The existence of a basis consisting of only two fundamental amplitudes, $R_F(x)$ and $S_F(x)$, reflects the symmetry properties of system (1), (2).

Thus, the solution $R(x, t)$, $S(x, t)$ subject to the input conditions $R_0(t)$, $S_0(t)$ can be represented as

$$\begin{aligned} R(x, t) &= R_F(x, t)R_0(t) - S_F^*(x, t)S_0(t), \\ S(x, t) &= S_F(x, t)R_0(t) + R_F^*(x, t)S_0(t). \end{aligned} \tag{4}$$

The fundamental amplitudes $R_F(x, t)$ and $S_F(x, t)$ completely characterize the diffraction properties of the grating. Equations (4) can be solved algebraically for R_F and S_F. In other words, the fundamental solutions can easily be calculated when the amplitudes $R_F(x, t)$ and $S_F(x, t)$ corresponding to a readout process are known and it is kept in mind that $|R_F(x, t)|^2 + |S_F(t, x)|^2 = 1$.

A necessary step in formulating feedback conditions is the expansion of R and S into transmitted (T) and diffracted (D) components. According to Eqs. (4) (see also Fig. 1), the contribution $R_F R_0$ to the amplitude R is precisely the transmitted part of the input R-beam, while $-S_F^* S_0$ is the diffracted part of the input S-beam. Similarly, $S_F R_0$ and $R_F^* S_0$ are the D and T components of $S(x, t)$.

Finally, we note that

$$\eta = |S_F(x_0, t)|^2 \equiv 1 - |R_F(x_0, t)|^2$$

is an important observable characteristic (diffraction efficiency) of a grating. It can be determined experimentally by shutting off either R- or S-beam momentarily and calculating the ratio $|R(x_0, t)|^2/|S(0, t)|^2$ or $|S(x_0, t)|^2/|R(0, t)|^2$, respectively.

2.3. Feedback Conditions

Hereinafter, we assume that the input R-wave amplitude is constant and represent the input S-wave amplitude as $S(0, t) = S_0 \exp(i\varphi_s)$, where $S_0 = \text{const}$ and $\varphi_s = \varphi_s(t)$ is the feedback-controlled phase. The time dependence of this phase determines the observed properties of the system. The dimensionless crystal thickness x_0 and the input intensity ratio $r_0 = |R_0|^2/|S_0|^2$ are used as adjustable parameters.

The phase shift between the diffracted and transmitted components of S, $S_F(x_0)R_0$ and $R_F^*(x_0)S_0 \exp(i\varphi_s)$, at the output end of the crystal is

$$\Phi_s = \arg[R_0 S_0^* R_F(x_0) S_F(x_0) \exp(-i\varphi_s)]. \tag{5}$$

The conditions $\Phi_s = \pm\pi/2$ are called perfect feedback conditions. They can be satisfied by an appropriate choice of the input phase φ_s. Since the fundamental amplitudes $R_F(x_0)$ and $S_F(x_0)$ can be expressed as algebraic functions of $R(x_0)$ and $S(x_0)$, the conditions $\Phi_s = \pm\pi/2$ can be interpreted as a nonlinear relation between φ_s and the output amplitudes of recording waves. The quantity Φ_s is well defined only if $\eta \neq 0, 1$ ($|S_F(x_0)R_F(x_0)| \equiv \sqrt{\eta(1-\eta)}$ does not vanish).

In experiments, the feedback is implemented by using a fast and weak input-phase modulation $\delta\varphi_s = \psi_d \sin(\omega t)$, where $\psi_d \ll 1$ and $\omega \gg 1$. Such oscillations of φ_s do not affect the recording process and serve as a marker of the T component of the signal wave. Owing to interference of the D and T components, the output intensity $|S(x_0)|^2$ has a component that oscillates with double frequency as $\cos(2\omega t)$ and amplitude

$$I_{2\omega} = \frac{1}{2}|R_0 S_0| \sqrt{\eta(1-\eta)}\,\psi_d^2 \cos\Phi_s.$$

Using $\pm I_{2\omega}$ as an error signal, one can adjust the input phase so that $\Phi_s = \pm\pi/2$. This method is effective so long as the product $\eta(1 - \eta)$ is not too close to zero. It is important that the present characterization of the feedback relies only on general properties of the fundamental amplitudes R_F and S_F, while the refractive-index grating may be spatially nonuniform.

It was shown experimentally that the $+\pi/2$ feedback can be used (under certain constraints on x_0 and r_0) to

attain η as high as $\eta \approx 1$ and hold it at this level [15]. When the feedback sign is reversed, the diffraction efficiency falls to zero. On the other hand, a numerical analysis showed [18] that the diffraction efficiency reaches a value of 1 or 0 in a time $t \approx 1$. After that, the use of the conditions $\Phi_s = \pm\pi/2$ is not beneficial. Thus, the perfect feedback conditions should be applied to describe the initial evolution of the system and should be modified when its continuous operation is to be described.

It was supposed in [19] that inertia of the feedback loop ensures continuous operation of the system. To make use of the inertia, the perfect conditions $\Phi_s = \pm\pi/2$ were replaced with the time-dependent relation

$$\dot{\varphi}_s = \mp t_f^{-1}|R_0 S_0|\sqrt{\eta(1-\eta)}\cos\Phi_s, \tag{6}$$

where $t_f \ll 1$ is the feedback-loop response time and a dot denotes an ordinary time derivative. This relation admits a simple interpretation. So long as the coefficient of $\cos\Phi_s$ is large as compared to unity, i.e., $\eta(1-\eta)$ is not too close to zero, the input phase φ_s rapidly approaches a value ensuring that $\Phi_s \approx \text{const} - \varphi_s \approx \pm\pi/2$, and relation (6) is almost equivalent to the perfect feedback conditions. As the value of $\eta(1-\eta)$ approaches zero, inertia plays an increasingly important role and the value of Φ_s may deviate substantially from $\pm\pi/2$.

According to numerical simulations [19, 20], the use of (6) instead of the perfect conditions $\Phi_s = \pm\pi/2$ ensures continuous operation of the system. After an extremal is reached, the diffraction efficiency oscillates about either unity or zero, while the phase factor $\exp(i\varphi_s)$ exhibits a fast periodic variation superimposed on a linear drift,

$$\varphi_s = \Omega t + \varphi_p(t). \tag{7}$$

The part of Ω that is a multiple of $2\pi/T$ can be subsumed under the phase factor $\exp(i\varphi_p)$. The corresponding periodic oscillation is neither small nor harmonic, and its period T decreases with t_f. Furthermore, computations showed that the periodic state (attractor) changes qualitatively as x_0 and r_0 are varied. The qualitative changes are observed in detuning Ω, period T, and shape of the "periodic" component $\varphi_p(t)$.

The most important theoretical predictions were confirmed in special experiments conducted to measure $\varphi_s(t)$ [19]. It was also shown that the frequency response of the feedback is in good agreement with Eq. (6) [20]. Thus, substantial evidence was obtained in favor of the inertial feedback condition. The feedback-loop response time was estimated as $t_f \sim 10^{-3}$.

3. ANALYTICAL METHOD

The presence of fast oscillations (with period $T \ll 1$) makes it possible to develop a perturbation theory by using an averaging method. In view of expression (7) for the input phase φ_s, we write

$$R_F(x,t) = \bar{R}_F(x) + \tilde{R}_F(x,t), \tag{8}$$

$$S_F(x,t) = \exp(i\Omega t)[\bar{S}_F(x) + \tilde{S}_F(x,t)], \tag{9}$$

$$E(x,t) = \exp(-i\Omega t)[\bar{E}(x) + \tilde{E}(x,t)], \tag{10}$$

where the amplitudes $\bar{R}_F$, $\bar{S}_F$, and $\bar{E}$ are independent of time and $\tilde{E}$, $\tilde{R}$, and $\tilde{S}$ are T-periodic functions of time. The recording-wave amplitudes R and S can be represented in similar form. The frequency shift should be attributed to slow variations ($|\Omega| \lesssim 1$).

Because of the recording inertia, the fast component of the grating amplitude is small: $\tilde{E} \ll \bar{E}$. Since the fundamental amplitudes R_F and S_F characterize instantaneous diffraction by the grating, it also holds that $\tilde{R}_F \ll \bar{R}_F$ and $\tilde{S}_F \ll \bar{S}_F$. However, the fast components of the recording waves, $\tilde{R}$ and $\tilde{S}$, are not small as compared to the slow ones, $\bar{R}$ and $\bar{S}$, because fast oscillations of the input amplitude $S(0, t) = S_0\exp[i\varphi_s(t)]$ are instantly transferred into the crystal by diffraction on a static (or quasistatic) grating. These fast components contribute substantially to $\bar{E}$, because the product $\tilde{R}\tilde{S}^*$ appearing on the right-hand side of (3) has a constant component that is not small.

As a first step in developing a perturbation theory, we derive a closed system of equations for the slow amplitudes. Using (4) to express the amplitudes R and S^* in terms of R_F and S_F, substituting the results into (3), neglecting the fast components of the fundamental amplitudes, and retaining their slow components, we obtain

$$\begin{aligned}\bar{E} &= (1 - i\Omega)^{-1} \\ &\times [W_0\bar{R}_F\bar{S}_F^* + \varepsilon|R_0 S_0|(\bar{R}_F^2 - \bar{S}_F^{*2})],\end{aligned} \tag{11}$$

where

$$W_0 = |R_0|^2 - |S_0|^2 \equiv \frac{r_0 - 1}{r_0 + 1}$$

is the input intensity difference and $\varepsilon = \langle\exp(i\varphi_p(t))\rangle$ is obtained by averaging over T. Since φ_p is defined up to a constant, we can choose a suitable value of $\arg\varepsilon$. Let us choose $\varepsilon = |\varepsilon|$. The admissible values of the modulation parameter ε lie between zero and unity. Its lowest value corresponds to an extremely strong effect of periodic phase modulation; the highest one, to the absence of modulation.

The differential equations for $\bar{R}_F$ and $\bar{S}_F$ differ from Eqs. (1) and (2) for R and S only by the replacement of E with $\bar{E}$. They can be solved simultaneously with (11) (see next section). As a result, we obtain expressions for

$\bar{E}$, $\bar{R}_F$, and $\bar{S}_F$ as functions of x parameterized by arbitrary x_0, r_0, Ω, and ε. The limit case of $\varepsilon = 1$ corresponds to time-independent beam coupling in the absence of periodic modulation and in the presence of detuning Ω.

The importance of relations for averaged amplitudes is explained as follows. The parameters Ω and ε can be adjusted to satisfy the complex relation $\bar{R}_F(x_0) = 0$ or $\bar{S}_F(x_0) = 0$, i.e., to implement a periodic state with $\eta = 1$ or 0, respectively. This cannot be done without periodic modulation. Thus, the first-order perturbation theory can be used to find the required averaged feedback characteristics, but it does not provide a complete description of periodic states. This can be done in the next order of perturbation theory. It is also important that the slow amplitude $\bar{E}$ corresponding to $\eta = 1$ or 0 strongly depends on x. This property rules out the use of a spatially uniform model of the grating and plays an important role in substantiating the theory.

We schematically describe the calculations performed in second-order perturbation theory. Equations (3) and (8)–(10) are used to derive an equation for the fast grating amplitude:

$$\begin{aligned}\tilde{E}_t &= |R_0 S_0|[(\bar{R}_F^2 - \bar{S}_F^{*2})(\cos\varphi_p - \varepsilon)\\ &\quad - i\sin\varphi_p(\bar{R}_F^2 + \bar{S}_F^{*2})].\end{aligned} \tag{12}$$

Here, we neglect the terms that are small with respect to T and make use of $\langle\cos\varphi_p\rangle = \varepsilon$ and $\langle\sin\varphi_p\rangle = 0$. On the right-hand side of (12) (which is interpreted as a fast-oscillating driving force), the space- and time-dependent functions are completely uncoupled, and this equation is easily solved.

The fast components of the fundamental amplitudes obey an inhomogeneous linear system of equations, which is derived from (1) and (2) and solved by variation of constants. As a result, we have

$$\begin{aligned}\tilde{R}_F(x, t) &= A(x, t)\bar{R}_F(x) - B(x, t)\bar{S}_F^*(x),\\ \tilde{S}_F(x, t) &= A(x, t)\bar{S}_F(x) + B(x, t)\bar{R}_F^*(x),\end{aligned} \tag{13}$$

where

$$\begin{aligned}A(x, t) &= 2i\int_0^x \mathrm{Re}[\tilde{E}(x', t)\bar{R}_F^*(x')\bar{S}_F(x')]dx',\\ B(x, t) &= i\int_0^x [\tilde{E}^*(x', t)\bar{R}_F^2(x')\\ &\quad - \tilde{E}(x, t)\bar{S}_F^2(x')]dx'.\end{aligned} \tag{14}$$

When the output $\tilde{R}_F(x_0, t)$ and $\tilde{S}_F(x_0, t)$ are known, the feedback conditions can be used to find the function $\varphi_p(t)$ (see Section 5). Thus, the theory of periodic states is made self-consistent. It is important that either $\bar{R}_F(x_0)$ or $\bar{S}_F(x_0)$ vanishes, depending on the feedback sign.

4. CHARACTERIZATION OF SLOW MOTIONS

4.1. Fundamental Amplitudes $\bar{R}_F$ and $\bar{S}_F$

According to the foregoing analysis, the amplitudes $\bar{R}_F(x)$ and $\bar{S}_F(x)$ obey system (1), (2) (with E replaced by $\bar{E}$) and satisfy the boundary conditions $\bar{R}_F(0) = 1$ and $\bar{S}_F(0) = 0$. The grating amplitude $\bar{E}$ is given by (11) as a quadratic form of $\bar{R}_F$ and $\bar{S}_F^*$.

To find $\bar{R}_F$ and $\bar{S}_F$, we use the fact that the system is invariant under the linear unitary transformation

$$\begin{aligned}\bar{R}_F &= Q_-\bar{R}_1 + Q_+^*\bar{S}_1^*,\\ \bar{S}_F^* &= -Q_+\bar{R}_1 + Q_-^*\bar{S}_1^*\end{aligned} \tag{15}$$

to new amplitudes $\bar{R}_1$ and $\bar{S}_1$, where the complex parameters $Q_\pm$ satisfy the condition $|Q_+|^2 + |Q_-|^2 = 1$. In physical terms, the invariance means that a variety of wave pairs (corresponding to various boundary conditions), rather than one, can be coupled by diffraction on the same grating. It is important for the present study that the quadratic form $\bar{E}(\bar{R}_F, \bar{S}_F^*)$ can be transformed by (15) into $\bar{E} \propto \bar{R}_1\bar{S}_1^*$ under an appropriate choice of $Q_\pm$. The differential equations for $\bar{R}_1$ and $\bar{S}_1$ thus obtained are easily solved, because they are formally identical to those describing the two-beam coupling due to photorefractive nonlinearity [2, 3]. The inverse transformation is used to obtain the required expressions for the fundamental amplitudes.

For present purposes, we can write

$$Q_\pm = \frac{1}{\sqrt{2}}\left(1 \mp \frac{W_0}{g}\right)^{1/2}, \tag{16}$$

where, again, $W_0 = |R_0|^2 - |S_0|^2$ and

$$g = \sqrt{W_0^2 + \varepsilon^2(1 - W_0^2)} \tag{17}$$

is a positive parameter. It is obvious that $g \geq |W_0|$, where $g = 1$ in the absence of phase modulation.

In the new representation, we have $\bar{E} = i\gamma\bar{R}_1\bar{S}_1^*$, where

$$\gamma = g(i + \Omega)^{-1} \tag{18}$$

is interpreted as the complex optical gain exponent. The corresponding input amplitudes are $\bar{R}_1(0) = Q_-$ and

$\bar{S}_1(0) = Q_+$, and $\bar{R}_1$ and $\bar{S}_1$ as functions of x are expressed as

$$\bar{R}_1 = Q_-[Q_-^2 + Q_+^2\exp(2\gamma' x)]^{-\gamma/2\gamma'},$$
$$\bar{S}_1^* = Q_+[Q_+^2 + Q_-^2\exp(-2\gamma' x)]^{-\gamma/2\gamma'}. \quad (19)$$

Hence, the required expressions for $\bar{R}_F$ and $\bar{S}_F$ are

$$\bar{R}_F = \left[Q_-^2\exp\left(-\frac{\gamma x}{2}\right) + Q_+^2\exp\left(\frac{\gamma x}{2}\right)\right] \times [Q_-^2\exp(-\gamma' x) + Q_+^2\exp(\gamma' x)]^{(i-\Omega)/2\Omega}, \quad (20)$$

$$\bar{S}_F^* = 2Q_+Q_-\sinh\frac{\gamma x}{2} \times [Q_+^2\exp(\gamma' x) + Q_-^2\exp(-\gamma' x)]^{(i-\Omega)/2\Omega}. \quad (21)$$

In the general case, both $\bar{R}_F$ and $\bar{S}_F$ are complex quantities, and it can be shown that the sum of their squared magnitudes is unity. When $\varepsilon = 1$, expressions (20) and (21) correspond to the fundamental amplitudes in the absence of phase modulation and in the presence of a frequency shift Ω between the input waves.

4.2. Conditions for 100% and Zero Diffraction Efficiency

The values $\eta = |S_F(x_0)|^2$ and $1 - \eta = |R_F(x_0)|^2$ quantify the diffraction efficiency and transmittance of a grating. 100% diffraction efficiency (i.e., zero transmission) is ensured by the condition $R_F(x_0) = 0$, whereas the condition for zero diffraction efficiency (100% transmission) is $S_F(x_0) = 0$. Since expressions (20) and (21) for $\bar{R}_F$ and $\bar{S}_F^*$ contain two feedback-controlled parameters, Ω and ε, both $\eta = 1$ and $\eta = 0$ can be attained in the leading order of perturbation theory.

Using (20), one can readily show that the condition $\bar{R}_F(x_0) = 0$ ($\eta \approx 1$) is equivalent to

$$x_0 = \frac{\pi^2 j^2 + L^2}{g\pi j}, \quad \Omega = -\frac{L}{\pi j}, \quad (22)$$

where

$$L = \ln\frac{g - W_0}{g + W_0},$$

and $j = 1, 3, \ldots$ is a positive odd number. Since $g = g(\varepsilon, W_0)$, the first relation in (22) defines a sequence of branches of the function $\varepsilon_j(x_0, W_0)$, while the second one determines the frequency detuning as a function of x_0 and W_0 for the jth branch. Here, ε and Ω are even and odd functions of W_0, respectively. Figure 2a shows ε versus the input intensity ratio $r_0 \equiv (1 + W_0)/(1 - W_0)$ (which is better suited for representation) for the branch with $j = 1$ and several values of x_0. Finally, note that the actual values of $|\Omega|$ do not exceed 1.5.

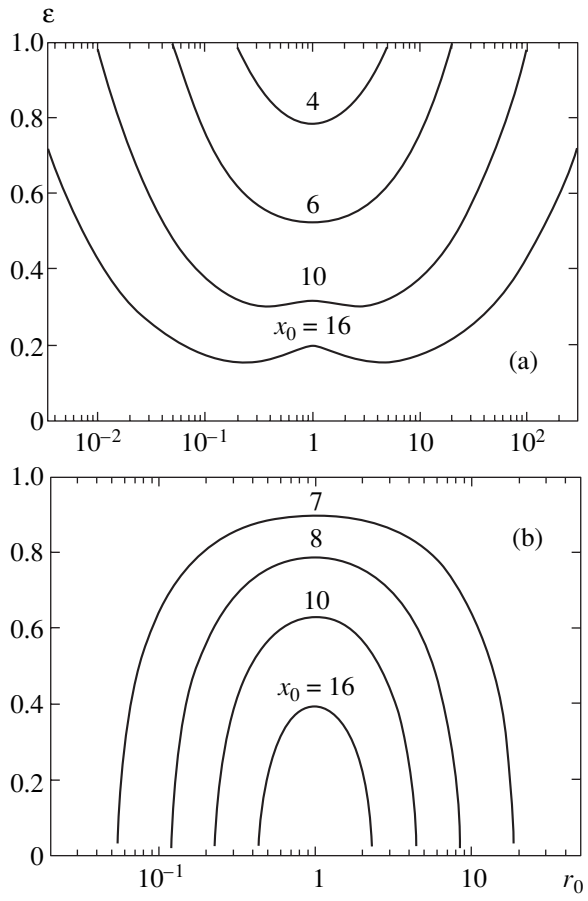


Fig. 2. Phase modulation parameter ε versus input intensity ratio $r_0 = |R(0)|^2/|S(0)|^2$ for several values of crystal thickness x_0: (a) $+\pi/2$ feedback; (b) $-\pi/2$ feedback.

Since $\varepsilon \le 1$, the admissible values of x_0 and $|\ln r_0|$ are bounded from below and above, respectively. Setting $\varepsilon = 1$ in (17) and (22), we obtain

$$x_0 = \pi j + \frac{\ln^2 r_0}{\pi j}. \quad (23)$$

This relation determines the boundary of the domain of x_0 and r_0 where $\eta = 1$. The minimal admissible crystal thickness, $x_0^{\min} = \pi$, corresponds to $r_0 = 1$ and the branch with $j = 1$. For this branch, $|\ln r_0|_{\max} = \sqrt{\pi(x_0 - \pi)}$; i.e., the range of admissible input intensity ratios rapidly increases with x_0. Figure 3 shows the function $x_0(r_0)$ for $j = 1$, 3, and 5. This demonstrates that the branch with $j = 1$ is the lowest in a wide range of parameters, $\pi \le x_0 \lesssim 12.5$ and $|\ln r_0| \lesssim 2$. This branch is of primary interest.

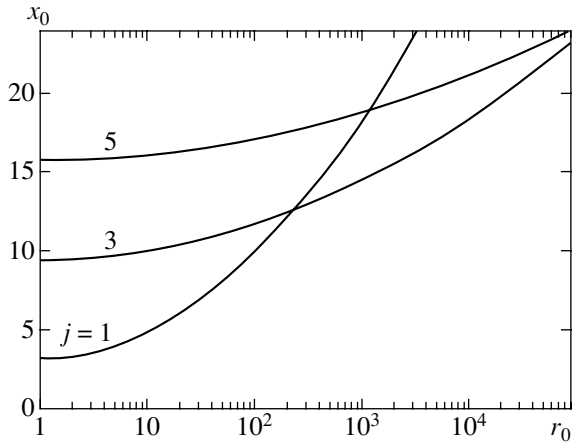


Fig. 3. Threshold curves of $x_0(r_0)$ for $+\pi/2$ feedback and $j = 1, 3, 5$.

Now, consider the condition $\bar{S}_F(x_0) = 0$ ($\eta \approx 0$). According to (21), it implies that $\gamma' = 0$ and $\gamma'' x_0 = -\pi j'$, where $j' = 2, 4, \ldots$. These relations are equivalent to

$$\varepsilon^2 = \frac{(\pi j'/x_0)^2 - W_0^2}{1 - W_0^2}, \quad \Omega = 0. \tag{24}$$

The minimal admissible thickness, $x_0^{\min} = 2\pi$, corresponds to $j' = 2$ and $W_0 = 0$ ($r_0 = 1$), twice as large as that obtained above. Figure 2b shows the $\varepsilon(r_0)$ for $j' = 2$ and several values of x_0. Also being an even function, it is characterized by some new properties. The value of ε decreases not only with increasing x_0, but also with increasing $|\log r_0|$. This means that the range of admissible values of r_0 is reduced by increasing the thickness x_0. Setting $\varepsilon = 0$, we obtain $|W_0| \le 2\pi/x_0$ for this interval. Finally, we note that the condition $\bar{S}_F(x_0) = 0$ is always satisfied when $\varepsilon = 0$ (the grating amplitude vanishes). Outside the interval indicated above, this trivial result is the only possible one.

When the growth exponent is an imaginary number ($\gamma' = 0$), there is no energy exchange between light waves, and expressions (20) and (21) for $\bar{R}_F$ and $\bar{S}_F^*$ are substantially simplified. For $j' = 2$, they reduce to

$$\bar{R}_F = \left(\cos\frac{\pi x}{x_0} + i\frac{W_0 x_0}{2\pi}\sin\frac{\pi x}{x_0}\right)\exp\left(-\frac{iW_0 x}{2}\right),$$

$$\bar{S}_F^* = -\left[i\left(1 - \frac{W_0^2 x_0^2}{4\pi^2}\right)^{1/2}\sin\frac{\pi x}{x_0}\right]\exp\left(-\frac{iW_0 x}{2}\right).$$

4.3. Spatial Distribution of the Grating Amplitude

Using (11), (20), and (21), one can easily find

$$\bar{E} = \frac{\varepsilon R_0 S_0}{1 - i\Omega}\left(\frac{\cosh\bar{a}}{\cosh(\gamma' x - \bar{a})}\right)^{1 - i/\Omega}. \tag{25}$$

The parameter $\bar{a}$ is defined by the relation $W_0 = g\tanh\bar{a}$. It can be positive or negative, depending on W_0 and x_0. The use of $\bar{a}$ is dictated by convenience considerations. The coefficient that multiplies the brackets is $\bar{E}(0)$.

Conditions (22) or (24) set the feedback parameters Ω and ε (and hence $\bar{a}$) and make the distribution $\bar{E}(x)$ dependent only on W_0 (or r_0) and x_0 for each jth or j'th branch. We restrict our analysis to the most important branches, $j = 1$ and $j' = 2$. Note also that the real functions $|\bar{E}(x)|$ and $\arg[\bar{E}(x)]$ are important quantitative and qualitative characteristics of the system describing the grating amplitude and the corresponding fringe distribution, respectively.

When $\eta = 1$, (22) yields the following simplifying relations: $\gamma' x_0 = -L = 2\bar{a}$ and $\Omega = 2\bar{a}/\pi$. Then, (25) reduces to

$$\left|\frac{\bar{E}(x)}{\bar{E}(0)}\right| = \frac{\cosh\bar{a}}{\cosh[\bar{a}(2x x_0^{-1} - 1)]}, \tag{26}$$

$$\arg\bar{E}(x) - \arg\bar{E}(0) = -\frac{\pi}{2\bar{a}}\ln\left|\frac{\bar{E}(x)}{\bar{E}(0)}\right|, \tag{27}$$

where the input parameters are expressed as

$$|\bar{E}(0)| = \frac{\varepsilon|R_0 S_0|}{\sqrt{1 + (2\bar{a}/\pi)^2}}, \quad \arg\bar{E}(0) = \arctan\frac{2\bar{a}}{\pi}. \tag{28}$$

According to (26), the profile $|\bar{E}(x)|$ is symmetric about the center of the crystal and $|\bar{E}(x)|/|\bar{E}(0)|$ increases from 1 to $\cosh\bar{a}$ from the boundaries ($x = 0, x_0$) toward the center ($x = x_0/2$). Thus, $\cosh\bar{a}$ can be interpreted as the degree of spatial nonuniformity of the grating profile. The spatial distribution of the grating phase is also symmetric about the center of the crystal. The corresponding degree of spatial nonuniformity is quantified by $(\pi/2\bar{a})\ln(\cosh\bar{a})$. Note that, by the definition of $\bar{a}$, it vanishes when $W_0 = 0$, i.e., when the input intensities are equal. In this special case, the grating is spatially uniform.

Figure 4 shows $|\bar{E}(x_0/2)/\bar{E}(0)|$ and $\arg[\bar{E}(x_0/2)/\bar{E}(0)]$ as functions of r_0 for several values of crystal thickness. It demonstrates that the effects due to nonuniformity, including fringe modulation and curvature, increase substantially with $|\log r_0|$.

The input parameter $|\bar{E}(0)|$ is an even function of $\log r_0$ with a maximum at $\log r_0 = 0$. As x_0 increases from 4 to 10, the maximum values decreases from 0.19 to 0.075. The parameter $\arg\bar{E}(0)$ is an odd function of $\log r_0$ close to $-\arg[\bar{E}(x_0/2)/\bar{E}(0)]$.

When $\eta = 0$ ($\gamma' = \Omega = 0$), the grating is spatially uniform and stationary for every admissible W_0 and x_0. According to (25), its amplitude (for $j' = 2$) is

$$\bar{E} = \exp(-iW_0x)\sqrt{\left(\frac{\pi}{x_0}\right)^2 - \left(\frac{W_0}{2}\right)^2}. \quad (29)$$

This case is much simpler than the preceding one.

At first glance, the combination of a nonzero grating amplitude with zero diffraction efficiency is a paradox. However, it is consistent with general principles and is explained by energy exchange between the *R*- and *S*-waves. In contrast to the well-known Kogelnik formula $\eta = \sin^2(|E|x_0)$, which is valid for a grating with a constant amplitude (E = const), the relations obtained here demonstrate the feasibility of gratings with $\eta = 0$ in a wide range of x_0 and r_0.

5. PERIODIC STATES

To derive closed equations for the "fast periodic" phase $\varphi_p(t)$, one must calculate the product

$$S_0R_0R_F(x_0, t)S_F(x_0, t)\exp(-i\Omega t),$$

which is contained in feedback condition (6) by virtue of (5). Using (8), (9), (13), (14), and the fact that either $\bar{R}_F(x_0)$ or $\bar{S}_F(x_0)$ vanishes for the periodic states controlled by the $\pm\pi/2$ feedback, we have

$$R_F(x_0)S_F(x_0)\exp(-i\Omega t) \approx \mp B(x_0). \quad (30)$$

The fast amplitude $\tilde{E}$ contained in the expression for B in (14) is determined from (12):

$$\tilde{E} = |R_0S_0|[(\tilde{R}_F^2 - \bar{S}_F^{*2})u - i(\bar{R}_F^2 + \bar{S}_F^{*2})v]. \quad (31)$$

The known amplitudes $\bar{R}_F$ and $\bar{S}_F$ depend on x, and the periodic functions u and v of time are related to $\varphi_p(t)$ by the equations

$$\dot{u} = -\cos\varphi_p - \varepsilon, \quad \dot{v} = \sin\varphi_p. \quad (32)$$

Substituting (31) into (14), we obtain

$$|R_0S_0|B(x_0, t) = -c_+v + ic_-u, \quad (33)$$

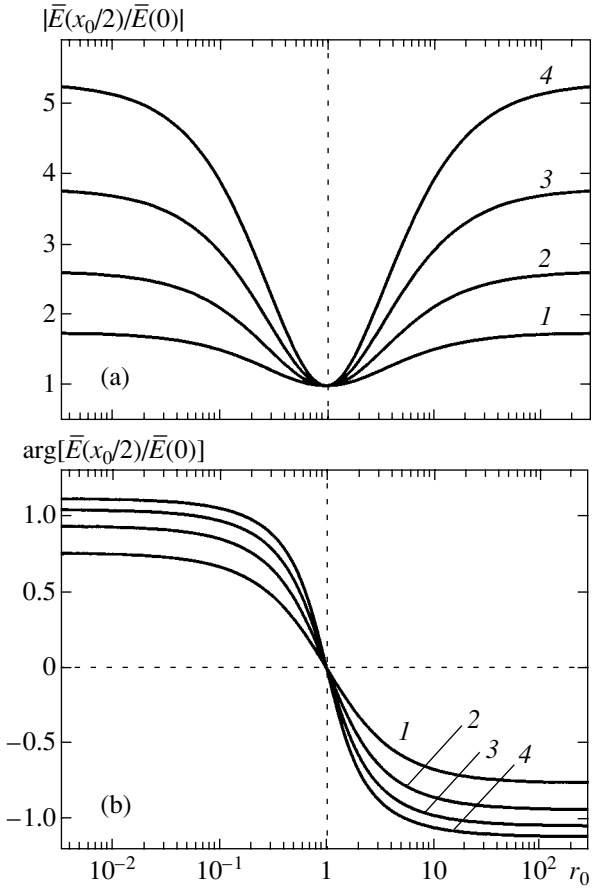


Fig. 4. Amplitude and phase distortion characteristics $|\bar{E}(x_0/2)|/|\bar{E}(0)|$ (a) and $\arg[\bar{E}(x_0/2)/\bar{E}(0)]$ (b) versus input intensity ratio for x = 4 (*1*), 6 (*2*), 8 (*3*), and 10 (*4*) in the case of $+\pi/2$ feedback.

where the coefficients

$$c_\pm = |R_0S_0|^2\int_0^{x_0}(1 \pm 2\bar{R}_F^2\bar{S}_F^2 - 2|\bar{R}_F\bar{S}_F|^2)dx \quad (34)$$

depend only on x_0 and r_0. The amplitudes $\bar{R}_F(x)$ and $\bar{S}_F(x)$ contained in the integrand are such that either $\eta = 1$ or $\eta = 0$. It is important that the coefficients $c_\pm$ are real. To demonstrate this and calculate some integrals, we recall the results obtained in the preceding section.

Using (20)–(22) and (24), one can readily show that

$$\bar{R}_F\bar{S}_F = \left\{\sinh\left[\bar{a}\left(\frac{2x}{x_0} - 1\right)\right] + \sinh\bar{a}\cos\frac{\pi x}{x_0} + i\cosh\bar{a}\sin\frac{\pi x}{x_0}\right\} \times \left\{2\cosh\bar{a}\cosh\left[\bar{a}\left(\frac{2x}{x_0} - 1\right)\right]\right\}^{-1} \quad (35)$$

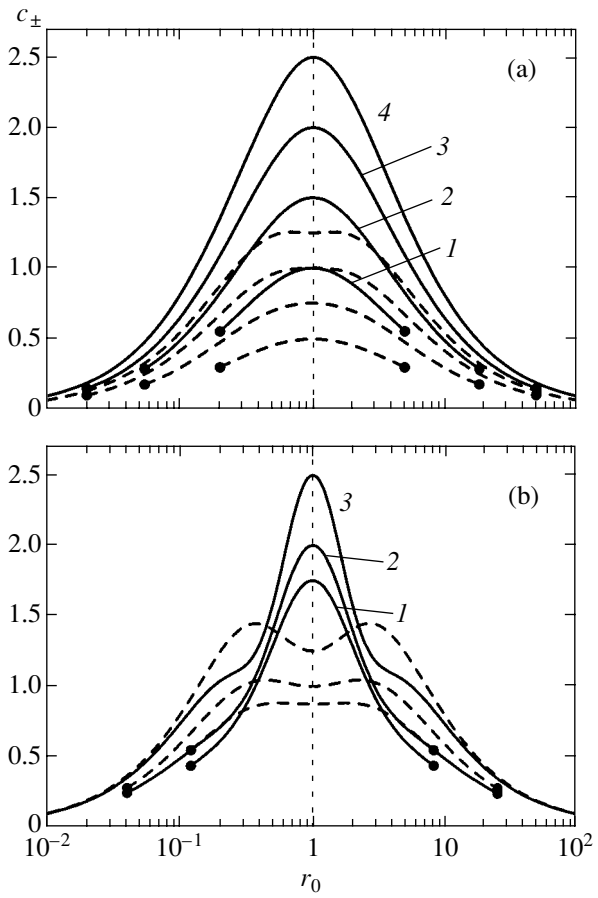


Fig. 5. Coefficients $c_\pm$ versus r_0 for several values of x_0: (a) $+\pi/2$ feedback; (b) $-\pi/2$ feedback. Solid curves and dashed curves represent c_- and c_+, respectively; (a) $x_0 = 4$ (*1*), 6 (*2*), 8 (*3*), and 10 (*4*); (b) $x_0 = 7$ (*1*), 8 (*2*), and 9 (*3*); symbols correspond to the endpoints of intervals of admissible values.

for $\eta \approx 1$, $j = 1$ and

$$\bar{R}_F\bar{S}_F = \frac{1}{2\cosh^2\bar{a}}\left(i\cosh\bar{a}\sin\frac{2\pi x}{x_0} - 2\sinh\bar{a}\sin^2\frac{\pi x}{x_0}\right) \quad (36)$$

for $\eta \approx 0$, $j' = 2$. In both cases, the real and imaginary parts of $\bar{R}_F(x)\bar{S}_F(x)$ have opposite parities with respect to the center of the crystal. This implies that the coefficients $c_\pm$ are real. Figure 5 shows $c_\pm(r_0)$ for several values of x_0 when $\eta = 1$ and 0. If $r_0 = 1$, then $c_+ = x_0/8$ and $c_- = x_0/4$. The curves presented here do not exhibit singular behavior of any type near the boundaries of the admissibility intervals. If $\eta \approx 1$, then c_+/c_- monotonically increases with x_0 and $|\log r_0|$ from 0.5 to approximately 0.7. If $\eta \approx 0$, this ratio can be as high as 1.0–1.2.

Finally, we use (5), (6), (30), and (33) to derive and equation for φ_p:

$$t_f\dot{\varphi}_p = c_-u\sin\varphi_p - c_+v\cos\varphi_p. \quad (37)$$

It is independent of the feedback sign. Proportional variation of c_+ and c_- is equivalent to renormalization of t_f. Ordinary differential equations (32) and (37) make up a closed system for the periodic functions φ_p, u, and v. This system is essentially nonlinear.

Note that $\bar{S}_F(x_0) \approx -i$ and $R_F(x_0, t) \approx -iB(x_0, t)$ in the case of $+\pi/2$ feedback. The periodic function of time

$$-iB = \frac{c_-u + ic_+v}{|R_0S_0|}$$

defines a closed trajectory on the complex plane in the neighborhood of the origin. This trajectory provides a convenient representation of the system's dynamics (related to experiment). In the case of $-\pi/2$ feedback, the trajectory of the system is described by the periodic function

$$S_F(x_0, t) \approx -B(x_0, t) = \frac{-ic_-u + c_+v}{|R_0S_0|}$$

close to zero.

It is understood that both period T and maximal deviation $\delta\eta_{max}$ of diffraction efficiency from 1 or 0 depend on the feedback response time t_f. The equations of motion obtained here can easily be used to determine these dependences. By normalizing t and the variables u and v to $\sqrt{t_f}$ (i.e., by changing to intrinsic reference values), the small parameter t_f is eliminated from the equation of motion. Hence, we have the scaling relations

$$T \propto \sqrt{t_f}, \quad \delta\eta_{max} \propto t_f. \quad (38)$$

They are consistent with results obtained by direct numerical simulation [19, 20]. The proportionality coefficients in (38) depend on the type of periodic state.

Finally, we describe some useful symmetry properties of system (32), (37). If it is solved by $\varphi_p(t)$, $u(t)$, and $v(t)$, then the combination of $-\varphi_p(t)$, $u(t)$, and $-v(t)$ provides another solution. Furthermore, time inversion is equivalent to sign reversal in u and v.

Since the system under analysis is autonomous, its order can be reduced. Setting $u(t) = u(\varphi_p(t))$ and $v(t) = v(\varphi_p(t))$, we obtain

$$\frac{1}{t_f}\frac{du}{d\varphi_p} = \frac{\cos\varphi_p - \varepsilon}{c_-u\sin\varphi_p - c_+v\cos\varphi_p},$$
$$\frac{1}{t_f}\frac{dv}{d\varphi_p} = \frac{\cos\varphi_p}{c_-u\sin\varphi_p - c_+v\cos\varphi_p}. \quad (39)$$

Next, we note that periodic conditions can be classified by the condition $\varphi_p(t + T) - \varphi_p(t) = 2\pi N$, where $N = 0$, ±1, … is an integer. Accordingly, the period of $u(\varphi_p)$ and $v(\varphi_p)$ is $2\pi N$. When $\varphi_p(t)$ is not a monotonic function (e.g., when $N = 0$), the functions $u(\varphi_p)$ and $v(\varphi_p)$ are two-valued. Since solutions with nonzero N and $-N$

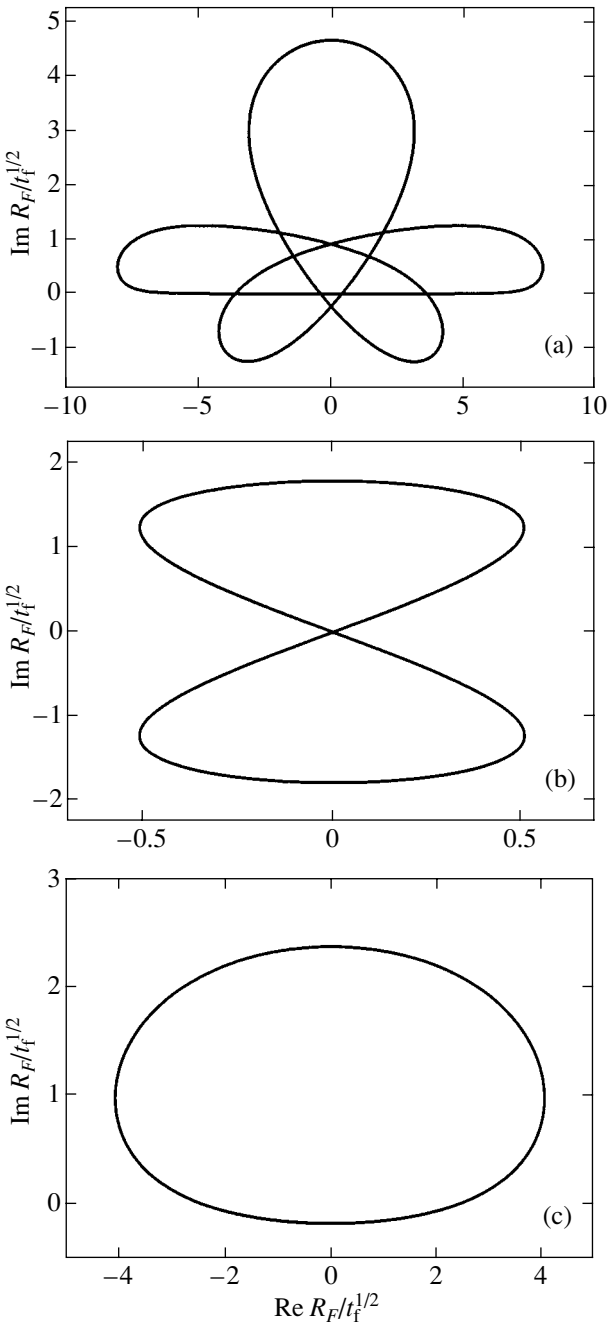


Fig. 6. Closed contours in the complex plane illustrating the time dependence of $R_F(x_0)$ for $x_0 = 6.6$ in the case of $+\pi/2$ feedback: $N = 3$ (a), 0 (b), and 1 (c); $r_0 = 1$ (a), 6 (b), and 1 (c).

are subject to the symmetry conditions indicated above, it is sufficient to consider the cases with $N \geq 0$. If the periodic functions $u(\varphi_p)$ and $v(\varphi_p)$ are known, then $\varphi_p(t)$ can be found by solving Eq. (37).

Nonlinear equations (39) cannot be integrated analytically. However, this can easily done numerically, e.g., by using the standard Mathematica system. Numerical integration is virtually equivalent to complete analysis of solutions.

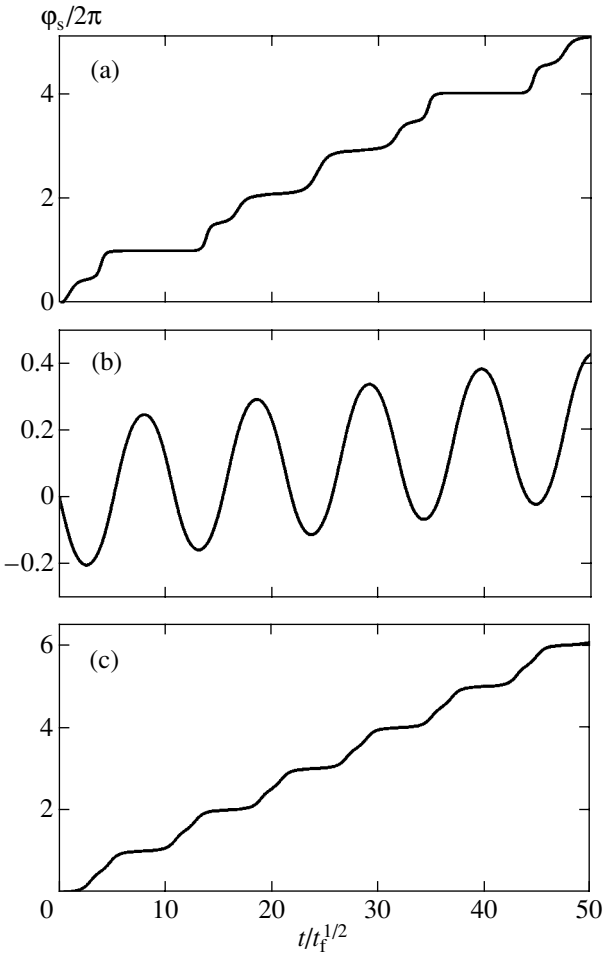


Fig. 7. Time dependence of input phase φ_s for the periodic states illustrated by Fig. 6.

Figure 6a shows the closed contour of $R_F(x_0, t)/\sqrt{t_f}$ in the complex plane calculated for $\eta \approx 1$, $x_0 = 6.6$, $r_0 = 1$, and $N = 3$. Over the period $T \approx 30.7\sqrt{\tau_f}$, the phase φ_p changes by 6π. Here, value of $\delta\eta_{max} = |R_F(x_0, t)|^2_{max}$ is approximately equal to $\approx 64.5t_f$. The corresponding time-dependent input phase φ_s is shown in Fig. 7a. Since $\Omega = 0$, it is similar to $\varphi_p(t)$. The phase exhibits three pronounced jumps within a period. These results are in good agreement with the results of direct numerical simulations and experimental measurements [19, 20]. Analogous behavior of $S_F(x_0, t)$ and $\varphi_p(t)$ is predicted for $\eta = 0$.

Figure 6b shows the contour calculated for $\eta = 1$, $x_0 = 6.6$, $r_0 = 6$, and $N = 0$. Here, the phase φ_p is periodic with period $T \approx 10.5\sqrt{t_f}$, and $\delta\eta_{max} \approx 3.2t_f$. The frequency detuning Ω is approximately 0.85 in this case. Owing to the detuning, the periodic oscillation of the input phase φ_s is superimposed on a linear growth (see

Fig. 7b). These trends are also consistent with numerical and experimental results.

Our calculations show that, generally, there exist periodic states with $N = 0, 1, 2, \ldots$ corresponding to the same values of x_0 and r_0. In other words, periodic solutions to system (32), (37) are not unique. Figures 6c and 7c show, respectively, the closed contour and $\varphi_s(t)$ corresponding to $N = 1$, $x_0 = 6.6$, and $r_0 = 1$. The contour has a single-lobed shape, the phase exhibits one jump over the period $T \approx 8.0\sqrt{t_f}$, and $\delta\eta \approx 17.5t_f$.

The number of solutions increases with distance from the threshold, i.e., from the boundary of the domain of periodic states (see Fig. 3). For simplicity, let us consider the case of $\eta \approx 1$ and $r_0 = 1$, when $x_0 - \pi = \pi(\varepsilon^{-1} - 1)$ is an intrinsic supercriticality parameter. In the near-threshold interval $\pi < x_0 \lesssim 4.2$, where ε varies from 1 to approximately 0.74, the periodic state with $N = 0$ is unique. As the threshold is approached, the amplitude of phase oscillation goes to zero as $(x_0 - \pi)^{1/2}$, while the period T decreases insignificantly. The absence of solutions with $N \neq 0$ in the near-threshold region should be expected, because the effects due to phase modulations would be eliminated via abrupt changes in φ_s (by multiples of 2π) and the system's inertia would preclude such jumps.

When $x_0 \gtrsim 4.2$, there exists a solution with $N = 1$ (in addition to one with $N = 0$). As x_0 approaches the lower boundary of the interval, it does not exhibit any threshold singularity and disappears via loss of stability. When $x_0 \gtrsim 5.7$ ($\varepsilon \lesssim 0.54$), there exists a periodic state with $N = 3$. As x_0 increases, it softly splits off from the state with $N = 1$ via period tripling (see also Figs. 6 and 7). When $x_0 \approx 6.7$ ($\varepsilon \approx 0.47$), a state with $N = 2$ splits off from the one with $N = 1$ via period doubling.

Analogous behavior is exhibited by the system when the distance from the separatrix is increased with decreasing $|\log r_0|$. The trends described above provide a basis for understanding the results of numerical and physical experiments presented in [19]. For thick crystals with $x_0 \gtrsim 12.5$, additional manifestations of threshold singularity can be expected as the branches with $j = 1$ and 2 are crossed (see Fig. 3) if the input intensity ratio is sufficiently large ($|\log r_0| \gtrsim 2.1$).

6. CONCLUSIONS

We have developed an efficient method for analyzing a new highly nonlinear model of feedback-controlled beam coupling and recording of gratings. The method makes use of the symmetries of time-dependent diffraction equations and fast response of the feedback. No close analogy can be found in the physics of nonlinear distributed systems. The method relies on a number of new concepts.

The new method is used to analyze the spatial structure of the refractive-index gratings corresponding to $\eta = 1$ and 0. Scaling relations are obtained for the oscillation period T and the oscillation amplitude $\delta\eta_{max}$. Some observable properties of various periodic states are described. A close relationship between the regimes with $\eta \approx 1$ and $\eta = 0$ is established. The scenario of the elimination of periodic states observed as the separatrix is approached is elucidated. These results supplement those obtained by direct numerical simulation.

It is remarkable that periodic states may not be unique. When the distance from the threshold is sufficiently large, there exist several periodic solutions characterized by different manifestations for the same input parameters. This finding suggests that stability of the periodic states should be examined. Moreover, we cannot rule out the possibility that memory effects are essential for implementation of a particular regime.

The local nonlinear response considered here is most pronounced in $LiNbO_3$ crystals, where photovoltaic charge transport plays a dominant role in grating formation [23]. This material is utilized in development of various devices based on diffraction effects and in most experiments on optical feedback. The typical dimensionless thickness of the crystals does not exceed 10, which corresponds to an actual thickness of several millimeters. Therefore, the branches with $j = 1$ and $j' = 0$ and 2 are most relevant, which justifies certain restrictions imposed in the course of our theoretical analysis.

Theoretical considerations suggest that experiments on $LiNbO_3$:Fe crystals of variable thickness with constant concentrations of Fe^{2+} and Fe^{3+} ions should be conducted to verify the predicted successive onset of new periodic states with increasing supercriticality parameter. The scaling relations for T and $\delta\eta_{max}$ should also be relatively easy to verify experimentally by varying the electronic circuit parameters that determine the feedback response time.

Nonlocal photorefractive response is associated with diffusive transport and is described by Eq. (3) with imaginary unit on the right-hand side. It is of particular interest for enhancement of optical gain. In systems based on this phenomenon, periodic states with $\eta \approx 1$ can be implemented under more stringent requirements for crystal parameters and experimental conditions. Analysis of these states must rely on a refined theory.

We believe that the systems with maximized and minimized diffraction efficiency examined above exemplify a more general class of optical devices, in which electronic feedback loops are responsible for implementation of useful functions. Development of an appropriate theory is a challenging task of great practical importance.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 03-02-16083.

REFERENCES

1. *Topics in Applied Physics*, Vols. 61 and 62: *Photorefractive Materials and Their Applications*, Ed. by P. Günter and J.-P. Huignard (Springer, Berlin, 1988 and 1989).
2. M. P. Petrov, S. I. Stepanov, and A. V. Khomenko, *Photorefractive Crystals in Coherent Optics* (Nauka, St. Petersburg, 1992).
3. L. Solymar, D. J. Webb, and A. Grunnet-Jepsen, *The Physics and Applications of Photorefractive Materials* (Clarendon Press, Oxford, 1996).
4. S. Stepanov, Rep. Prog. Phys. **57**, 39 (1994).
5. T. Tschudi, A. Herden, J. Goltz, *et al.*, IEEE J. Quantum Electron. **22**, 1493 (1986).
6. B. Ya. Zel'dovich, N. F. Pilipetskiĭ, and V. V. Shkunov, *Reversal of a Wave Front* (Nauka, Moscow, 1985).
7. M. Cronin-Golomb, B. Fischer, J. O. White, and A. Yariv, IEEE J. Quantum Electron. **20**, 12 (1984).
8. S. G. Odulov, M. S. Soskin, and A. I. Khizhnyak, *Dynamic Lattice Lasers* (Nauka, Moscow, 1990).
9. M. Segev, B. Crosignani, A. Yariv, *et al.*, Phys. Rev. Lett. **68**, 923 (1992); M. F. Shih, M. Segev, and G. Salamo, Phys. Rev. Lett. **78**, 2551 (1997).
10. A. Yariv, S. S. Orlov, and G. A. Rakuljic, J. Opt. Soc. Am. B **13**, 2513 (1996).
11. B. Sturman, S. Odoulov, and M. Goul'kov, Phys. Rep. **275**, 197 (1996).
12. A. Bledowski, J. Otten, K. H. Ringhofer, and B. Sturman, Zh. Éksp. Teor. Fiz. **102**, 406 (1992) [Sov. Phys. JETP **75**, 215 (1992)]; B. I. Sturman, A. I. Chernykh, and E. Shamonina, Zh. Éksp. Teor. Fiz. **114**, 1034 (1998) [JETP **87**, 563 (1998)].
13. M. V. Gorkunov, E. V. Podivilov, and B. I. Sturman, Zh. Éksp. Teor. Fiz. **121**, 551 (2002) [JETP **94**, 470 (2002)]; E. V. Podivilov, B. I. Sturman, H. Pedersen, and P. M. Johansen, Phys. Rev. Lett. **85**, 1867 (2000).
14. G. F. Calvo, B. Sturman, F. Agullo-Lopez, and M. Carrascosa, Phys. Rev. Lett. **84**, 3839 (2000); **89**, 033902 (2002).
15. A. Freschi and J. Frejlich, J. Opt. Soc. Am. B **11**, 1837 (1994).
16. P. M. Garcia, K. Buse, D. Kip, and J. Frejlich, Opt. Commun. **117**, 35 (1995).
17. P. M. Garcia, A. A. Freschi, J. Frejlich, and E. Krätzig, Appl. Phys. B **63**, 207 (1996).
18. V. P. Kamenov, K. H. Ringhofer, B. I. Sturman, and J. Frejlich, Phys. Rev. A **56**, R2541 (1997).
19. E. V. Podivilov, B. I. Sturman, S. G. Odoulov, *et al.*, Phys. Rev. A **63**, 053805 (2001).
20. E. V. Podivilov, B. I. Sturman, S. G. Odoulov, *et al.*, Opt. Commun. **192**, 399 (2001).
21. K. H. Ringhofer, V. P. Kamenov, B. Sturman, *et al.*, Phys. Rev. E **61**, 2029 (2000).
22. A. S. Sonin, *Introduction to the Physics of Liquid Crystals* (Nauka, Moscow, 1983).
23. B. I. Sturman and V. M. Fridkin, *Photovoltaic Effect in Media without Center of Symmetry and Related Phenomena* (Nauka, Moscow, 1992).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 908–917.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1041–1051.

ATOMS, SPECTRA, RADIATION

Expansion of Bose–Einstein Condensates Confined in Quasi-One-Dimensional or Quasi-Two-Dimensional Traps

A. M. Kamchatnov

Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia
e-mail: kamch@isan.troitsk.ru

Received October 31, 2003

Abstract—Solutions to the Gross–Pitaevskii equations are obtained in the hydrodynamic approximation for a repulsive Bose gas that expands after a quasi-one-dimensional or quasi-two-dimensional trap is removed. The results are expressed in terms of measurable parameters, such as the initial condensate size and the oscillation frequencies of trapped particles. Three-dimensional effects are calculated by a variational method. The analytical results are in good agreement with available experimental data. *© 2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The properties of a Bose–Einstein condensate in which particle motion is "frozen" or reduced to zero-point oscillations in one or two directions are the subject of intensive studies [1–14]. In experiments, cigar-shaped quasi-one-dimensional condensates are created by using optical dipole traps [1]. A quasi-two-dimensional condensate was created in an array of disc-shaped traps provided by the periodic potential of a laser beam [2]. When the traps are sufficiently deep, the motion along the array is frozen and the condensate splits into several independent condensates confined in separate potential wells.

Important experimental information about the properties of a Bose–Einstein condensate confined in a three-dimensional trap can be extracted by measuring the time-dependent density of the expanding atomic cloud after the trapping potential is switched off. In the mean-field approximation, the dynamics of a dilute condensate is described by the Gross–Pitaevskii equation [14]

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + V_{\text{ext}}(\mathbf{r})\psi + g|\psi|^2\psi, \qquad (1)$$

where

$$V_{\text{ext}}(\mathbf{r}) = \frac{1}{2}m(\omega_x^2x^2 + \omega_y^2y^2 + \omega_z^2z^2)$$

is the trapping potential,

$$g = 4\pi\hbar^2a_s/m \qquad (2)$$

is the nonlinear coupling constant associated with an atom–atom scattering length a_s, and the condensate wave function ψ is normalized to the number of atoms

$$\int|\psi|^2d\mathbf{r} = N. \qquad (3)$$

If the number of atoms is sufficiently large, then the Gross–Pitaevskii equation can be transformed into hydrodynamic equations that admit simple self-similar solutions describing both oscillations of a gas in a parabolic trapping potential and its free three-dimensional expansion after the potential is switched off [15–18]. This theory is perfectly consistent with experiment.

A different situation arises when some degrees of freedom of the expanding condensate remain frozen. Recently, condensate expansion was investigated in quasi-one-dimensional waveguides [1] and in systems of two-dimensional discs [2]. This promising line of research was pursued in several studies. In [19], quasi-one-dimensional condensate expansion was analyzed without taking into account the transverse "quantum pressure." In [20], the effects due to quantum pressure were taken into account for steady states, in which case only the two transverse modes contribute to the pressure. In [13], the ground states of condensates confined in cigar- and disc-shaped traps were calculated by a variational method, but no analysis of the dynamics of condensate expansion was presented.

In this paper, an analytical study of quasi-one-dimensional and quasi-two-dimensional condensate expansion is presented. Conditions are formulated under which the three-dimensional Gross–Pitaevskii equation can be reduced to analogous equations in fewer coordinates. These equations are solved in the hydrodynamic approximation under initial conditions corresponding to a trapped condensate in equilibrium before the trap is switched off. The condensate expands either along the axis of a quasi-one-dimensional waveguide or in the plane of a quasi-two-dimensional trap. However, if the conditions for reduction to Gross–Pitaevskii equations of lower dimension are violated, then the gas flow is three-dimensional. Three-dimensional effects in the flow are calculated by a variational

method. Finally, it is shown that the theoretical results agree with experiment.

2. QUASI-ONE-DIMENSIONAL AND QUASI-TWO-DIMENSIONAL CONDENSATE EXPANSION WITHOUT THREE-DIMENSIONAL EFFECTS

It is well known that the Gross–Pitaevskii equation can be formulated as a principle of least action with the action functional

$$S = \int L dt, \quad L = \int \mathcal{L} d\mathbf{r}, \tag{4}$$

where the Lagrangian density is

$$\mathcal{L} = \frac{i\hbar}{2}(\psi_t^*\psi - \psi_t\psi^*) + \frac{\hbar^2}{2m}|\nabla\psi|^2 + V_{\text{ext}}|\psi|^2 + \frac{1}{2}g|\psi|^4. \tag{5}$$

In the case of a cigar- or disc-shaped trap, one can readily find conditions under which the tightly restrained degrees of freedom are frozen and the Gross–Pitaevskii equation reduces to a one- or two-dimensional equation, respectively. Even though this problem has been considered more than once, we briefly review here the basic points of the derivation in order to identify the essential parameters of the theory and formulate conditions for its applicability.

2.1. One-Dimensional Expansion

If the longitudinal frequency ω_z for an axially symmetric trap is much less than the transverse trap frequency $\omega_\perp$,

$$\lambda = \omega_z/\omega_\perp \ll 1, \tag{6}$$

and the transverse zero-point energy is much higher than the nonlinear interaction energy per atom, then the transverse motion reduces to the ground state of particle oscillation, with the amplitude

$$a_\perp = (\hbar/m\omega_\perp)^{1/2}.$$

Denoting by Z_0 the characteristic size of the condensate along the axis of a cigar-shaped trap, one can use the estimate

$$N \sim |\psi|^2 a_\perp^2 Z_0$$

(see (3)) to write the corresponding condition as follows (e.g., see [11]):

$$Na_s/Z_0 \ll 1. \tag{7}$$

If this condition is satisfied, then the condensate wave function can be factorized:

$$\psi(\mathbf{r}, t) = \phi(x, y)\Psi(z, t), \tag{8}$$

where

$$\phi(x, y) = \frac{1}{\sqrt{\pi}a_\perp}\exp\left(-\frac{x^2 + y^2}{2a_\perp^2}\right) \tag{9}$$

is the wave function of the ground state of transverse motion. Substituting (8) and (9) into (4) and (5) and integrating the result over the condensate's cross section, one obtains the action expressed in terms of the one-dimensional Lagrangian density

$$\mathcal{L}_{1D} = \frac{i\hbar}{2}(\Psi_t^*\Psi - \Psi_t\Psi^*) + \frac{\hbar^2}{2m}|\Psi_z|^2 + \frac{1}{2}m\omega_z^2 z^2|\Psi|^2 + \frac{g}{4\pi a_\perp^2}|\Psi|^4. \tag{10}$$

Then, the evolution of $\Psi(z, t)$ obeys the one-dimensional Gross–Pitaevskii equation

$$i\hbar\Psi_t = -\frac{\hbar^2}{2m}\Psi_{zz} + \frac{1}{2}m\omega_z^2 z^2\Psi + g_{1D}|\Psi|^2\Psi, \tag{11}$$

where

$$g_{1D} = \frac{g}{2\pi a_\perp^2} = \frac{2\hbar^2 a_s}{ma_\perp^2} \tag{12}$$

is an effective coupling constant and Ψ is normalized as

$$\int |\Psi|^2 dz = N. \tag{13}$$

Equation (11) determines the longitudinal dynamics of a condensate in a cigar-shaped trap.

By the well-known substitution

$$\Psi(z, t) = \sqrt{\rho(z, t)}\exp\left(\frac{im}{\hbar}\int^z v(z', t)dz'\right), \tag{14}$$

Eq. (11) is transformed into the system

$$\rho_t + (\rho v)_z = 0, \tag{15}$$

$$v_t + vv_z + \frac{g_{1D}}{m}\rho_z + \omega_z^2 z + \frac{\hbar^2}{2m^2}\left(\frac{\rho_z^2}{4\rho^2} - \frac{\rho_{zz}}{2\rho}\right)_z = 0. \tag{16}$$

In Eq. (16), the last term ("quantum pressure") can be neglected if it is much smaller than the nonlinear term, i.e., if

$$\frac{a_\perp}{Z_0} \ll \frac{Na_s}{a_\perp}. \tag{17}$$

Then, Eq. (16) reduces to

$$v_t + vv_z + \frac{g_{1D}}{m}\rho_z + \omega_z^2 z = 0. \tag{18}$$

Combined with Eq. (15), it constitutes the hydrodynamic approximation describing the evolution of a condensate.

The time-independent solution of the hydrodynamic equations is the well-known Thomas–Fermi distribution of a one-dimensional condensate:

$$\rho(z) = \frac{3N}{4Z_0}\left(1 - \frac{z^2}{Z_0^2}\right), \quad v = 0, \tag{19}$$

where the integration constant Z_0 (longitudinal half-length of the condensate) can be expressed in terms of the number of atoms N as

$$Z_0 = (3Na_s a_\perp^2 \lambda^{-2})^{1/3}. \tag{20}$$

Applicability conditions (7) and (17) for the one-dimensional hydrodynamic approximation can be rewritten by substituting (20) as follows (see [11]):

$$\sqrt{\lambda} \ll Na_s/a_\perp \ll 1/\lambda. \tag{21}$$

Now, assume that the longitudinal trapping potential is switched off and the condensate can freely expand along the longitudinal axis. At the same time, it remains transversely confined, and its transverse motion remains frozen in ground state (9). Accordingly, the expansion can be described by hydrodynamic equations (15) and (18) subject to initial conditions (19). An analogous problem in nonlinear optics was solved long ago [22], with a "pressure" ρ_z in (18) having the opposite sign, and its solution was recently applied to describe three-dimensional condensate expansion [15–18]. This approach is used here to analyze the case when the condensate expands into a "waveguide." A solution to Eqs. (15) and (18) is sought in the form

$$\rho(z, t) = \frac{3N}{4Z_0}\frac{1}{b_z(t)}\left(1 - \frac{z^2}{Z_0^2 b_z^2(t)}\right), \qquad v(z, t) = z\alpha_z(t), \tag{22}$$

where $b_z(t)$ and $\alpha_z(t)$ satisfy the conditions

$$b_z(0) = 1, \quad \alpha_z(0) = 0. \tag{23}$$

Substituting (22) into (15) and (18) yields

$$a_z(t) = \dot{b}_z(t)/b_z(t), \tag{24}$$

and the equation

$$\ddot{b}_z = \omega_z^2/b_z^2 \tag{25}$$

for $b_z(t)$. The latter equation can easily be integrated to obtain an implicit formula for b_z as a function of t:

$$\sqrt{2}\omega_z t = \sqrt{b_z(b_z - 1)} + \frac{1}{2}\ln[2b_z - 1 + 2\sqrt{b_z(b_z - 1)}] \tag{26}$$

(see [23], where this solution was applied to describe the quasi-one-dimensional initial stage of the condensate expansion that follows after a disc-shaped trap is switched off). The expression for $\alpha(t)$ in terms of $b_z(t)$ yields the velocity field:

$$v(z, t) = \frac{\sqrt{2}\omega_z z}{b_z(t)}\sqrt{1 - \frac{1}{b_z(t)}}. \tag{27}$$

The leading edge of the density distribution moves as

$$z_{\max}(t) = Z_0 b_z(t) \tag{28}$$

with the maximum velocity

$$v_{\max}(t) = \frac{dz_{\max}}{dt} = \sqrt{2}Z_0\omega_z\sqrt{1 - \frac{1}{b_z(t)}}. \tag{29}$$

At $t \gg \omega_z^{-1}$,

$$b_z(t) \approx \sqrt{2}\omega_z t, \quad t \gg \omega_z^{-1}, \tag{30}$$

the density and velocity distributions simplify to

$$\rho(z, t) \approx \frac{3N}{4v_{\max}t}\left(1 - \frac{z^2}{(v_{\max}t)^2}\right), \qquad v(z, t) \approx \frac{z}{t}, \quad t \gg \omega_z^{-1}, \tag{31}$$

and the maximum velocity tends to the constant value

$$v_{\max} \approx \sqrt{2}Z_0\omega_z, \quad t \gg \omega_z^{-1}. \tag{32}$$

These formulas describe inertial motion when the density is so small that the nonlinear pressure does not accelerate the gas any longer. Formula (32) is suitable for comparison with experiment, because the asymptotic value of the maximum velocity is expressed in terms of measurable parameters: the longitudinal trap frequency ω_z and the initial half-width Z_0 of the longitudinal Thomas–Fermi profile.

Expression (31) yields the asymptotic velocity distribution

$$\rho(v)dv = \frac{3N}{4v_{\max}}\left(1 - \frac{v^2}{v_{\max}^2}\right)dv, \quad |v| \le v_{\max}. \tag{33}$$

The mean kinetic energy is

$$\bar{E} = \frac{m}{2N}\int v^2\rho(v)dv = \frac{1}{5}E_{\max}, \qquad E_{\max} = \frac{1}{2}mv_{\max}^2. \tag{34}$$

2.2. Two-Dimensional Expansion

Two-dimensional condensate dynamics are observed when the longitudinal trap frequency ω_z is

much higher than the radial trap frequency $\omega_\perp$, i.e., when inequality (6) is replaced with the reverse one:

$$\lambda = \omega_z/\omega_\perp \gg 1. \quad (35)$$

Now, assume that the motion along the z axis is frozen, i.e., the zero-point energy associated with the oscillation amplitude $a_z = (\hbar/m\omega_z)^{1/2}$ is much higher than the nonlinear energy. By virtue of the estimate

$$N \sim |\psi|^2 R_0^2 a_z,$$

where R_0 is the radius of the density distribution in the plane (x, y) of the trap, this condition leads to the inequality

$$\frac{Na_s}{a_z} \ll \left(\frac{R_0}{a_z}\right)^2. \quad (36)$$

If it holds, the condensate wave function can again be factorized:

$$\psi(\mathbf{r}, t) = \phi(z)\Psi(x, y, t), \quad (37)$$

where

$$\phi(z) = \frac{1}{\pi^{1/4}a_z^{1/2}}\exp\left(-\frac{z^2}{2a_z^2}\right) \quad (38)$$

is the longitudinal ground-state wave function. By substituting (37) and (38) into (4) and (5) and integrating the result over the longitudinal coordinate, the action is expressed in terms of the effective two-dimensional Lagrangian density

$$\mathcal{L}_{2D} = \frac{i\hbar}{2}(\Psi_t^*\Psi - \Psi_t\Psi^*) + \frac{\hbar^2}{2m}(|\Psi_x|^2 + |\Psi_y|^2) + \frac{1}{2}m\omega_\perp^2(x^2 + y^2)|\Psi|^2 + \frac{g}{2\sqrt{2\pi}a_z}|\Psi|^4. \quad (39)$$

The corresponding Euler–Lagrange equation is the two-dimensional Gross–Pitaevskii equation

$$i\hbar\Psi_t = -\frac{\hbar^2}{2m}\Delta_\perp\Psi + \frac{1}{2}m\omega_\perp^2(x^2 + y^2)\Psi + g_{2D}|\Psi|^2\Psi, \quad (40)$$

where $\Delta_\perp = \partial_x^2 + \partial_y^2$ is the transverse Laplace operator, g_{2D} is an effective coupling constant expressed as

$$g_{2D} = \frac{g}{\sqrt{2\pi}a_z} = \frac{2\sqrt{2\pi}\hbar^2 a_s}{ma_z}, \quad (41)$$

and Ψ is normalized as

$$\int |\Psi|^2 dxdy = N. \quad (42)$$

Equation (40) describes the two-dimensional transverse dynamics of a condensate in a disc-shaped trap.

By the substitution

$$\Psi(\mathbf{r}_\perp, t) = \sqrt{\rho(\mathbf{r}_\perp, t)}\exp\left(\frac{im}{\hbar}\int^{\mathbf{r}_\perp}\mathbf{v}(\mathbf{r}'_\perp, t)d\mathbf{r}'\right), \quad (43)$$

where $\mathbf{r}_\perp = (x, y)$ and $\mathbf{v} = (v_x, v_y)$, Eq. (40) is transformed into the system

$$\rho_t + \nabla_\perp(\rho\mathbf{v}) = 0, \quad (44)$$

$$\mathbf{v}_t + (\mathbf{v}\nabla_\perp)\mathbf{v} + \frac{g_{2D}}{m}\nabla_\perp\rho + \omega_\perp^2\mathbf{r} + \frac{\hbar^2}{2m^2}\nabla_\perp\left[\frac{(\nabla_\perp\rho)^2}{4\rho^2} - \frac{\Delta_\perp\rho}{2\rho}\right] = 0, \quad (45)$$

where $\nabla_\perp = (\partial_x, \partial_y)$ is the transverse gradient operator. The quantum pressure can be neglected if it is much lower than the nonlinear pressure, i.e.,

$$1 \ll Na_s/a_z. \quad (46)$$

Then, Eq. (45) reduces to the hydrodynamic equation

$$\mathbf{v}_t + (\mathbf{v}\nabla_\perp)\mathbf{v} + \frac{g_{2D}}{m}\nabla_\perp\rho + \omega_\perp^2\mathbf{r} = 0. \quad (47)$$

The time-independent solution of the hydrodynamic equations is the Thomas–Fermi profile

$$\rho(r) = \frac{2N}{\pi R_0^2}\left(1 - \frac{r^2}{R_0^2}\right), \quad \mathbf{v} = 0, \quad (48)$$

where $r^2 = x^2 + y^2$ and the radius R_0 of the density distribution is determined by the number of atoms N:

$$R_0 = \left(\frac{16}{\sqrt{2\pi}}Na_s a_z^3\lambda^2\right)^{1/4}. \quad (49)$$

This inequality can be used to rewrite inequality (36) in a more convenient form, and the applicability condition for the two-dimensional Thomas–Fermi approximation becomes

$$1 \ll Na_s/a_z \ll \lambda^2. \quad (50)$$

After the transverse trapping potential is switched off, the condensate begins to expand radially, remaining bounded longitudinally. The radial expansion is described by hydrodynamic equations (44) and (47) subject to initial conditions (48). Now, the solution is sought in the form

$$\rho(r, t) = \frac{2N}{\pi R_0^2}\frac{1}{b_\perp^2(t)}\left(1 - \frac{r^2}{R_0^2 b_\perp^2(t)}\right), \quad v(r, t) = r\alpha_\perp(t), \quad (51)$$

where v is the radial velocity component, and $b_\perp(t)$ and

$\alpha_\perp(t)$ satisfy the initial conditions

$$b_\perp(0) = 1, \quad \alpha_\perp(0) = 0. \tag{52}$$

The substitution of (51) into (44) and (47) yields an equation relating $b_\perp(t)$ to $\alpha_\perp(t)$,

$$\alpha_\perp = \dot{b}_\perp/b_\perp,$$

and a differential equation for $b_\perp(t)$,

$$\ddot{b}_\perp = \omega_\perp^2/b_\perp^3. \tag{53}$$

The last equation is solved under the initial conditions

$$b_\perp(0) = 1, \quad \dot{b}_\perp(0) = b_\perp(0)\alpha_\perp(0) = 0$$

to obtain

$$b_\perp(t) = \sqrt{1 + \omega_\perp^2 t^2}, \tag{54}$$

and hence

$$\alpha_\perp(t) = \dot{b}_\perp/b_\perp = \omega_\perp^2 t/(1 + \omega_\perp^2 t^2).$$

Thus, simple expressions are obtained for the radial density and velocity distributions:

$$\rho(r, t) = \frac{2N}{\pi R_0^2}\frac{1}{1 + \omega_\perp^2 t^2}\left(1 - \frac{r^2}{R_0^2(1 + \omega_\perp^2 t^2)}\right),$$
$$v(r, t) = \frac{\omega_\perp^2 rt}{1 + \omega_\perp^2 t^2}. \tag{55}$$

The leading edge of the radial density distribution moves as

$$r_{\max}(t) = R_0\sqrt{1 + \omega_\perp^2 t^2} \tag{56}$$

with the maximum velocity

$$v_{\max}(t) = \frac{dr_{\max}}{dt} = \frac{R_0\omega_\perp^2 t}{\sqrt{1 + \omega_\perp^2 t^2}}. \tag{57}$$

At $t \gg \omega_\perp^{-1}$,

$$\rho(r, t) \approx \frac{2N}{\pi v_{\max} t^2}\left(1 - \frac{r^2}{(v_{\max} t)^2}\right),$$
$$v(r, t) \approx \frac{r}{t}, \tag{58}$$

where

$$v_{\max} \approx R_0\omega_\perp, \quad t \gg \omega_\perp^{-1}. \tag{59}$$

As in the one-dimensional case, these formulas describe inertial motion. Again, the maximum velocity is expressed in terms of measurable parameters, the radial frequency $\omega_\perp$ of the trap before it was switched off and the initial Thomas–Fermi radius R_0.

The asymptotic velocity distribution is

$$\rho(v)dv = \frac{4N}{v_{\max}^2}\left(1 - \frac{v^2}{v_{\max}^2}\right)v dv, \tag{60}$$

and the mean energy is

$$\bar{E} = \frac{1}{3}E_{\max}, \quad E_{\max} = \frac{1}{2}mv_{\max}^2. \tag{61}$$

3. THREE-DIMENSIONAL EFFECTS: A VARIATIONAL APPROACH

If condition (7) or (36) is violated, then motion along the smaller dimension of the condensate is not frozen and must be taken into account. As shown in [24], this can be done by applying a simple variational method. However, when the trap is highly anisotropic and the number of atoms in the condensate is sufficiently large, the trial distribution along the larger dimension should be the equilibrium Thomas–Fermi profile, which differs substantially from the Gaussian distribution assumed in [24]. This method was applied to calculate the ground states of condensates in [13]. Here, it is applied to condensate dynamics.

3.1. Cigar-Shaped Trap

In the case of a cigar-shaped trap with Thomas–Fermi axial density distribution, the variational condensate wave function has the form

$$\psi = A\exp\left(-\frac{r^2}{2w_\perp^2}\right)\sqrt{1 - \frac{z^2}{w_z^2}}$$
$$\times \exp\left[\frac{i}{2}(\alpha_\perp^2 r^2 + \alpha_z^2 z^2)\right], \tag{62}$$

where the parameters A, $w_\perp$, w_z, $\alpha_\perp$, and α_z are functions of time. It is assumed that $Na_s/Z_0 \sim 1$, i.e., the Thomas–Fermi limit radial profile is not reached and the radial wave function can be well approximated by a Gaussian distribution. The parameter A is related to the widths $w_\perp$ and w_z by normalization condition (3), which yields

$$A = \left(\frac{3N}{4\pi w_\perp^2 w_z}\right)^{1/2}. \tag{63}$$

Substituting (62) into (4) and (5) and integrating the result, one obtains the averaged Lagrangian

$$\frac{L}{N} = \left[\frac{\hbar^2}{2}\frac{d\alpha_\perp}{dt} + \frac{\hbar^2}{2m}\left(\frac{1}{w_\perp^4} + \alpha_\perp^2\right) + \frac{1}{2}m\omega_\perp^2\right]w_\perp^2$$
$$+ \left(\frac{\hbar^2}{2}\frac{d\alpha_z}{dt} + \frac{\hbar^2}{2m}\alpha_z^2 + \frac{1}{2}m\omega_\perp^2\lambda^2\right)\frac{w_z^2}{5} + \frac{3Na_s\hbar^2}{5m}\frac{1}{w_\perp^2 w_z}, \tag{64}$$

where the $1/w_z^4$ term is neglected. Indeed, this term is much less than α_z^2, because the estimates

$$w_z \sim Z_0, \quad \alpha_z \sim (m/\hbar)\dot{w}_z/w_z \sim m\omega_z/\hbar \sim \lambda/a_\perp^2$$

imply that the condition $1/w_z^2 \ll \alpha_z$ is equivalent to the condition $\lambda^{1/2} \ll Na_s/a_\perp$ for the applicability of the Thomas–Fermi approximation in the longitudinal direction (see (21)).

Lagrangian (64) entails the well-known expressions

$$\alpha_\perp = \frac{m}{\hbar}\frac{1}{w_\perp}\frac{dw_\perp}{dt}, \quad \alpha_z = \frac{m}{\hbar}\frac{1}{w_z}\frac{dw_z}{dt} \tag{65}$$

and the equations of motion for the widths

$$\ddot{w}_\perp + \omega_\perp^2 w_\perp = \frac{\hbar^2}{m^2 w_\perp^3} + \frac{6Na_s\hbar^2}{5m^2}\frac{1}{w_\perp^3 w_z}, \tag{66}$$

$$\ddot{w}_z + \lambda^2\omega_\perp^2 w_z = \frac{3Na_s\hbar^2}{m^2}\frac{1}{w_\perp^2 w_z^2}. \tag{67}$$

These equations differ from those derived in [24] by numerical factors and by the absence of the term corresponding to longitudinal quantum pressure (it would be incorrect to retain this term in the approximation employed here). In the dimensionless variables

$$b_\perp = w_\perp/a_\perp, \quad b_z = w_z/Z_0, \quad \tau = \omega_\perp t \tag{68}$$

Eqs. (66) and (67) are rewritten as

$$\frac{d^2 b_\perp}{d\tau^2} + b_\perp = \frac{1}{b_\perp^3} + \frac{2q}{5}\frac{1}{b_\perp^3 b_z}, \tag{69}$$

$$\frac{d^2 b_z}{d\tau^2} + \lambda^2 b_z = \frac{\lambda^2}{b_\perp^2 b_z^2}, \tag{70}$$

where the parameter

$$q = (\lambda Z_0/a_\perp)^2 \tag{71}$$

characterizes the radial nonlinear pressure. If $q \ll 1$, then the second term on the right-hand side of (69) can be neglected to obtain the time-independent solution $b_{\perp 0} = b_{z0} = 1$, which corresponds to the one-dimensional approximation (see Section 2). In this case, Eq. (70) describing free longitudinal expansion obviously reduces to Eq. (25).

The equilibrium values of $b_\perp$ and b_z are determined by the equations

$$b_{\perp 0} = b_{z0}^{-3/2}, \quad \frac{1}{b_{z0}^6} = 1 + \frac{2q}{5}\frac{1}{b_{z0}}, \tag{72}$$

which differ from analogous equations obtained in [21] only by notation. They describe the state of the condensate before expansion.

The condensate expansion that follows after the longitudinal potential is switched off is described by the equations

$$\frac{d^2 b_\perp}{d\tau^2} + b_\perp = \frac{1}{b_\perp^3} + \frac{2q}{5}\frac{1}{b_\perp^3 b_z}, \quad \frac{d^2 b_z}{d\tau^2} = \frac{\lambda^2}{b_\perp^2 b_z^2}. \tag{73}$$

They can readily be solved numerically under the initial conditions

$$\begin{aligned} b_\perp(0) &= b_{\perp 0}, \quad \dot{b}_\perp(0) = 0, \\ b_z(0) &= b_{z0}, \quad \dot{b}_z(0) = 0, \end{aligned} \tag{74}$$

where $b_{\perp 0}$ and b_{z0} are determined by (72). Figure 1 shows the functions of time

$$b_\perp = w_\perp/a_\perp, \quad z_{\max} = Z_0 b_z(t) = Z b_z(t)/b_{z0},$$

where $Z = Z_0 b_{z0}$ is the initial half-width of the condensate, for several values of N. When

$$\sqrt{\lambda}a_\perp/a_s \ll N \ll a_\perp/(\lambda a_s),$$

we have the analytical solution obtained in Section 2. If

$$N > a_\perp/(\lambda a_s),$$

then the expansion starts from a radius larger than the zero-point oscillation amplitude $a_\perp$, and $b_\perp \longrightarrow 1$ as $\tau \longrightarrow \infty$. At $\tau \gg 1$, the longitudinal expansion approaches an inertial motion characterized by constant velocities of atoms. The maximum velocity is readily found by using the conservation law corresponding to Eq. (73):

$$\begin{aligned} \frac{1}{2}\left(\frac{db_z}{d\tau}\right)^2 + \frac{5\lambda^2}{2q}\left[\left(\frac{db_\perp}{d\tau}\right)^2 + b_\perp^2 + \frac{1}{b_\perp^2}\right] \\ + \frac{\lambda^2}{b_\perp^2 b_z} = \text{const}. \end{aligned} \tag{75}$$

At $\tau = 0$, it yields the initial values in (74). As $\tau \longrightarrow \infty$,

$$db_z/d\tau \longrightarrow \dot{b}_{z,\max}, \quad b_\perp \longrightarrow 1, \quad b_z \longrightarrow \infty,$$

and therefore

$$\dot{b}_{z,\max} = \lambda\sqrt{\frac{2}{b_{\perp 0}^2 b_{z0}} + \frac{5}{q}\left(b_{\perp 0} - \frac{1}{b_{\perp 0}}\right)^2}. \tag{76}$$

Then, Eqs. (72) are used to find

$$v_{\max} = \left.\frac{dz_{\max}}{dt}\right|_{t\to\infty} = \frac{2Z\omega_z b_{\perp 0}}{\sqrt{1 + b_{\perp 0}^2}}. \tag{77}$$

When $q \ll 1$, i.e., $b_{\perp 0} = 1$ and $Z = Z_0$, this expression

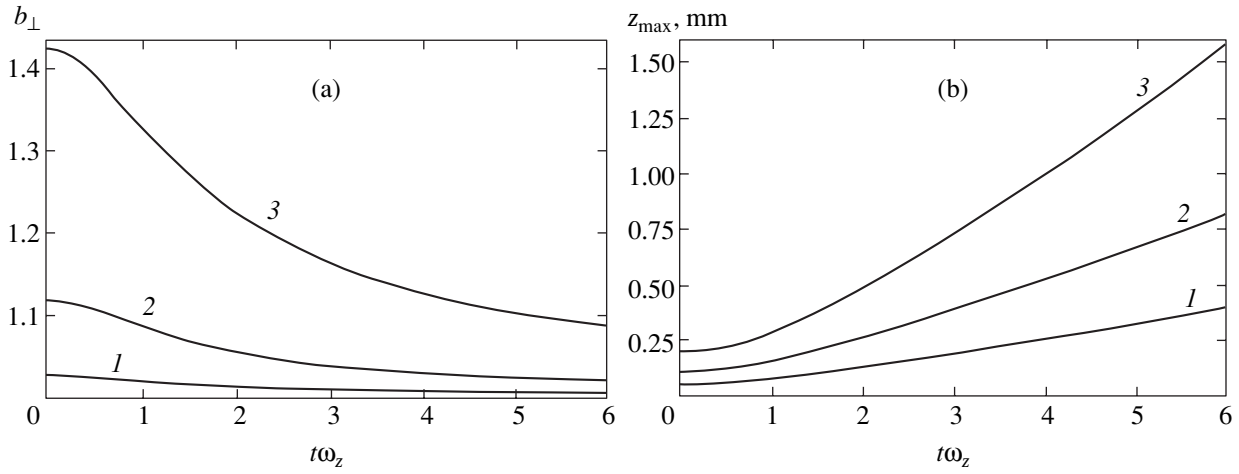


Fig. 1. Radius $b_\perp = w_\perp/a_\perp$ (a) and longitudinal size z_{max} (b) vs. time $\tau = t\omega_z$ for a one-dimensional expanding condensate in a cigar-shaped trap at $a_\perp = 5$ μm, $a_s = 5$ nm, and $\lambda = 0.05$. Curves *1*, *2*, and *3* correspond to $N = 10^3$, 10^4, and 10^5, respectively.

yields the one-dimensional velocity given by (32). When $q \gg 1$, i.e., $b_{\perp 0} \gg 1$, it reduces to

$$v_{max} \approx 2\omega_z Z.$$

Thus, not only the longitudinal condensate size increases, but also the ratio of v_{max} to $\omega_z R$ increases from $\sqrt{2}$ to 2 as the number of atoms in the condensate increase.

Both asymptotic velocity distribution and expression for the mean energy retain their form given by (33) and (34), respectively, where v_{max} is now given by (77). Figure 2 illustrates the dependence of the mean energy on N.

3.2. Disc-Shaped Trap

In the case of a disc-shaped trap with Thomas–Fermi radial density distribution, the variational wave function has the form

$$\psi = A\sqrt{1 - \frac{r^2}{w_\perp^2}}\exp\left(-\frac{z^2}{2w_z^2}\right) \times \exp\left[\frac{i}{2}(\alpha_\perp^2 r^2 + \alpha_z^2 z^2)\right], \quad (78)$$

where the time-dependent parameters A, $w_\perp$, w_z, $\alpha_\perp$, and α_z are related by normalization condition (3):

$$A = \left(\frac{2N}{\pi^{3/2}}\frac{1}{w_\perp^2 w_z}\right)^{1/2}. \quad (79)$$

Substituting (78) into (4) and (5) and integrating the result, one obtains the Lagrangian

$$\frac{L}{N} = \left(\frac{\hbar^2}{2}\frac{d\alpha_\perp}{dt} + \frac{\hbar^2}{2m}\alpha_\perp^2 + \frac{1}{2}m\omega_\perp^2\right)\frac{w_\perp^2}{3} + \left[\frac{\hbar^2}{2}\frac{d\alpha_z}{dt} + \frac{\hbar^2}{2m}\left(\frac{1}{w_z^4} + \alpha_z^2\right) + \frac{1}{2}m\omega_\perp^2\lambda^2\right]\frac{w_z^2}{2} + \frac{8}{3\sqrt{2\pi}}\frac{Na_s\hbar^2}{m}\frac{1}{w_\perp^2 w_z}, \quad (80)$$

where the $1/w_\perp^4$ term is neglected, because this term is

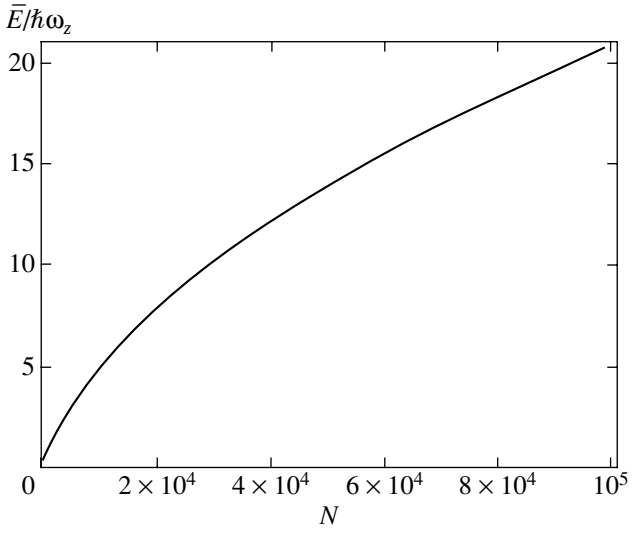


Fig. 2. Mean energy of atoms in a condensate (measured in energy quanta $\hbar\omega_z$ of longitudinal oscillations) after one-dimensional expansion in a cigar-shaped trap vs. number of atoms.

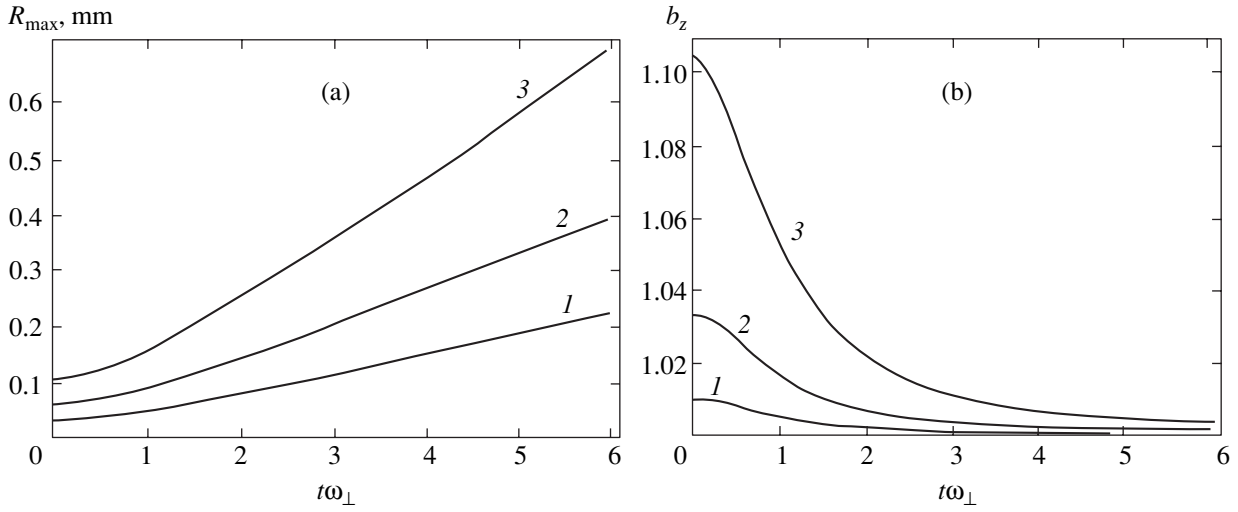


Fig. 3. Radius R_{max} (a) and longitudinal size $b_z = w_z/a_z$ (b) vs. time $\tau = t\omega_\perp$ for a two-dimensional expanding condensate in a disc-shaped trap at $a_z = 5$ μm, $a_s = 5$ nm, and $\lambda = 20$. Curves *1*, *2*, and *3* correspond to $N = 10^3$, 10^4, and 10^5, respectively.

much less than $\alpha_\perp^2$ by virtue of the applicability condition $Na_s/a_z \gg 1$ for the radial Thomas–Fermi approximation.

Lagrangian (80) entails expressions (65) and the equations of motion for the widths

$$\ddot{w}_\perp + \omega_\perp^2 w_\perp = \frac{16}{\sqrt{2\pi}}\frac{Na_s\hbar^2}{m^2}\frac{1}{w_\perp^3 w_z}, \quad (81)$$

$$\ddot{w}_z + \lambda^2\omega_\perp^2 w_z = \frac{\hbar^2}{m^2 w_z^3} + \frac{16}{\sqrt{2\pi}}\frac{Na_s\hbar^2}{m^2}\frac{1}{w_\perp^2 w_z^2}, \quad (82)$$

which are rewritten in the dimensionless variables

$$b_\perp = w_\perp/R_0, \quad b_z = w_z/a_z, \quad \tau = \omega_\perp t, \quad (83)$$

where $a_z = a_\perp/\sqrt{\lambda}$, as

$$\frac{d^2 b_\perp}{d\tau^2} + b_\perp = \frac{1}{b_\perp^3 b_z}, \quad (84)$$

$$\frac{d^2 b_z}{d\tau^2} + \lambda^2 b_z = \frac{\lambda^2}{b_z^3} + \frac{\lambda^2 q}{3}\frac{1}{b_\perp^2 b_z^2}. \quad (85)$$

The parameter

$$q = (R_0/\lambda a_z)^2 \quad (86)$$

characterizes the longitudinal nonlinear pressure. If $q \ll 1$, then the second term on the right-hand side of (85) can be neglected to obtain the time-independent solution

$$b_\perp = b_z = 1,$$

which corresponds to the two-dimensional approximation considered in Section 2. In this case, Eq. (84) obviously reduces to (53).

The equilibrium values of $b_\perp$ and b_z are determined by the equations

$$b_{z0} = b_{\perp 0}^{-4}, \quad \frac{1}{b_{\perp 0}^{16}} = 1 + \frac{q}{3}\frac{1}{b_{\perp 0}^6}, \quad (87)$$

which again are identical to the equations obtained in [21], except for notation. They describe the state of the condensate before expansion. The expansion in the plane of the trap is described by the equations

$$\frac{d^2 b_\perp}{d\tau^2} = \frac{1}{b_\perp^3 b_z}, \quad \frac{d^2 b_z}{d\tau^2} + \lambda^2 b_z = \frac{\lambda^2}{b_z^3} + \frac{\lambda^2 q}{3}\frac{1}{b_\perp^2 b_z^2}. \quad (88)$$

They can readily be solved numerically under initial conditions (74), where $b_{\perp 0}$ and b_{z0} are determined by (87). Figure 3 shows the functions of time

$$R_{max} = Rb_\perp(t)/b_{\perp 0}, \quad b_z = w_z/a_z,$$

where $R = R_0 b_{\perp 0}$ is the initial radius of the condensate, for several values of N. When

$$a_z/a_s \ll N \ll a_z\lambda^2/a_s,$$

we have the analytical solution obtained in Section 2. If

$$N > a_z\lambda^2/a_s$$

then the expansion starts from a longitudinal size w_{z0} larger than the zero-point oscillation amplitude a_z (i.e., $b_{z0} > 1$), and $b_z \longrightarrow 1$ as $\tau \longrightarrow \infty$. At $\tau \gg 1$, the radial expansion approaches an inertial motion characterized by constant velocities of atoms. The maximum velocity

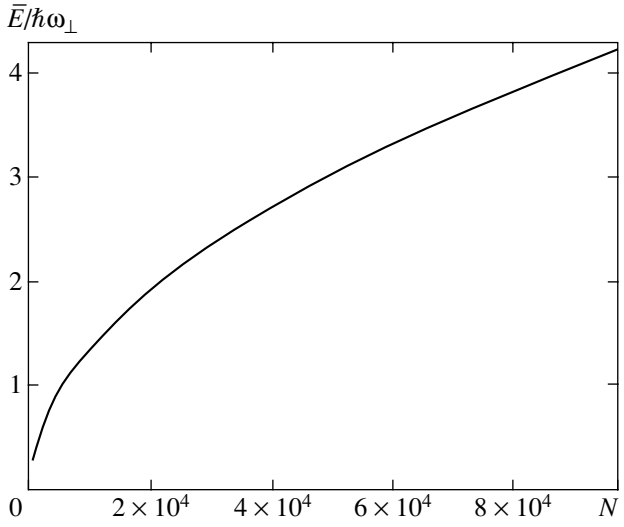


Fig. 4. Mean energy of atoms in a condensate (measured in energy quanta $\hbar\omega_\perp$ of longitudinal oscillations) after two-dimensional expansion in a disc-shaped trap vs. number of atoms.

is found by using the conservation law corresponding to Eq. (88):

$$\left(\frac{db_\perp}{d\tau}\right)^2 + \frac{3}{\lambda^2 q}\left[\frac{1}{2}\left(\frac{db_z}{d\tau}\right)^2 + \frac{\lambda^2}{2}\left(b_z^2 + \frac{1}{b_z^2}\right)\right] + \frac{1}{b_\perp^2 b_z} = \text{const}, \tag{89}$$

which yields

$$\dot{b}_{\perp,\max} = \sqrt{\frac{1}{b_{\perp 0}^2 b_{z0}} + \frac{3}{2q}\left(b_{z0} - \frac{1}{b_{z0}}\right)^2}. \tag{90}$$

Accordingly,

$$v_{\max} = \left.\frac{dR_{\max}}{dt}\right|_{t\to\infty} = \omega_\perp R\sqrt{\frac{1+3b_{z0}^2}{2(1+b_{z0}^2)}}. \tag{91}$$

When $q \ll 1$, i.e., $b_{z0} = 1$ and $R = R_0$, this expression yields the two-dimensional velocity given by (59). When $q \gg 1$, i.e., $b_{z0} \gg 1$, it reduces to

$$v_{\max} \approx \sqrt{3/2}\,\omega_\perp R.$$

Both the asymptotic velocity distribution and expression for the mean energy retain their form given by (60) and (61), respectively, where $v_{\max}$ is now given by (91). Figure 4 illustrates the dependence of the mean energy on N.

4. DISCUSSION

Let us compare the theoretical results with available experimental data.

An experimental study of condensate expansion in a quasi-one-dimensional trap was reported in [1], where the longitudinal half-length of the condensate was $Z \approx 100$ μm and the time of transition to inertial motion was approximately 20 ms, which corresponds to the trap frequency $w_z \approx 50$ s^{-1}. Under these conditions, expression (32) yields

$$v_{\max} \approx \sqrt{2}Z\omega_z \approx 7 \text{ mm/s}.$$

In view of the assumptions made, good agreement between this estimate and the experimental value $v_{\max} \approx 5.9$ mm/s is achieved without introducing any adjustable parameters.

Expansion of a Bose–Einstein condensate in a system of two-dimensional disc-shaped traps was investigated in [2]. The results of that study showed that no longitudinal expansion took place, and the condensate was effectively divided into separate condensate "pancakes" confined in separate potential wells. Therefore, the two-dimensional theory can be applied to describe the radial expansion of each particular condensate. According to [2], the maximum radial velocity was $v_{\max} \approx 1.5$–1.7 mm/s when the initial radius was $R \approx 13$ μm and the radial trap frequency was $\omega_\perp \approx 132$ s^{-1}. Expression (59) predicts

$$v_{\max} \approx R\omega_\perp \approx 1.7 \text{ mm/s},$$

which is in good agreement with the experimental value. Thus, the mean-field theory provides a good description of condensate expansion in fewer coordinates as well. Deviations from the theoretical predictions would point to the existence of a condensate that could not be described by the mean-field theory.

ACKNOWLEDGMENTS

I thank F.Kh. Abdullaev, V.A. Brazhnyi, A. Gammal, V.V. Konotop, R.A. Kraenkel, and L. Tomio for valuable discussions. This work was supported by the Russian Foundation for Basic Research, project no. 01-01-00696.

REFERENCES

1. K. Bongs, S. Burger, S. Dettmer, *et al.*, Phys. Rev. A **63**, 031602 (2001).
2. O. Morsch, M. Cristiani, J. H. Müller, *et al.*, Phys. Rev. A **66**, 021601 (2002).
3. A. Görlitz, J. M. Vogels, A. E. Leanhardt, *et al.*, Phys. Rev. Lett. **87**, 130402 (2001).
4. M. Greiner, I. Bloch, O. Mandel, *et al.*, Phys. Rev. Lett. **87**, 160405 (2001).
5. S. Dettmer, D. Hellweg, P. Ryytty, *et al.*, Phys. Rev. Lett. **87**, 160406 (2001).

6. A. E. Leanhardt, A. P. Chikkatur, D. Kielpinski, *et al.*, Phys. Rev. Lett. **89**, 040401 (2002).
7. D. S. Petrov, M. Holzmann, and G. V. Shlyapnikov, Phys. Rev. Lett. **84**, 2551 (2000).
8. D. S. Petrov, G. V. Shlyapnikov, and J. T. M. Walraven, Phys. Rev. Lett. **85**, 3745 (2000).
9. V. Dunjko, V. Lorent, and M. Olshanii, Phys. Rev. Lett. **86**, 5413 (2001).
10. P. Pedri, L. Pitaevskii, S. Stringari, *et al.*, Phys. Rev. Lett. **87**, 220401 (2001).
11. C. Menotti and S. Stringari, Phys. Rev. A **66**, 043610 (2002).
12. A. Gammal, L. Tomio, and T. Frederico, Phys. Rev. A **66**, 043619 (2002).
13. K. K. Das, M. D. Girardeau, and E. M. Wright, Phys. Rev. Lett. **89**, 110402 (2002).
14. L. Pitaevskii and S. Stringari, *Bose–Einstein Condensation* (Clarendon Press, Oxford, 2003).
15. Y. Castin and R. Dum, Phys. Rev. Lett. **77**, 5315 (1996).
16. Yu. Kagan, G. V. Shlyapnikov, and J. T. M. Walraven, Phys. Rev. Lett. **76**, 2670 (1996).
17. Yu. Kagan, E. L. Surkov, and G. V. Shlyapnikov, Phys. Rev. A **54**, R1753 (1996).
18. F. Dalfovo, C. Minniti, S. Stringari, and L. Pitaevskii, Phys. Lett. A **227**, 259 (1997).
19. J. A. Stickney and A. A. Zozulya, Phys. Rev. A **65**, 053612 (2002).
20. L. Plaja and L. Santos, Phys. Rev. A **65**, 035602 (2002).
21. K. K. Das, Phys. Rev. A **66**, 053612 (2002).
22. V. I. Talanov, Pis'ma Zh. Éksp. Teor. Fiz. **2**, 218 (1965) [JETP Lett. **2**, 138 (1965)].
23. V. A. Brazhnyi, A. M. Kamchatnov, and V. V. Konotop, Phys. Rev. A **68**, 035603 (2003).
24. V. M. Pérez-García, H. Michinel, J. I. Cirac, *et al.*, Phys. Rev. Lett. **77**, 5320 (1996); Phys. Rev. A **56**, 1424 (1997).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 918–923.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1052–1057.

ATOMS, SPECTRA, RADIATION

Cross Sections of Electron Capture by Boron Ions in Gaseous Media

I. S. Dmitriev, Ya. A. Teplova*, Yu. A. Belkova, and Yu. A. Faĭnberg

Skobel'tsyn Research Institute of Nuclear Physics, Moscow State University, Moscow, 119899 Russia
**e-mail: teplova@anna19.npi.msu.su*

Received December 3, 2003

Abstract—We study the cross sections $\sigma_{i,i-1}$, $\sigma_{i,i-2}$, and $\sigma_{i,i-3}$ of capture of one, two, and three electrons by boron ions with charges i = 1–5 and velocities $V = (1.83–5.50)V_0$ in gaseous media with atomic numbers Z_t varying from 1 to 54. The oscillatory form of the Z_t dependence of electron capture cross section by boron ions, which has been established for lighter ions, is confirmed. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Theoretical estimates of electron capture cross sections $\sigma_{i,i-1}$ by various ions of light elements, based on the Thomas–Fermi statistical model, proved that the cross sections increase monotonically with the nuclear charge Z_t of atoms in the medium [1, 2]. Analysis of experimental data in the range of relatively low velocities $V = (1–2)V_0$ ($V_0 = 2.19 \times 10^8$ cm/s), for which the electron capture cross sections attain their maximal value, also indicates that cross sections $\sigma_{i,i-1}$ on the average increase monotonically in proportion to $Z_t^{1/3}$ [2]. However, subsequent experimental studies revealed certain deviations from this type of dependence of $\sigma_{i,i-1}$ on Z_t. The cross sections of electron capture by boron ions with velocity $V = 1.83V_0$ in neon, determined from experimental data, are smaller than the corresponding values in nitrogen by a factor of 1.5–2 [3]. Analogous deviations are also observed for higher velocities of ions. Analysis of the electron capture cross section by nitrogen and neon ions for $V = 3.65V_0$ indicates that the corresponding cross sections in neon are considerably higher than in nitrogen and argon [4, 5]. Investigations of electron capture cross sections by hydrogen and helium ions in various media in a wide range of ion velocities proved that the variation of the value of cross sections $\sigma_{i,i-1}$ upon an increase in Z_t has the form of oscillations [6, 7]. Thus, we can assume that the actual Z_t dependence of the electron capture cross section by light ions in the velocity range $V > 2V_0$ substantially differs from a monotonic dependence, which is apparently due to the shell structure of atoms of the medium.

In this study, we investigate the effect of structural features of atoms of the medium on the electron capture cross sections $\sigma_{i,i-1}$ measured for boron ions with charges i = 1–5 for velocities $V = (1.83–5.50)V_0$ in gases H_2, He, N_2, Ne, Ar, Kr, and Xe; we also calculate theoretically the electron capture cross sections by the same ions in media with Z_t = 1–54. This work continues our experiments with boron ions reported in [3]. The ion velocities V will be given in atomic units of V_0.

2. EXPERIMENTAL TECHNIQUE

Experiments with boron ions were carried out on the setup described in [3, 8]. Boron ions ^{11}B with charges i = 1, 2, and 3 and velocities V = 1.83, 2.74, 3.65, and 5.50 were extracted from an accelerator and directed to the collision chamber after their passage through a recharging device, where the primary ions were transformed into beams of boron atoms B^{i+} with charges i = 1–5. The charge distribution of boron ions after their passage through the collision chamber was measured with the help of a detection system consisting of a magnetic analyzer and a block of identical counters. The capture cross sections $\sigma_{i,i-m}$ for one, two, and three electrons, where m = 1, 2, 3, were determined from the obtained charge distributions using the method described in [9]. The errors in determining the cross sections were mainly determined by the errors in estimating the gas layer thickness in the collision chamber (~10%) and by the statistical spread in the results of several sets of measurements, which amounted on average to 10–15% in the capture cross sections for one electron, 20–30% in the capture cross sections for two electrons, and 50–70% in the capture cross section for three electrons. The cross sections of electron capture by boron ions for all above-mentioned velocities in atomic and molecular hydrogen [10] and in He, N_2, Ar, as well as Kr, were measured earlier [9, 11]. All cross sections $\sigma_{i,i-m}$ for boron ions in Ne and Xe were determined for the first time in this study. The values of cross sections $\sigma_{i,i-m}$ for ion velocities V = 2.74, 3.65, and 5.50 are given in the table. Since the values of cross sections obtained in the present study and in [9, 11] coin-

Cross sections $\sigma_{i,\,i-m}$ of electron capture (in units of 10^{-16} cm^2/at)

i, $i-m$	H_2	He	N_2	Ne	Ar	Kr	Xe
			$V = 2.75$				
1, 0	0.018	0.09	0.04	0.19	0.10		
2, 1	0.08	0.36	0.43	0.85	0.63	0.9	
2, 0		0.0036	0.3	0.018	0.004		
3, 2	0.29	0.8	1.0	1.60	1.7	2.3	
3, 1	0.0017	0.04	0.04	0.07	0.05	0.1	
4, 3	0.54	2.0	2.1	3.6	4.0		
4, 2		0.05	0.1	0.24	0.22		
4, 1				0.02	0.01		
			$V = 3.65$				
1, 0			0.012				0.03
2, 1	0.008	0.05	0.09	0.18	0.10	0.13	0.16
2, 0			0.01				0.01
3, 2	0.016	0.12	0.24	0.5	0.35	0.50	0.80
3, 1		0.003	0.012	0.007		0.015	0.03
4, 3	0.05	0.40	0.80	1.9	1.0	1.60	2.4
4, 2		0.03	0.03	0.09	0.063	0.14	0.2
4, 1			0.005	0.01		0.03	0.012
5, 4	0.12	0.70	1.4	3.0	1.80	2.1	3.0
5, 3		0.2	0.16	0.35	0.30	0.60	0.71
5, 2			0.03	0.04		0.1	0.1
			$V = 5.50$				
3, 2	0.002	0.008	0.03	0.13	0.035		
3, 1				0.001			
4, 3	0.0043	0.03	0.13	0.30	0.35	0.48	
4, 2		0.0002	0.0035	0.004	0.016	0.027	
5, 4	0.009	0.07	0.25	0.70	0.50	0.83	
5, 3	0.0006	0.0012	0.045	0.055	0.18		
5, 2			0.0008	0.006	0.014		

cided to within 20–30%, the table contains the averaged values of cross sections.

3. THEORETICAL MODELS FOR CALCULATING ELECTRON CAPTURE CROSS SECTIONS

In accordance with the analysis carried out in [3], the total cross section of electron capture by an ion with charge i and velocity $V \sim 1$–2 is proportional to the number of vacancies $p_{i-1}(n)$ in the unfilled electron shell closest to the nucleus and characterized by the principal quantum number n:

$$\sigma_{i,\,i-1} = p_{i-1}(n)\pi a_0^2 \frac{I_{i-1}}{I_0} Z_t^{1/3} V^{-3}, \qquad (1)$$

where I_{i-1} is the binding energy of the electron being captured, $I_0 = 13.6$ eV, and $a_0 = 5.29 \times 10^{-9}$ cm.

For boron ions, the values of $p_{i-1}(n)$ decrease from $p_2(2) = 8$ for B^{3+} ions to $p_0(2) = 6$ for B^+ ions. As the ion velocity subsequently increases to $V \sim 3$–5, the values

of cross sections $\sigma_{i,i-1}$ can be estimated from the Bohr formula for the cross sections of electron capture by light ions in media with $Z_t \geq 2$ [1, 2, 12]:

$$\sigma_{i,i-1} = 4\pi a_0^2 i^3 Z_t^{1/3} V^{-5}. \quad (2)$$

In the case of electron capture by a fast ion from the filled shell n_t of an atom of the medium to the ion shell n_f, the partial cross section $\sigma_{i,i-1}(n_t \longrightarrow n_f)$ can be calculated in the Oppenheimer–Brinkman–Kramers (OBK) approximation [13, 14]. The values of $\sigma_{i,i-1}^{OBK}$ were calculated for the cross sections of electron capture by boron ions with charges i = 1–5 and velocities V = 1.83, 2.74, 3.65, and 5.50 in media with Z_t = 1–54. The method for calculating cross sections is analogous to that proposed in [6] for helium ions.

The main drawback in the OBK approximation is that it gives exaggerated values of calculated electron capture cross sections as compared to the experimental values. For this reason, the calculated values of cross sections were normalized to the experimental data obtained for boron ions for the above velocities in nitrogen. The normalization coefficient

$$R(V, Z) = \frac{\sigma_{i,i-1}^{\exp}}{\sigma_{i,i-1}^{OBK}}$$

for velocities V = 1.83, 2.74, 3.65, and 5.50 for boron ions is $R(V, 5)$ = 0.03, 0.06, 0.09, and 0.15, respectively. It is known from previous studies that the value of $R(V, 1)$ for hydrogen (Z = 1) varies from 0.13 for V = 3.65 to 0.18 for V = 5.50 [6]. For helium ions (Z = 2), the value of $R(V, 2)$ increases from 0.04 for V = 1.2 to 0.18 for V = 3.65–10 [7]. These data have made it possible to represent the normalization coefficient in the form of the empirical expression

$$R(V, Z) = 0.295[(0.54V/Z)^{-1.6} + 1]^{-1}. \quad (3)$$

The values of $R(V, Z)$ determined using formula (3) are in accordance with the most of the above values of $R(V, Z)$ to within ±30%. Formula (3) is a modification of the analogous expression from [14].

4. ANALYSIS OF OBTAINED RESULTS

Figure 1 shows the values of the electron capture cross sections $\sigma_{i,i-1}$ as a function of Z_t, obtained in experiments and calculated using various approximations. For velocity V = 1.83, cross sections $\sigma_{i,i-1}$ for boron ions with charges i = 2, 3 increase with Z_t in proportion to $Z_t^{1/3}$. The values of cross sections determined on the basis of formula (1) are in satisfactory agreement with the experimental data. A noticeable discrepancy is observed only for cross sections of electron capture by B^{2+} ions in nitrogen and neon. The calculations of the electron capture cross sections carried out in the OBK approximation lead to a qualitatively different dependence, which is not monotonic, but has the form of alternating maxima and minima with the general increase in the value of $\sigma_{i,i-1}$ with Z_t in proportion to $Z_t^{1/3}$. Calculations carried out in the OBK approximation make it possible to describe the deviation of the electron capture cross section from those calculated by formula (1) for nitrogen and neon (see Fig. 1a). As the ion velocity increases, the dependence of the electron capture cross section on Z_t becomes essentially nonmonotonic. The deviation of the experimental data from the results calculated by formula (2) for an ion velocity of V = 3.65 increases, especially upon an increase in the ion charge. The applicability of the OBK approximation is more justified for increasing ion velocity [6]; Figs. 1b and 1c clearly demonstrate the improvement in matching calculated electron capture cross sections with experimental data for V = 5.5.

The results of this study show that maxima are formed due to the shell structure of atoms of the medium and the resonant nature of the dependence of total cross sections $\sigma_{i,i-1}$ on Z_t. Each of these maxima corresponds to the maximal contribution of partial electron capture cross sections from individual atomic shells *K*, *L*, and *M* of the medium. The results presented in Fig. 1 confirm this conclusion for boron ions. In the framework of the OBK approximation, it is also possible to explain the shift of the maxima towards high values of Z_t with increasing ion velocity. Indeed, the OBK cross sections attain their maximal values when the electron binding energy in one of the atomic shells of the medium becomes equal to $I_t = (I_V + 2I_{\text{ion}})/3$, where I_t and I_{ion} are the electron binding energy in the target atom and the fast ion, respectively, and $I_V = \mu V^2/2$ is the energy transferred by an electron [6, 7]. With increasing ion velocity, the value of I_t for which the electron capture cross section is maximal increases, leading to the corresponding increase in the value of Z_t at which the given maximum is observed.

The emergence of oscillations in the dependence of the electron capture cross section on Z_t is mainly due to the enhancement of the effect of internal and external screening on the average orbital velocity of electrons from outer shells of atoms in the medium upon an increase in the principal quantum number *n*. This leads to an increase in the relative difference between the cross sections for states with adjacent values of *n*. Incomplete filling of outer electron shells deepens the minima in the dependence of cross sections $\sigma_{i,i-1}$ on Z_t. It can be seen from the above results that the largest amplitude of oscillations (i.e., the difference between the maximal and minimal values of cross sections) for boron ions in the range of Z_t = 7–18 (N, Ne, Ar) is observed for an ion velocity of V = 3.65. In this case,

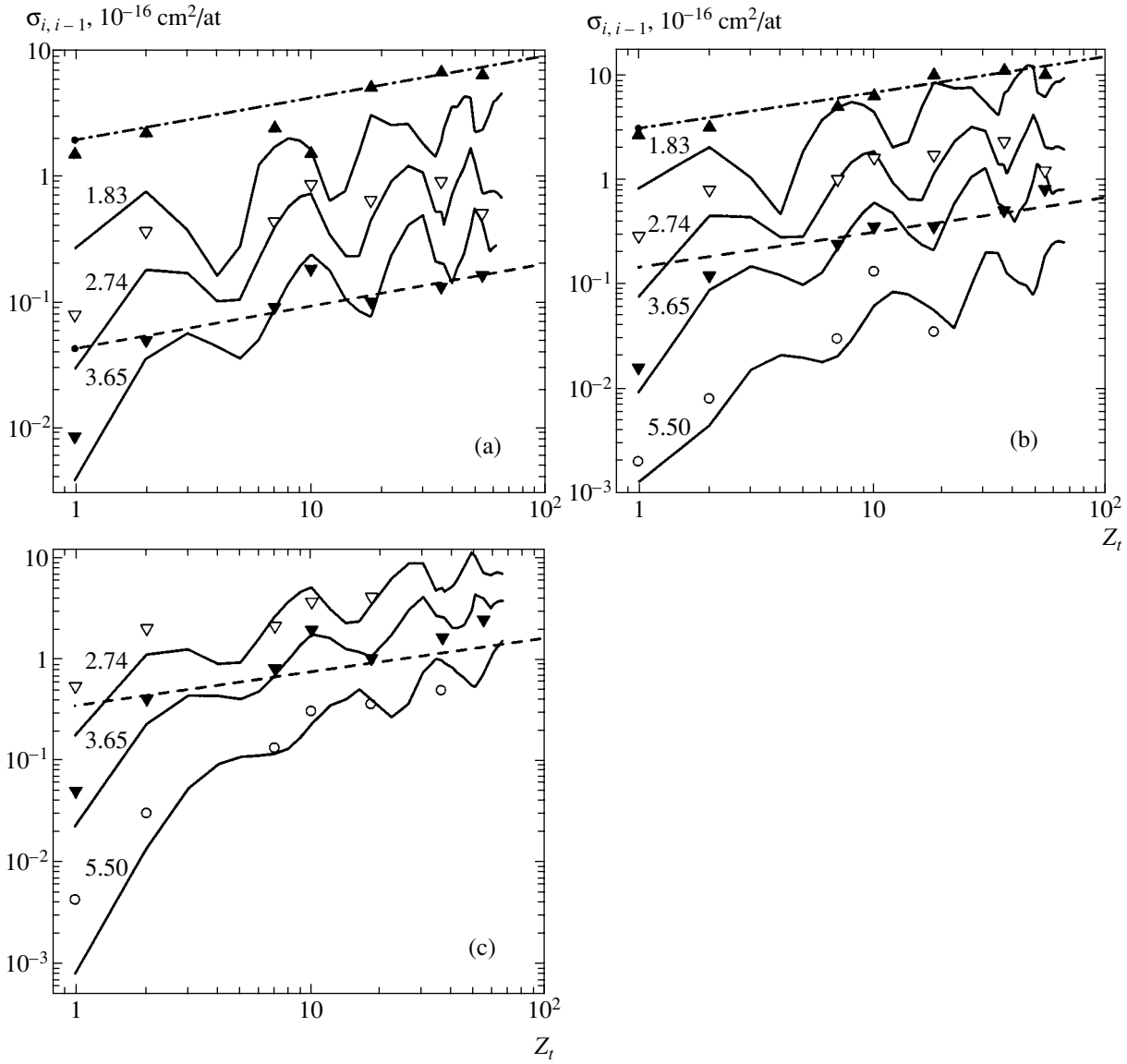


Fig. 1. Dependence of cross sections $\sigma_{i,i-1}$ of electron capture by boron ions (a) B^{2+}, (b) B^{3+}, and (c) B^{4+} on Z_t. Symbols correspond to experimental values; solid curves describe the results of normalized calculations in the OBK approximation; dot-and-dash lines correspond to calculations by formula (1) for $V = 1.83$, and dashed lines, to calculations by formula (2) for $V = 3.65$. The values of velocities V (in units of V_0) are shown on the curves.

the amplitude of oscillations decreases with increasing boron ion charge and attains its lowest value for nuclei of B^{5+} (Figs. 1 and 2) With a further increase in the velocity of ions, the oscillatory dependence of the cross sections on Z_t is gradually weakened and becomes virtually monotonic again for $V > 20$ [6].

Figure 2 shows for comparison the available experimental and theoretical values of cross sections $\sigma_{Z,Z-1}$ of electron capture by nuclei of hydrogen H^+ [6], helium He^{2+} [7], boron B^{5+}, and nitrogen N^{7+} [5] for a velocity of $V = 3.65$ in various media. In the range of $Z_t = 2$–54, the values of $\sigma_{Z,Z-1}$ increase on average in proportion to $Z_t^{1/3}$, approximately in accordance with the results of calculations based on the Bohr formula (2). However, the figure shows the deviation of the experimental data from this dependence. For example, for $Z_t = 1$, the values of cross sections $\sigma_{Z,Z-1}$ for the nuclei in question differ from those calculated by formula (2) approximately by an order of magnitude. In contrast to the model of independent electrons, the calculations carried out in the OBK approximation are in qualitative agreement with the experimental dependence of cross sections $\sigma_{Z,Z-1}$ on Z_t (including the case of $Z_t = 1$). The above results show that the dependence of the electron capture cross sections on Z_t for the nuclei in question

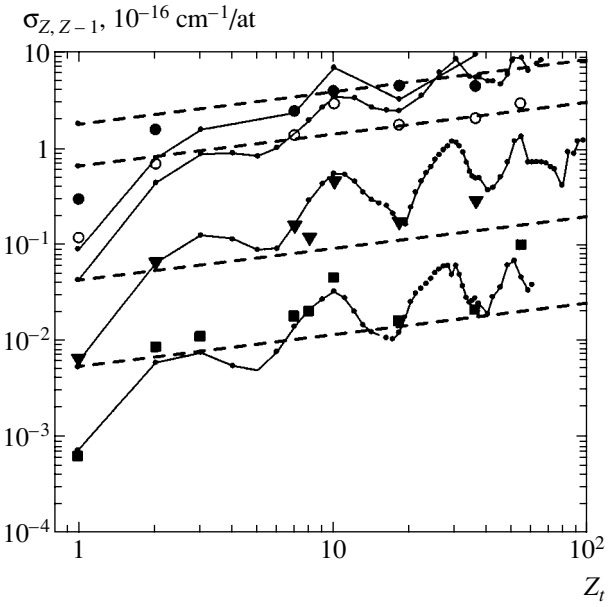


Fig. 2. Cross sections $\sigma_{Z,Z-1}$ of electron capture by nuclei of hydrogen, helium, boron, and nitrogen as functions of Z_t for velocity $V = 3.65$. Experiment: (■) H^+ [4], (▼) He^{2+} [3], (○) B^{5+}, our results, and (●) N^{7+} [5]. Solid curves correspond to the results of normalized calculations in the OBK approximation; dashed straight lines correspond to calculations based on formula (2).

exhibits oscillations with the amplitude decreasing with increasing the nuclear charge Z of the ion.

An analysis of capture cross sections of two electrons was also carried out using the experimental data for boron ions with velocities of $V = 1.19$ and 1.82. We calculated the ratio $\eta_2 = \sigma_{i,i-2}/\sigma_{i,i-1}$ of cross section of capture of two electrons to the cross sections of capture of one electron by boron ions. Within the experimental error, these ratios remain constant for the electron capture cross sections in all gaseous media studied, except krypton and xenon. This enabled us to average η_2 over various media; as a result, the minimal value of $\eta_2 = 0.004 \pm 0.001$ was obtained for capture of two electrons by B^+ ions with the formation of negative B^- ions. For boron ions with charges $i = 2$–5, ratio η_2 is 0.09 ± 0.05.

The cross section $\sigma_{i,i-3}$ of capture of three electrons is 3–5 times smaller than the cross section $\sigma_{i,i-2}$ for capture of two electrons. The ratios $\eta_3 = \sigma_{i,i-3}/\sigma_{i,i-1}$ for capture of three electrons were obtained only for boron ion velocities of $V = 3.65$ and 5.50 and were approximately equal to 0.02 ± 0.01.

5. CONCLUSIONS

It has been established that the values of cross sections $\sigma_{i,i-1}$ obtained from the experimental data increase on the average with Z_t in proportion to $Z_t^{1/3}$ and can be successfully described by the model of independent electrons in the low-velocity range. However, the analysis of experimental data revealed deviations from the monotonic dependence of $\sigma_{i,i-1}$ on Z_t: in the range of low velocities, the cross sections in neon are smaller than the corresponding values in nitrogen and argon by a factor of 2–3. As the ion velocity increases, the dependence of the electron capture cross section on Z_t becomes essentially nonmonotonic.

The calculations carried out in the OBK approximation successfully describe the oscillations observed in the dependence of cross sections $\sigma_{i,i-1}$ on Z_t. We proposed a unified empirical formula for various ions of light elements to normalize the electron capture cross sections $\sigma_{i,i-1}$ calculated in the OBK approximation.

The results of our calculations make it possible to analyze the variation of oscillations in the cross section upon a change in the ion charge and velocity. With increasing V, the position of the extrema in the dependence of cross sections $\sigma_{i,i-1}$ on Z_t is shifted towards higher values of Z_t. The amplitude of oscillation depends both on the ion charge i and on the nuclear charge Z of the ion. An increase in the ion charge reduces the amplitude of oscillations, which attains its minimal value for nuclei. With increasing nuclear charge of the ion, the amplitude of oscillations also decreases if we consider the cross sections of electron capture by the nuclei of the elements; in this case, for electron capture by ions, the dependence of cross sections on Z is more complicated. Analysis of the dependence of the oscillation amplitude on velocity V makes it possible to determine the velocity range in which oscillations must be taken into account in the dependence of $\sigma_{i,i-1}$ on Z_t.

Thus, it has been established that the cross sections of electron capture by light ions ($Z = 1$–7) with velocities $V = 2$–10 as a function of Z_t exhibit oscillations due to peculiarities of the shell structure of atoms of the medium.

REFERENCES

1. V. S. Nikolaev, Usp. Fiz. Nauk **85**, 678 (1965) [Sov. Phys. Usp. **8**, 269 (1965)].
2. N. Bohr, *The Penetration of Atomic Particles through Matter*, 3rd ed. (Munksgaard, Copenhagen, 1960; Inostrannaya Literatura, Moscow, 1950).
3. I. S. Dmitriev, Ya. A. Teplova, and Yu. A. Faĭnberg, Zh. Éksp. Teor. Fiz. **116**, 1539 (1999) [JETP **89**, 830 (1999)].
4. I. S. Dmitriev, Yu. A. Tashaev, V. S. Nikolaev, *et al.*, Zh. Éksp. Teor. Fiz. **73**, 1684 (1977) [Sov. Phys. JETP **46**, 884 (1977)].
5. I. S. Dmitriev, Ya. A. Teplova, and Yu. A. Faĭnberg, Zh. Éksp. Teor. Fiz. **108**, 75 (1995) [JETP **81**, 40 (1995)].

6. I. S. Dmitriev, N. F. Vorob'ev, Zh. M. Konovalova, *et al.*, Zh. Éksp. Teor. Fiz. **84**, 1987 (1983) [Sov. Phys. JETP **57**, 1157 (1983)].

7. V. S. Nikolaev, I. S. Dmitriev, Ya. A. Teplova, and Yu. A. Faĭnberg, Vestn. Mosk. Univ., Ser. 3: Fiz., Astron. **86** (5), 21 (1994).

8. Ya. A. Teplova and I. S. Dmitriev, in *Development of Scientific Research on Nuclear and Atomic Physics in Institute of Nuclear Physics of Moscow State University* (Mosk. Gos. Univ., Moscow, 1994), p. 132.

9. V. S. Nikolaev, I. S. Dmitriev, L. N. Fateeva, and Ya. A. Teplova, Zh. Éksp. Teor. Fiz. **40**, 989 (1961) [Sov. Phys. JETP **13**, 695 (1961)].

10. T. V. Goffe, M. V. Shah, and H. B. Gilbody, J. Phys. B **12**, 3763 (1979).

11. V. S. Nikolaev, L. N. Fateeva, I. S. Dmitriev, and Ya. A. Teplova, Zh. Éksp. Teor. Fiz. **41**, 89 (1961) [Sov. Phys. JETP **14**, 67 (1961)].

12. N. Bohr and J. Lindhard, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **28**, 1 (1957).

13. M. C. Brinkman and H. A. Kramers, Proc. R. Acad. Sci. Amsterdam **33**, 973 (1930).

14. V. S. Nikolaev, Zh. Éksp. Teor. Fiz. **51**, 1263 (1966) [Sov. Phys. JETP **24**, 847 (1966)].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 924–935.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1058–1070.

ATOMS, SPECTRA, RADIATION

Processes in Condensed Inert Gases Involving Excess Electrons¶

E. B. Gordon[a] and B. M. Smirnov[b]

[a]*Institute of Problems of Chemical Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*

[b]*Institute for High Temperatures, Russian Academy of Sciences, Moscow, 127412 Russia*

e-mail: smirnov@orc.ru, smirnov@oivtran.iitp.ru

Received December 17, 2003

Abstract—Drift of an excess electron in dense and condensed inert gases in external electric field and excitation of atoms by electron impact in these systems are analyzed. The effective potential energy surface for an excess electron at a given electric field strength consists of wells and hills, and the actions of neighboring atoms are therefore separated by saddles of the potential energy. At such atomic densities that the difference of interaction potentials for an excess electron between neighboring wells and hills of the potential energy surface becomes small, the electron mobility is large. This is realized for heavy inert gases (Ar, Kr, Xe) with a negative scattering length of an electron on individual atoms. In these cases, the average potential energy of the electron interaction with atoms corresponds to attraction at low atomic densities and to repulsion at high densities. The transition from attraction to repulsion at moderate atomic densities leads to a maximum of the electron mobility. A gas model for electron drift in condensed inert gases is constructed on the basis of this character of interaction. Due to high electron mobility, condensed inert gases provide high efficiency of transformation of the electric field energy into the energy of emitting photons through drifting electrons. It is shown that, although the role of formation of autodetaching states in the course of electron drift is more important for condensed inert gases than for rare gases, this effect acts weakly on exciton production at optimal atomic densities. The parameters of a self-maintained electric discharge in condensed inert gases as a source of ultraviolet radiation are discussed from the standpoint of electron drift processes. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The mobility K and drift velocity w of a free electron in gases is inversely proportional to the number density N of atoms, i.e., the quantities KN and wN in a gas are independent of its density. At high densities, this law is violated because of two reasons. First, electron scattering proceeds simultaneously on several atoms, and second, the interaction between atoms changes the atomic system, which affects the character of the electron interaction in this system. For inert gases, where atoms conserve their individuality in a condensed system, the behavior of an excess electron injected in a condensed gas is different depending on its sort. In light inert gases, He and Ne, an excess electron compels to displace the surrounding atoms. This structure change moves together with the electron, and, therefore, the electron effective mass is on the order of the atomic mass, and, hence, the reduced electron mobility KN in these condensed gases (He and Ne) is substantially lower than in gases.

As follows from experiments [1–15], different behavior of the reduced electron mobility occurs in heavy inert gases, Ar, Kr, and Xe. The electron effective mass in condensed heavy inert gases is on the order of the free electron mass, and as the atomic number density increases, the reduced mobility has a tendency to decrease; it experiences a sharp jump in a narrow range of the atomic number densities (see Fig. 1 for the case of xenon [7, 11]). Table 1 [16] demonstrates this behavior of the mobility of the excess electron. In Table 1, K_{gas} is the electron zero-field mobility in gases at room temperature, K_{tr} is the electron zero-field mobility at the triple point, and K_{max} is the maximum electron zero-

Table 1. Parameters of drift of an excess electron in rare gases

	Ar	Kr	Xe
T_{max}, K	155	170	223
N_{max}, 10^{22} cm^{-3}	1.2	1.4	1.2
$K_{max}N_{max}$, 10^{24} (cm V s)$^{-1}$ [14]	22	64	72
T_{tr}, K	85	117	163
N_{liq}, 10^{22} cm^{-3}	2.1	1.8	1.4
$K_{tr}N_{liq}$, 10^{24} (cm V s)$^{-1}$ [1–6, 8, 10, 14]	10 ± 1	29 ± 5	28 ± 10
V_0, V [9, 12, 13, 15]	−0.3	−0.5	−0.8
$K_{gas}N$, 10^{23} (cm V s)$^{-1}$ [16]	12	0.62	0.17

¶This article was submitted by the authors in English.

field mobility for the liquid state. This table also contains the temperatures T_{max} of the maximum electron mobility, the triple point temperatures T_{tr}, and the number densities of atoms N_{max}, N_{liq} corresponding to these temperatures. We note that the difference in the electron mobilities for solid and liquid inert gases is not significant; i.e., high mobility of an excess electron cannot be explained by the order distribution of atoms. The quantity V_0 in Table 1 is the maximum difference of the electric potentials acting on the electron located inside and outside a condensed inert gas. It is energetically profitable for an excess electron to be located inside the inert gas, and the minimum of V_0 corresponds to the atomic number densities that are related to the maximum of the electron mobility. In addition, high mobility of an excess electron is observed only in a narrow range of atomic number densities, to be considered below. As follows from the data in Table 1, the maximum reduced mobilities of the excess electrons in inert gases, as well as the reduced mobilities at the triple point, significantly exceed those at gaseous densities. This difference is especially high for xenon, which is the main object of our consideration. In addition, the maximum reduced electron mobilities in inert gases exceed those in metals of high conductivity. Indeed, the reduced electron mobility K_eN_a is equal to 2.9 and 3.1 in the units 10^{24} (cm V s)$^{-1}$ for copper and silver, respectively. We note that, because electrons are degenerate in these metals, a typical electron velocity near the Fermi surface is much greater than the thermal velocity of a free electron. Therefore, although the specific mobilities of an excess electron in condensed inert gases significantly exceed those in metals, the ratios of the electron free mean path to the distance between nearest atoms (or the lattice constant) have the same order of magnitude for both condensed inert gases and metals.

Some theories [17–21] explain the high mobility of an excess electron by the Ramsauer effect in electron scattering on an individual atom, but such approaches are just models because they are correct only for gases. Of course, the gaseous approach for the electron scattering is the simplest one, but it does not allow us to describe the electron behavior in a wide range of atomic number densities with a small number of fitting parameters. The high mobility of an excess electron in condensed inert gases has a fundamental meaning because it is evidence of a weak interaction between the electron and this matter at such densities. In what follows, we consider just this range of the atom number densities corresponding to high values of the electron mobility, and our task is to explain the nature of this phenomenon. Analyzing the properties of the total potential energy that acts on the electron from a condensed inert gas, we show that it varies, on average, from attraction to repulsion as the atomic number density increases. The maximum electron mobility corresponds to the transition from attraction to repulsion, and in what follows, we analyze the problem of drift of an excess electron in condensed inert gases from this standpoint.

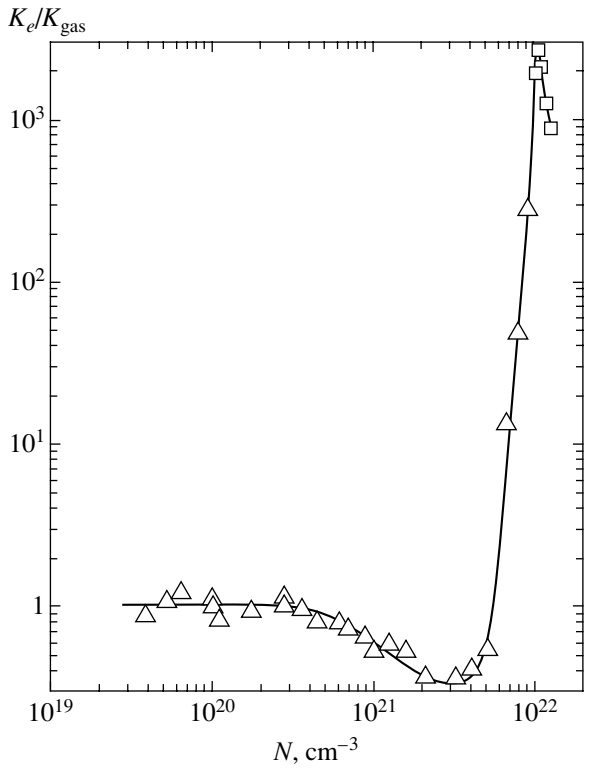


Fig. 1. The dependence of reduced electron mobility in xenon on the density of atoms according to experiments [7, 11].

The high mobility of excess electrons corresponds to a low resistance of condensed inert gases at such atom number densities if the excess electrons propagate in this matter under the action of an external electric field. When the electron energy reaches the threshold of atom excitation, it is consumed in the formation of excitons; in this range of atom number densities, formation of excitons by electron impact is an effective process. But formation of autodetachment states may affect the efficiency and rate of exciton formation in condensed inert gases. These autodetachment states are bound states of excited and excess electrons with a binding energy of about 0.4 eV [22]. Formation of autodetachment states of inert gas atoms impedes excitation of atoms, and we consider this problem below.

The high electron mobility in condensed inert gases has fundamental meaning and can be applied to transform the energy of an external electric field into the energy of photons in the vacuum ultraviolet (VUV) spectral range through excess electrons moving in condensed inert gases. This method, which requires the creation of a self-maintaining electric discharge in condensed inert gases, was suggested in [23, 24] and was then experimentally proved for xenon [25–27]. Because the excited inert gas atoms are characterized by a high excitation energy, such a method makes it possible to effectively convert the electric energy into

Table 2. Parameters of the interaction potential of two identical atoms of inert gases (R_e, D) and reduced parameters of condensed inert gases near the triple point [28]

	Ar	Kr	Xe	Average
R_e, Å	3.76	4.01	4.36	–
D, K	143	200	278	–
a, Å	3.755	3.992	4.335	–
a/R_e	1.00	0.99	1.01	1.005 ± 0.013
$N_0 = R_e^{-3}$, 10^{22} cm^{-3}	1.88	1.55	1.21	–
T_{tr}/D	0.587	0.578	0.570	0.579 ± 0.007
$p_{tr}R_e^3/D$, 10^{-3}	1.9	1.7	1.7	1.9 ± 0.2
V_{liq}/R_e^3	0.879	0.884	0.855	0.88 ± 0.02
V_{sol}/R_e^3	0.77	0.76	0.74	0.76 ± 0.01
ε_{sub}/D	6.5	6.7	6.7	6.5 ± 0.3
$\Delta H_{fus}/D$	0.990	0.980	0.977	0.98 ± 0.02
ΔS_{fus}	1.69	1.70	1.71	1.68 ± 0.03
r_W/R_e	0.639	0.641	0.627	0.64 ± 0.01

Table 3. Reduced parameters of an excess electron in condensed inert gases

	Ar	Kr	Xe
N_{sol}/N_0	1.30	1.31	1.34
N_{liq}/N_0	1.13	1.13	1.17
N_{max}/N_0	0.68	0.90	0.99
N_*/N_0	2.05	1.54	1.38
eV_0/ε_{sub}	–3.8	–4.3	–4.9

radiation because elastic scattering of electrons is weak. This problem of energy conversion is analyzed in this paper in the context of the behavior of excess electrons in condensed inert gases.

2. INTERACTION OF AN EXCESS ELECTRON IN CONDENSED INERT GASES

A feature of condensed inert gases is the small binding energy between atoms in comparison with typical atomic energies. This implies that the interaction between neighboring atoms in solid and liquid inert gases is relatively small and allows us to use the similarity law for various parameters of dense and condensed inert gases [28]. The reason for this is that the parameters of inert gases are governed by pairwise interactions of atoms, and the pair interaction potentials are known for inert gas atoms with a high (several percent) accuracy [29–32]. The pair character of atom interaction makes it possible to express certain parameters of inert gases via the parameters of the interaction potential of two atoms and ensures validity of the scaling laws for various parameters of bulk inert gases. Representing the pair interaction potential of two atoms in the form of a potential well, we use two interaction parameters, the depth D of this potential well and the equilibrium interatomic distance R_e that corresponds to the minimum of the interaction potential. Table 2 gives some reduced parameters of heavy inert gases [28] and confirms the validity of the similarity law for them. Here, a is the lattice constant, and all the inert gases have the face-centered cubic lattice; T_{tr} is the triple point temperature; V_{liq} and V_{sol} are the specific volumes per atom for the liquid and solid states, respectively, at the triple point; ε_{sub} is the sublimation energy per atom for the crystal at the triple point; ΔH_{fus} is the fusion energy (the energy consumed in melting) per atom; ΔS_{fus} is the entropy jump per atom at melting; and r_W is the Wigner–Seitz radius for the liquid state. As follows from Table 2, the reduced parameters are the same for different rare gases within the accuracy of several percent. Hence, the bound systems of inert gas atoms have a simple nature and can be treated as systems of classical bound atoms. We use this in the subsequent analysis.

Although scaling is not valid for an electron in condensed inert gases, it is convenient to express the reduced parameters of an excess electron, which allows us to compare the electron parameters with those for the interaction of atoms. The number densities N_{sol} and N_{liq} (where $N_{sol} = 1/V_{sol}$ and $N_{liq} = 1/V_{liq}$) in Table 3 correspond to the solid and liquid states of inert gases at the triple point, N_{max} is the number density in Table 1 at which the zero-field electron mobility has a maximum, $N_0 = R_e^{-3}$, and $N_* = \sqrt{2}/(2\bar{r})^3$, where $\bar{r}$ is the mean radius of the valence electron in a given atom and the values of these radii are taken from [33–35]. Hence, N_* is the number density of balls of radius $\bar{r}$ if these balls form a close-packed crystal lattice. In accordance with the Pauli exchange interaction, an excess electron cannot be located inside atoms, and, as follows from the data in Table 3, the excluded volume for the location of an excess electron is comparable to the total volume inside condensed inert gases at the triple point. As can be seen, the similarity law is not valid for N_*/N_0. In addition, the ratio N_{liq}/N_* grows as we transfer from Ar to Xe; this ratio expresses a typical part of space inside a liquid inert gas where an excess electron may not be located.

We now consider the problem of interaction of an excess electron inside a liquid inert gas from another standpoint, analyzing the behavior of the potential energy surface for this electron as the atom density

increases starting from low values. At low atomic densities, when an electron is located in a gas, it interacts with individual atoms independently. In regions between atoms far from them, the interaction potential is zero and nonzero interaction occurs only near regions occupied by atoms. On the basis of the Fermi formula [36], the interaction potential between the electron and atoms can be represented as

$$U(\mathbf{r}) = \sum_i \frac{2\pi\hbar^2}{m_e} L\delta(\mathbf{r} - \mathbf{R}_i), \quad (1)$$

where, $\mathbf{r}$ is the electron coordinate, $\mathbf{R}_i$ is the coordinate of the ith atom, and L is the electron–atom scattering length. Because the scattering length is negative for Ar, Kr, and Xe, this interaction potential corresponds to attraction in regions where atoms are located. Therefore, the potential energy surface consists of regions inside atoms with a sharp electron repulsion, regions near each atom with electron attraction, and regions between atoms with zero interaction potential. The region between atoms with zero interaction potential shrinks as the number density of atoms increases, and when the distance between nearest neighbors is comparable with the electron orbit size, the potential energy surface takes the form of wells and saddles, which separate regions of individual atoms. This potential energy surface resembles that describing interaction of bound atoms in clusters [37, 38]. In reality, attraction corresponds only to an average interaction of an electron of zero energy with an individual atom in a gas, and it leads to a redshift of spectral lines emitted by excited atoms located in inert gases [39]. The exchange interaction of a test electron with electrons of the inert gas atoms is accompanied by repulsion if this electron penetrates an internal atom region occupied by other electrons. The volume of repulsion near each atom is approximately $1/N_*$, where the values of N_* are given in Table 3, and this implies that high electron mobility is absent at high gas pressures.

The interaction potential between a test electron and an individual inert gas atom can therefore be repulsion at small distances from the atom and attraction at longer distances, which are of the order of the electron scattering length. Correspondingly, the attraction in the region of location of individual atoms dominates in the interaction potential of a test electron with the system of inert gas atoms at low densities of atoms. Evidently, as the number density of atoms increases, the attraction part of the interaction potential disappears, and, therefore, there is an atomic density when an average interaction potential becomes zero. This atomic density corresponds to the transition from attraction to repulsion for the total interaction potential, and the mobility of slow electrons obtains a maximum, because a typical energy difference between wells or hills in the potential energy becomes minimal at such an atomic density. This electron behavior is expected, however, if the electron scattering length on an individual atom is negative, i.e., for Ar, Kr, and Xe.

Thus, there is a strong repulsion of an excess electron in heavy inert gases near each core due to the exchange interaction with internal atomic electrons, and a test electron does not penetrate the atom because of the Pauli exclusion principle. Average electron interaction with an individual atom corresponds to attraction because of the negative electron–atom scattering length. At intermediate atomic densities, these interactions compensate each other, on average, and the electron mobility has a maximum at such densities. We are guided by the liquid state of condensed inert gases, because the order distribution of atoms is not important for the nature of this high electron mobility. Therefore, the electron mobilities for the solid and liquid states do not differ, in principle, but this mechanism of high electron mobility corresponds to a narrow density range, while high mobility of a solid inert gas can be observed in a wide range of atomic densities.

3. DRIFT AND MOBILITY OF AN EXCESS ELECTRON IN HEAVY INERT GASES

Guided by a range of high mobilities of an excess electron in condensed inert gases where the interaction of this electron with the environment is weak, we use a gas approach for electron scattering. Keeping in mind that the potential energy surface for an individual electron consists of wells and hills, and that hills separate the action of an individual core, we use the gaseous approach for electron scattering where each core is a scattering center for an excess electron. If the cross section for scattering on an individual core is small compared to the square atom size, we reduce the problem of electron motion inside a condensed inert gas to scattering on individual cores, as takes place in gases. We note that this model is valid only in the case of high electron mobility in a condensed system if the electron interaction with this system is weak.

In the case of electron motion in a gas, when a test electron is scattered subsequently on individual atoms, the zero-field electron mobility K_e is given by [40, 41]

$$K_e = \frac{e}{3m_e}\left\langle \frac{1}{v^2}\frac{d}{dv}\left(\frac{v^3}{\nu}\right)\right\rangle, \quad (2)$$

where m_e is the electron mass; v is the electron velocity; averaging is taken over the distribution of electrons with respect to velocities; $\nu = Nv\sigma^*$ is the rate of the electron–atom scattering, with N being the atom number density; and σ^* is the diffusion cross section of the electron–atom elastic scattering. For simplicity, we consider below the case where the cross section is inde-

Table 4. Parameters of the gas model for drift of an excess electron in liquid rare gases [16]

	Ar	Kr	Xe
$\sigma_{tr}/\pi a_{tr}^2$	0.012	0.005	0.005
$\sigma_{max}/\pi a_{max}^2$	0.004	0.002	0.002
λ_{tr}/a_{tr}	65	150	170
λ_{max}/a_{max}	200	450	480

pendent of the collision velocity and formula (2) takes the form [40, 41]

$$K_e = 0.53\frac{e\lambda}{\sqrt{m_e T}}, \text{ or } K_e N = 0.53\frac{e}{\sigma^*\sqrt{m_e T}}, \quad (3)$$

where $\lambda = (N\sigma^*)^{-1}$ is the mean free path of an electron in a gas and T is the temperature. This connection of the reduced electron mobility and the scattering cross section allows us to express the diffusion cross section σ^* of the electron–atom scattering in a gas via the reduced mobility, which is given by

$$\sigma^* = \frac{0.5}{\sqrt{m_e T}K_e N}. \quad (4)$$

In particular, applying this gas model to gaseous xenon and using the experimental values [7] of the reduced mobility $K_e N = 3 \times 10^{22}$ (cm V s)$^{-1}$ at the temperature $T = 236$ K, we obtain from formula (4) the diffusion cross section $\sigma^* = 52$ Å^2 at this temperature. We introduce the critical number density of atoms N_{cr} that characterizes the transition from the gaseous system to the condensed one and corresponds to the relation $\sigma^* = \pi r_W^2$, where r_W is the Wigner–Seitz radius ($\pi r_W^2 = 2.3 \times 10^{-15}$ cm^2 for liquid xenon). According to formula (4), the critical number density is

$$N_{cr} = \frac{3\sqrt{\pi}}{4(\sigma^*)^{3/2}}, \quad (5)$$

and it is equal to $N_{cr} = 4 \times 10^{21}$ cm^{-3} for xenon. This is the transient gas–liquid number density for xenon.

We treat the density range of the maximum electron mobility in xenon on the basis of the gas model. Taking the maximum zero-field electron mobility from Table 1 ($K_e N = 7.2 \times 10^{25}$ (cm V s)$^{-1}$ at the atom number density $N_{max} = 1.2 \times 10^{22}$ cm^{-3}), we obtain on the basis of formula (4) $\sigma^* = 1.4 \times 10^{-18}$ cm^2, which is small in comparison with $\pi r_W^2 = 2.3 \times 10^{-15}$ cm^2, and, hence, the gas model is applicable in this case. Thus, composing electron scattering in a condensed inert gas at the atomic density of the maximum electron mobility as a result of the electron interaction with small independent scatterers located on atom cores, we use formula (4) to find the diffusion cross sections for electron scattering σ_{max} and σ_{tr} that correspond to the atom number densities N_{max} and N_{tr} (see Table 1). These cross sections are given in Table 4 [16], where a_{tr} and a_{max} are the distances between nearest neighbors at these atomic densities, and λ_{max} and λ_{tr} are the electron mean free paths for these densities. As can be seen, the gas approach is valid because the effective cross sections are relatively small, whereas the mean free paths for electrons are relatively large.

We consider the data in Table 1 from another standpoint, composing the potential energy surface for an excess electron in the form of wells and hills near each core if the well depth is relatively small. We take the interaction potential of the electron with each atomic center in the form

$$U(r) = -U_0 \exp\left(-\frac{r^2}{a^2}\right), \quad a \le r_W, \quad (6)$$

where r is the distance from the atom's center, a is the range of atomic forces, and r_W is the Wigner–Seitz radius for the condensed system. The electron scattering on an individual center is weak if $U_0 \ll \varepsilon$, where ε is the electron energy. We can then use perturbation theory for electron scattering (the Born approximation), and the differential cross section of the electron scattering on an individual center in the Born approximation for the interaction electron–core potential (6) is [42]

$$d\sigma = \frac{\pi a^2}{4}\left(\frac{m_e U_0 a^2}{\hbar^2}\right)^2 \exp\left(-\frac{K^2 a^2}{2}\right) d\Omega, \quad (7)$$

where

$$K = 2q\sin(\theta/2)$$

is the variation of the electron wave vector as a result of scattering, θ is the scattering angle, q is the initial electron wave vector, and

$$d\Omega = \pi d\cos\theta$$

is the solid angle element. From this, we have the diffusion cross section of elastic scattering on each scattered center, given by

$$\sigma^* = \frac{\pi^2 a^2}{16}\left(\frac{U_0}{\varepsilon}\right)^2. \quad (8)$$

This consideration is valid for $U_0 \ll \varepsilon$, i.e., in the range of parameters where the electron mobility is high in a condensed system. In the case of xenon at the atomic density of the maximum electron mobility, we obtain $U_0/\varepsilon = 0.05$ if $a = r_W$ in the interaction potential (6). Thus, from different standpoints, we find that the interaction of an excess electron with the environment is relatively weak at atomic densities where the electron

mobility has a maximum and the diffusion cross section of the electron at each core is also relatively small.

As follows from the above analysis, the gas model can be valid for the mobility of an excess electron in condensed inert gases in some range of atomic densities. This means that, in the case where we compose the potential energy surface for an individual electron inside a condensed inert gas in the form of wells and hills, the amplitude of the electron scattering on an individual well or hill is less than the distance between the nearest neighboring atoms. We can also use another criterion of the gas model validity: when the electron mean free path in a condensed inert gas is large compared to the distance between neighboring atoms. This allows us to use the gas model for electron scattering, according to which an electron is scattered independently on neighboring nonuniformities of the potential energy surface. This leads to the classical theory of electron kinetics in gases in an external electric field in this case of an excess electron in condensed inert gases. This theory was elaborated for kinetics of electrons in semiconductors and gases [44–50]. It is represented in contemporary publications [41, 51, 52], and we use this theory below for excitation of atoms in condensed inert gases.

Although the criterion for the weakness of the interaction of an excess electron in condensed inert gases at high electron mobilities is fulfilled, as well as the criterion for the gas approach for propagation of an excess electron, the character of electron motion is in reality more complex. Indeed, due to the exchange interaction of an excess electron with atomic electrons, an excess electron cannot penetrate the atoms. Hence, if we consider electron scattering at atomic cores to be independent, the cross section of scattering at each core is on the order of the atomic radius squared, which significantly exceeds the values in Table 4 that follow from the mobility data. Therefore, a large mean free path for electrons inside liquid inert gases may be explained by collective effects in simultaneous electron scattering on several cores, and the above gas model has a qualitative character. Nevertheless, because of its simplicity, we use the above gas model for the electron scattering in liquid inert gases as a result of the interaction with independent scatterers.

4. EXCITATIONS IN CONDENSED INERT GASES BY ELECTRONS DRIFTING IN AN EXTERNAL ELECTRIC FIELD

In analyzing the electron behavior in a gas of independent scatterers, we use the classical theory [44–50] of electron motion in gases under the action of an external electric field. The basis of this theory is the small change in electron energy when there is a significant change in the electron momentum as a result of elastic electron–atom scattering due to the small ratio of the electron and atomic masses. This allows us to expand the velocity distribution function of the electron $f(\mathbf{v})$ over the spherical harmonics, and this distribution function has the form

$$f(\mathbf{v}) = f_0(v) + v_x f_1(v), \tag{9}$$

where v_x is the electron velocity component along the electric field and $vf_1(v) \ll f_0(v)$ according to the basic concept. Although the antisymmetric part of the distribution function $f_1(v)$ is small, it is of importance because the electric field acts on electrons via this component of the distribution function. The set of equations for the distribution function in the case of only the elastic electron–atom scattering and in neglecting inelastic processes has the form (see, e.g., [51])

$$a\frac{df_0}{dv} = -\nu v f_1, \quad \frac{a}{3v^2}\frac{d}{dv}(v^3 f_1) = I_{ea}(f_0) \tag{10}$$

in the stationary case, where $a = eE/m_e$, E is the electric field strength, $\nu = Nv\sigma(v)$ is the rate of electron collisions with atoms, and $I_{ea}(f)$ is the electron–atom collision integral. This implies a general expression (2) for the electron mobility, where an average is taken over the spherical component f_0 of the distribution function.

We consider the simplest case where $\sigma(v) \sim 1/v$ and the rate ν is independent of v. We then obtain the expressions

$$f_0(\varepsilon) \sim \exp\left(-\frac{\varepsilon}{T_e}\right), \quad w = \frac{a}{\nu} = \frac{eE}{m_e\nu}, \qquad T_e = T + \frac{Ma^2}{3\nu^2}, \tag{11}$$

where ε is the electron energy (which is to be used along with the electron velocity) and T is the gas temperature. We note that we ignore collisions between electrons and that the parameter T_e does not correspond to the definition of the electron temperature, but coincides with it in the expression for distribution function (11). The drift velocity is proportional to the electric field strength, and this dependence is related to liquid xenon as long as inelastic collisions are weak [53]. Hence, this simple dependence $\nu(v)$ describes the electron behavior in liquid xenon. On the contrary, the electron behavior in gaseous xenon is more complex because of a nonmonotonic velocity dependence for the electron–atom cross section due to the Ramsauer effect. Such a dependence leads to the maximum of the electron mobility as a function of the electric field strength [7], which also follows from detailed calculations [54] for gaseous xenon. Below, we use the simplest dependence $\sigma(v) \sim 1/v$ for the analysis of electron kinetics in liquid xenon.

Also taking into consideration excitation of atoms by electron impact and assuming that, above the excitation threshold $\Delta\varepsilon$, the electron loses the energy by atom excitation, we find that the distribution function is zero at the excitation threshold $f_0(\Delta\varepsilon) = 0$. This gives the energy

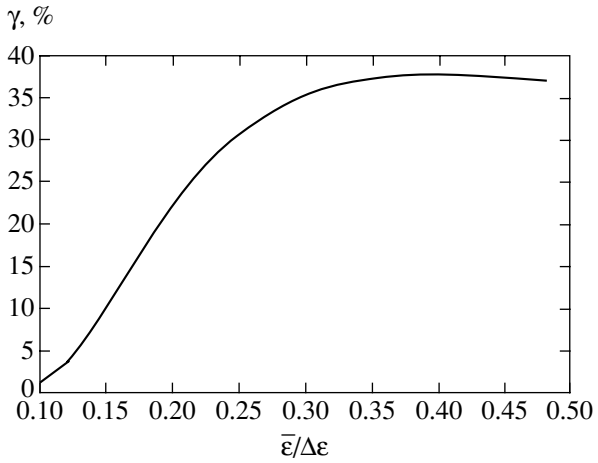


Fig. 2. The efficiency of excitation of atoms in an atomic system if the rate ν of the electron–atom elastic scattering is independent of the electron velocity v.

distribution function $f_0(\varepsilon)$ of an excess electron [55] in the form

$$f_0(\varepsilon) = \begin{cases} C[\varphi_0(\varepsilon) - \varphi_0(\Delta\varepsilon)] & \varepsilon \le \Delta\varepsilon, \\ 0, & \varepsilon \ge \Delta\varepsilon, \end{cases} \qquad (12)$$

where C is the normalization constant and $\varphi_0(\varepsilon)$ is given by formula (11),

$$\varphi_0(\varepsilon) = \exp\left(-\frac{\varepsilon}{T_e}\right). \qquad (13)$$

In the regime under consideration, the drift velocity is proportional to the electric field strength E even at high fields and the parameter $T_e \sim E^2$ at high fields, while the average electron energy is restricted by the value $(3/7)\Delta\varepsilon$ in the limit of high electric fields. In this regime, electrons acquire energy from the external field and lose it in elastic collisions with atoms. When an electron reaches the excitation energy $\Delta\varepsilon$, a forming excited atom emits a photon. The electron energy becomes zero after atom excitation, and the process of the increase of the electron energy repeats. Therefore, the rate of atom excitation is determined by the flux in the energy space, and, on the basis of the indicated energy balance, we have the rate of atom excitation given by [16]

$$\frac{dN_*}{dt} = \frac{4}{\sqrt{\pi}}\left(\frac{\Delta\varepsilon}{T_e}\right)^{3/2} N_e \frac{m_e}{M}\nu \exp\left(-\frac{\Delta\varepsilon}{T_e}\right), \qquad (14)$$
$$T \ll T_e \ll \Delta\varepsilon,$$

where N_* is the number density of excited atoms, N_e is the number density of excess electrons, and $\Delta\varepsilon$ is the excitation energy.

We now find the portion γ of the power acquired by the electrons from the external electric field and lost to atom excitation. We assume that the power acquired from the field is transformed below the excitation threshold mostly into the atom thermal energy as a result of elastic collisions between electrons and atoms, and this power per one electron is eEw, where w is the electron drift velocity. From formula (14), we then have

$$\gamma = \frac{\Delta\varepsilon \dfrac{dN_*}{dt}}{eEwN_e} = \frac{4}{3\sqrt{\pi}}\left(\frac{\Delta\varepsilon}{T_e}\right)^{3/2} \exp\left(-\frac{\Delta\varepsilon}{T_e}\right), \qquad (15)$$
$$T \ll T_e \ll \Delta\varepsilon,$$

where

$$T_e = \frac{Ma^2}{3\nu^2} = \frac{Mw^2}{3}$$

is the effective electron temperature. Figure 2 gives the dependence of the efficiency of atom excitation γ on the electron energy $\bar{\varepsilon} = 3T_e/2$ under these conditions [16]. Formally, this expression has a maximum at $T_e = 2\Delta\varepsilon/5$, where $\xi = 0.61$, but because $\varepsilon \le 3\Delta\varepsilon/7$, the above consideration is valid below this limit. It follows from (15) that the transformation efficiency γ is significant even at low values of $T_e/\Delta\varepsilon$.

In this consideration, we implicitly use the criterion

$$\nu_{ex} \gg \frac{m_e}{M}\nu,$$

where ν_{ex} is a typical excitation rate of atoms in collisions with a test electron. To obtain a more specific criterion, we analyze the other limiting case of slow atom excitation. The energy distribution function then has the form

$$f_0(\varepsilon) = C\varphi_0(\varepsilon)$$

and differs from that in formula (9) by the absence of decay of fast electrons in the excitation process. Correspondingly, the rate of the atom excitation is equal to [16, 41]

$$\frac{dN_*}{dt} = N_e N_a k_q \frac{g_*}{g_0} \exp\left(-\frac{\Delta\varepsilon}{T_e}\right), \qquad (16)$$

where k_q is the rate constant of quenching resonantly excited atoms by a slow electron, and g_0 and g_* are the statistical weights of the ground and excited atomic states. We have used the principle of detailed balance between the excitation and inverse quenching processes [51, 56], which is convenient because the quenching rate constant is independent of the electron energy for a slow electron ($T_e \ll \Delta\varepsilon$). The values of the quenching rate constants for resonantly excited atoms

Table 5. Parameters of the lowest resonantly excited states of inert gas atoms [16, 56]

Atom (state)	$\Delta\varepsilon$, eV	J^*, eV	τ, ns	k_q, 10^{-9} cm^3/s	$1/k_q\tau$, 10^{16} cm^{-3}	k_{el}, 10^{-8} cm^3/s	$(m_e/M)k_{el}$, 10^{-14} cm^3/s
Ar ($1s_2$)	11.62	4.14	10	0.82	12	0.15	2.0
Ar ($1s_4$)	11.83	3.93	2	3.9	13		
Kr ($1s_2$)	10.03	3.97	3.5	3.9	7.3	2.8	18
Kr ($1s_4$)	10.64	3.36	3.2	3.5	8.9		
Xe ($1s_2$)	8.44	3.69	3.6	7.0	4.0	10	43
Xe ($1s_4$)	9.57	3.43	3.5	4.6	6.2		

of inert gases are given in Table 5. We there give the values of k_{el} for thermal collisions on the basis of the electron mobilities in gases according to the data in Table 1.

Comparing the excitation rates according to formulas (13) and (16), we find them to be simultaneously valid if the criterion

$$k_q \gg \frac{m_e}{M}\left(\frac{\Delta\varepsilon}{T_e}\right)^{3/2} k_{el} \qquad (17)$$

is satisfied, where the rate constant of the elastic electron–atom scattering is introduced as $k_{el} = \nu/N_a$. Using the gas model for elastic scattering of electrons on atoms, we take the quenching rate constants for electron–atom collisions and radiative times of excited atoms in a condensed inert gas to be close to those in a gas, whereas the rate constants for the elastic electron scattering are much less in condensed state. This means that criterion (17) is even more valid for condensed inert gases than for their gaseous phase. In addition, this leads to a high efficiency of transformation of the electric field energy into excitations in electronic excitation for condensed inert gases.

We note [57] that, in spite of the simplicity of condensed inert gases as a system of bound atoms in the ground state (see Table 2), elementary excitations in this system—excitons—have a complex structure. One more peculiarity of excitons in condensed inert gases in comparison with excitations in a gaseous system is due to the interaction of an excited atom with the environment. In gases, excited atoms are formed as a result of the electron impact, and these excited atoms emit radiation. In condensed inert gases, an excited atom is transformed very fast into a diatomic excimer molecule and lives in such a form. Therefore, radiation of a condensed inert gas is characterized by a broad band for a quasimolecular exciton, and luminescence is redshifted in comparison with the spectral line of atom emission, and one can expect the luminescence quantum yield to be close to unity.

There are two types of quasimolecular excitons depending on the total spin of an excited electron and the core. Because the electron spin is zero for the ground state, the radiative time of an exciton with the total spin one is much greater than the radiative time of an exciton with zero total spin [58]. This fact is taken into account in the analysis of exciton kinetics.

5. ROLE OF AUTODETACHING STATES IN GENERATION OF EXCITONS

Because of weak electron–atom interaction in condensed inert gases for atomic densities of high electron mobility, an excess electron experiences a weak friction when it is drifting in an external electric field. Therefore, the efficiency is high for conversion of the electron energy obtained from the field into excitation of atoms. But this weak electron–matter interaction is the reason for the increased role of autodetaching states in the formation of excitons. We consider this problem in what follows.

Excitation and decay of autodetaching states in the course of motion of an excess electron in condensed inert gases proceeds according to the scheme

$$e + \mathrm{A} \longrightarrow (\mathrm{A}^-)^{**}, \qquad (18)$$

where A is the inert gas atom. The parameters of autodetaching states $(A^-)^{**}$ of inert gas negative ions are analogous to those of $H^-(2s^2)$ [59], where the autodetaching state is placed about 0.4 eV lower than the atom excitation energy and the lifetime of these autodetaching states is approximately 10^{-14}–10^{-13} s. Because the typical time of radiative decay of an ionic state via the emission of a UV photon is about 10^{-9}–10^{-8} s, its probability of occurring during the decay of an autodetaching state is small. However, this process can be repeated, and the process of photon emission through channel (18) may be significant. If channel (18) is realized, the spectrum of radiation is characterized by longer wavelengths than that from A* but shorter than from A_2^* [22].

Electron capture in an autodetaching state leads to an additional decrease in the electron distribution func-

tion, and the capture cross section σ_{res} is given by the Breight–Wigner formula [42]

$$\sigma_{res} = \frac{\pi\hbar^2}{2m_e\varepsilon}\frac{\Gamma^2}{(\varepsilon-\varepsilon_{res})^2+\Gamma^2/4}, \quad (19)$$

where Γ is the width of the autodetaching level and ε_{res} is the electron energy for this resonance. We first determine formally a decrease in the energy distribution function of electrons, considering electron capture on the autodetaching level as the elastic electron scattering. We can then represent the cross section of the electron–atom elastic scattering as the sum of two parts,

$$\sigma = \sigma_0 + \sigma_{res}, \quad (20)$$

where σ_0 smoothly depends on the electron energy and σ_{res} is the resonant part of the elastic cross section. Restricting ourselves to the Druyvesteyn case of the electron distribution function [44], where the typical electron energy significantly exceeds the thermal energy of atoms, we find that, instead of formula (12), the distribution function is given by

$$\varphi_0(\varepsilon) = \exp\left(-\int^{\varepsilon} d\varepsilon\left(\frac{Ma^2}{3\nu^2}\right)^{-1}\right). \quad (21)$$

If we assume the rate of elastic electron–atom scattering ν to be independent of the electron energy, formula (19) transforms to formula (12) when the resonant part of electron scattering is neglected.

The character of the electron interaction with an autodetaching state consists in the electron capture on this level and in the subsequent decay of the resonance level, which can lead to a change in the direction of the electron velocity. Thus, the capture of an electron on the resonance level is similar to elastic scattering of the electron with a change in the direction of its motion. Assuming that $\sigma_{res} \gg \sigma_0$ at the resonance, we find that formation of the autodetaching state under the above consideration leads to a jump in the exponent (19), and this jump is equal to

$$\xi = \int^{\varepsilon} d\varepsilon\left(\frac{Ma^2}{3\nu_{res}^2}\right)^{-1}, \quad (22)$$

where $\nu_{res} = N\nu\sigma_{res}$, which implies that, above the resonance, the distribution function acquires the factor $\exp(-\xi)$. This means that distribution function (13) transforms into

$$\varphi_0(\varepsilon) = \exp\left(-\frac{\varepsilon}{T_e}\right), \quad \varepsilon < \varepsilon_{res},$$

$$\varphi_0(\varepsilon) = \exp\left(-\frac{\varepsilon}{T_e} - \xi\right), \quad \varepsilon > \varepsilon_{res}.$$

Under the assumption that the rate ν of the electron–atom scattering is independent of ε, the exponent is given by

$$\xi = \int\frac{3d\varepsilon}{Ma^2}(N\nu\sigma_{res})^2 = \frac{3\pi\Gamma}{4Ma^2}(Nk_{res})^2, \quad (23)$$

where

$$k_{res} = \frac{4\pi\hbar^2}{v_{res}}, \quad v_{res} = \sqrt{\frac{2\varepsilon_{res}}{m_e}}. \quad (24)$$

In particular, for $\sigma_{res} = 8$ eV (xenon), we obtain $k_{res} = 1.4\times10^{-7}$ cm^3/s and this estimate is valid for both gaseous and condensed states when the gas model is applicable.

Thus, the role of the autodetaching level in electron kinetics is governed by exponent (23). Taking $k_{el} = \nu/N_a$, where k_{el} is the rate constant of the elastic electron–atom scattering aside the resonance, we obtain

$$\xi = \frac{\pi\Gamma}{4T_e}\left(\frac{k_{res}}{k_{el}}\right)^2. \quad (25)$$

Using formula (25) for the estimate in gaseous and condensed xenon, we take $\Gamma \approx 0.01$ eV, $T_e \sim 3$ eV, $\sigma_{res} \approx 8$ eV, and the value of k_{el} from Table 5. This gives $\xi \approx 0.005$ for gaseous xenon; i.e., autodetaching states do not affect the electron distribution function there. Applying the gas model to condensed xenon, we can evaluate the rate constant of the electron–atom scattering from the known zero-field mobility of electrons as $k_{el} = 2.4\times10^{-11}$ cm^3/s at the number density corresponding to the mobility maximum. Substituting this small rate constant in (25), we obtain $\xi \sim 10^5$. Thus, as supposed in [22], formation of autodetaching states is not essential for excitation of atoms in gases even at high atomic densities. However, it may be important in condensed inert gases at densities where the electron mobility is high.

The autodetaching states are of importance in electron kinetics and atom excitation if $\xi \gg 1$; we determine this limit below. The probability $dw(\varepsilon)$ of the electron energy being in a range from ε to $\varepsilon + d\varepsilon$ after the decay of an autodetaching state according to formula (19) is given by

$$dw = \frac{ds}{\pi(1+s^2)}, \quad s = \frac{2(\varepsilon-\varepsilon_{res})}{\Gamma}. \quad (26)$$

We suppose that autodetaching is not essential if $\sigma_{res} \le \sigma_0$, i.e., if

$$s^2 \ge s_0^2 = \frac{2\pi\hbar^2}{m_e\varepsilon_{res}\sigma_0}, \quad (27)$$

and the probability w that $s \ge s_0$, when the autodetach-

ing state can be ignored, is equal to

$$w = \int_{s_0}^{\infty} dw = \frac{1}{\pi s_0}. \quad (28)$$

In particular, for xenon at the atomic number density of the maximum electron mobility, we have $s_0 \approx 14$. The presence of an autodetaching state therefore acts as a barrier in the kinetics of electrons in the space of electron energy. In the case of xenon, the probability of passing through this barrier is approximately 0.02. Of course, this decreases the efficiency of transformation of the energy of the external electric field into the energy of emitting photons through electrons that excite atoms in condensed inert gases.

This effect decreases the efficiency γ of transformation of the electric field energy into the energy of VUV photons because it leads to a drop in the distribution function. However, in the limit

$$T_e \gg \Delta\varepsilon - \varepsilon_{\text{res}}, \quad (29)$$

this influence on the efficiency coefficient γ is small if the criterion

$$wk_q \gg \frac{m_e}{M}\left(\frac{\Delta\varepsilon}{T_e}\right)^{3/2} k_{\text{el}} \quad (30)$$

is satisfied instead of criterion (17). Indeed, because $\Delta\varepsilon - \varepsilon_{\text{res}} \approx 0.4$ eV, criterion (29) is valid if the efficiency γ is not small. Next, both the influence of electron capture in the autodetaching state and the atom excitation lead to a decrease of the distribution function, but because these effects are not separated, all the fast electrons lose their energy for the excitation of atoms due to criterion (30). Thus, although autodetaching states affect the efficiency of exciton production by drifting the excess electrons in condensed inert gases, this is evidently inessential at optimal atomic densities. An indirect confirmation of this is the efficiency of 18% for conversion of the electric energy into VUV radiation that is observed in solid xenon [26, 27]. Formula (15) gives this value at $T_e = 2.5$ eV ($\Delta\varepsilon/T_e = 3.2$).

Because of high efficiency, it is advantageous to create a self-maintaining electric discharge in condensed inert gases for generation of VUV photons, as suggested in [23, 24]. Such an emission was observed in liquid xenon [25] in the form of a broad line near the central wavelength of 175 nm, when electrons were transported from a cold field-emission cathode at moderate electric field strengths. Generation of VUV photons in solid xenon was achieved in experiments [26, 27]. As follows from the above analysis, the energy distribution function of the excess electrons is zero at electron energies above the excitation threshold, and, hence, the direct ionization of atoms by electron impact is impossible in condensed inert gases. In addition, the number density of the excess electrons is relatively small, and stepwise ionization does not proceed in condensed inert gases. Therefore, to support a discharge, ionization processes are required outside a condensed inert gas. In particular, within the framework of the experimental scheme of [26, 27], a photocathode is placed near solid xenon and secondary electrons form as a result of absorption of VUV photons by the photocathode. A gap filled with gaseous xenon makes it possible to avoid the self-diffusion effect that decreases by several orders of magnitude the probability of an electron entering a media in the case of a photocathode in direct contact with condensed matter [60]. In solid xenon 1 mm thick at an electric tension of 1 kV, approximately 20 photons may be formed per one electron. As a result, a self-maintaining discharge is created in this scheme.

We now make some evaluations for this scheme with a layer of condensed xenon and a layer of gaseous xenon that are governed separately because of using a grid. For definiteness, we take xenon at the triple point, such that the temperature is equal to 163 K and the pressure of gaseous xenon is 0.8 atm, which corresponds to the number density of xenon atoms approximately 4×10^{19} cm^{-3}. Taking the characteristic energy in liquid xenon about 3 eV, which provides an efficiency of conversion of the electric energy into VUV radiation of about 20%, we obtain the electric field strength for liquid xenon approximately 100 V/cm. After passing the liquid layer, each electron creates five VUV photons per centimeter of its path. The drift velocity of electrons in liquid xenon is 2×10^5 cm/s under these conditions. In the gaseous layer, an average energy of 2 eV is attained at a reduced electric field strength of $E/N \approx 1$ Td [54], which corresponds to 400 V/cm. We note that the number k of electrons forming per electron in a condensed inert gas is equal to

$$k = \frac{\gamma E l}{\Delta\varepsilon}, \quad (31)$$

where l is the layer thickness. Therefore, it is advantageous to increase the electric field strength in a layer of a condensed inert gas. This is not valid for a gaseous layer, because, along with excitation processes, direct ionization of atoms by electron impact proceeds, which restricts both the layer thickness and the electric field strength. The total electric current density is limited by heat transport processes, and this value is measured in μA/cm^2.

On the basis of the experience of experimental study [25–27] and from the above analysis, one can simplify the scheme of a self-maintaining electric discharge in a condensed inert gas and improve the discharge parameters if the electric discharge is a generator of VUV radiation. First, a CsI photocathode is useful for this goal, because its efficiency for the electron emission is almost three orders of magnitude higher at $\lambda = 172$ nm than that for a zinc photocathode [60] that was used in experiments [26, 27]. Second, it is convenient to apply an alternating voltage to a layer from condensed inert gases. In this manner, using a suitable frequency of the

electric field ($f \approx 1$ MHz), one can increase the path that an electron passes in a layer. As a result, in a layer with a thickness of several millimeters, an electron traverses a path of several meters. Next, due to the increase in the electron lifetime in a discharge, stepwise ionization processes may be important, increasing the electron number density. This makes it possible to improve the parameters of a self-maintaining electric discharge.

6. CONCLUSIONS

High electron mobility is observed in heavy condensed inert gases (Ar, Kr, Xe) in a narrow range of atomic densities. The widespread explanation of this effect [17–21] by the Ramsauer effect in the electron scattering on an individual atom is not correct because of the large distance of the electron–atom scattering in comparison with the distance between neighboring atoms at these atomic densities. We have shown that the nature of high electron mobility is connected to the transition from an attracting interaction potential to a repulsing one between an excess electron and an atom ensemble. High electron mobility is accompanied by weak electron interaction with atoms of a condensed inert gas, which allows us to use a gas model for electron scattering inside this system. In reality, electron scattering in condensed inert gases is not reduced to scattering on individual cores, i.e., collective effects are important in these processes. Therefore, a quantitative description is used in experimental results, and the parameters corresponding to the electron mobility maximum do not admit similarity law for different inert gases. In addition, the analysis of electron kinetics in condensed inert gases in external fields shows that the processes of formation of autodetaching states lead to a decrease in the electron energy distribution function with increasing electron energy. Nevertheless, formation of autodetaching states may be not essential for the transformation of the electric field energy into vacuum ultraviolet radiation if a condensed inert gas is transparent to electron drift.

The results of this analysis can be used for creating a self-maintaining electric discharge in condensed inert gases to generate ultraviolet radiation. The basis for this is the existing versions of such a discharge [25–27] together with the above analysis of the nature of processes involving electrons. This allows us to take the next step in constructing a new version of self-maintaining electric discharge in condensed inert gases.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (project nos. 04-03-32684 and LSS-1953.2003.2).

REFERENCES

1. H. Schnyders, S. A. Rice, and L. Meyer, Phys. Rev. **150**, 127 (1966).
2. B. Halpern *et al.*, Phys. Rev. **156**, 351 (1967).
3. L. S. Miller, S. Howe, and W. Spear, Phys. Rev. **166**, 871 (1968).
4. J. A. Jahnke, L. Meyer, and S. A. Rice, Phys. Rev. A **3**, 734 (1971).
5. T. Kimura and G. R. Freeman, Can. J. Phys. **52**, 2220 (1974).
6. K. Yoshino, U. Sowada, and W. F. Schmidt, Phys. Rev. A **14**, 438 (1976).
7. S. S.-S. Huang and G. R. Freeman, J. Chem. Phys. **68**, 1355 (1978).
8. S. S.-S. Huang and G. R. Freeman, Phys. Rev. A **24**, 714 (1981).
9. R. Reiniger, U. Asaf, and I. T. Steinberger, Chem. Phys. Lett. **90**, 287 (1982).
10. E. M. Gushchin, A. A. Kruglov, and I. M. Obodovskiĭ, Zh. Éksp. Teor. Fiz. **82**, 1114 (1982) [Sov. Phys. JETP **55**, 650 (1982)].
11. V. V. Dmitrienko, A. S. Romanuk, S. I. Suchkov, and Z. M. Uteshev, Zh. Tekh. Fiz. **53**, 2343 (1983).
12. R. Reininger, U. Asaf, and I. T. Steinberger, Phys. Rev. B **28**, 3193 (1983).
13. R. Reininger, U. Asaf, I. T. Steinberger, *et al.*, Phys. Rev. B **28**, 4426 (1983).
14. F. M. Jacobsen, N. Gee, and G. B. Freeman, Phys. Rev. A **34**, 2329 (1986).
15. A. K. Al-Omari, K. N. Altmann, and R. Reininger, J. Chem. Phys. **105**, 1305 (1996).
16. B. M. Smirnov, Usp. Fiz. Nauk **172**, 1411 (2002).
17. J. Lekner, Phys. Rev. **158**, 130 (1967).
18. M. H. Cohen and J. Lekner, Phys. Rev. **158**, 305 (1967).
19. B. E. Springett, J. Jortner, and M. H. Cohen, J. Chem. Phys. **48**, 2720 (1968).
20. V. M. Atrazhev and I. T. Iakubov, J. Phys. C **14**, 5139 (1981).
21. V. M. Atrazhev and E. G. Dmitriev, J. Phys. C **18**, 1205 (1985).
22. E. B. Gordon and A. F. Shestakov, Fiz. Nizk. Temp. **27**, 1192 (2001) [Low Temp. Phys. **27**, 883 (2001)].
23. E. B. Gordon, O. S. Rzhevskiĭ, and V. V. Khmelenko, Quantum Electron. **21**, 209 (1994).
24. E. B. Gordon, V. V. Khmelenko, and O. S. Rzhevskii, Chem. Phys. Lett. **217**, 605 (1994).
25. A. S. Schussler, J. Burghoorn, P. Wyder, *et al.*, Appl. Phys. Lett. **77**, 2786 (2000).
26. E. B. Gordon, G. Frossati, and A. Usenko, Zh. Éksp. Teor. Fiz. **123**, 846 (2003) [JETP **96**, 846 (2003)].
27. A. Usenko, G. Frossati, and E. B. Gordon, Phys. Rev. Lett. **90**, 153201 (2003).
28. B. M. Smirnov, Usp. Fiz. Nauk **171**, 1291 (2001) [Phys. Usp. **44**, 1229 (2001)].
29. R. A. Aziz and M. J. Slaman, Chem. Phys. **130**, 187 (1989).
30. A. K. Dham, A. R. Allnatt, W. J. Meath, and R. A. Aziz, Mol. Phys. **67**, 1291 (1989).

31. R. A. Aziz and M. J. Slaman, J. Chem. Phys. **92**, 1030 (1990).
32. A. K. Dham, W. J. Meath, A. R. Allnatt, *et al.*, Chem. Phys. **142**, 173 (1990).
33. E. Clementi and C. Roetti, At. Data Nucl. Data Tables **14**, 177 (1974).
34. A. D. McLean and R. S. McLean, At. Data Nucl. Data Tables **26**, 197 (1981).
35. A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Énergoatomizdat, Moscow, 1986; Springer, Berlin, 1985).
36. E. Fermi, Nuovo Cimento **11**, 157 (1934).
37. R. S. Berry, in *Theory of Atomic and Molecular Clusters*, Ed. by J. Jellinek (Springer, Berlin, 1999).
38. D. Wales, Adv. Chem. Phys. **115**, 1 (2000).
39. H. S. W. Massey, E. H. S. Burhop, and H. B. Gilbody, *Electronic and Ionic Impact Phenomena*, 2nd ed. (Claredon Press, Oxford, 1969).
40. L. G. H. Huxley and R. W. Crompton, *The Diffusion and Drift of Electrons in Gases* (Wiley, New York, 1974; Mir, Moscow, 1977).
41. B. M. Smirnov, *Physics of Ionized Gases* (Wiley, New York, 2001).
42. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory,* 3rd ed. (Nauka, Moscow, 1976; Oxford Univ. Press, Oxford, 1980).
43. A. J. Dahm, in *Progress in Low Temperature Physics*, Ed. by D. W. Brewer (North-Holland, Amsterdam, 1985), Vol. 9.
44. M. J. Druyvesteyn, Physica (Amsterdam) **10**, 61 (1930).
45. B. I. Davydov, Phys. Z. Sowjetunion **8**, 59 (1935).
46. P. M. Morse, W. P. Allis, and E. P. Lamar, Phys. Rev. **48**, 412 (1935).
47. B. I. Davydov, Phys. Z. Sowjetunion **12**, 269 (1937).
48. W. P. Allis and J. Allen, Phys. Rev. **52**, 703 (1937).
49. L. Ye. Gurevich, *Grounds of Physical Kinetics* (GITTL, Leningrad, 1940).
50. W. P. Allis, Ann. Phys. (Leipzig) **21**, 383 (1956).
51. B. M. Smirnov, *Physics of Weakly Ionized Gas* (Nauka, Moscow, 1978).
52. E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).
53. L. V. Lukin, Chem. Phys. **291**, 261 (2003).
54. V. M. Atrazhev, I. V. Chernysheva, and T. D. Doke, Jpn. J. Appl. Phys. **41**, 1572 (2002).
55. A. V. Eletskiĭ and V. D. Kulagin, Fiz. Plazmy (Moscow) **5**, 98 (1979) [Sov. J. Plasma Phys. **5**, 55 (1979)].
56. B. M. Smirnov, *Physics of Atoms and Ions* (Springer, New York, 2003).
57. I. Ya. Fugol, Adv. Phys. **27**, 1 (1978).
58. E. Morikawa *et al.*, J. Chem. Phys. **91**, 1469 (1989).
59. H. S. W. Massey, *Negative Ions* (Cambridge Univ. Press, Cambridge, 1976; Mir, Moscow, 1979).
60. E. Aprile *et al.*, Nucl. Instrum. Methods Phys. Res. A **338**, 328 (1994).

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 936–944.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1071–1081.

PLASMA, GASES

Electric Field Cumulation in Dissipative Structures of Gas-Discharge Plasmas

Ph. I. Vysikaĭlo

State Research Center of the Russian Federation, Troitsk Institute for Innovation and Thermonuclear Research, Troitsk, Moscow oblast, 142190 Russia
e-mail: filvys@rambler.ru, filvys@triniti.ru

Received August 5, 2003

Abstract—The process of cumulation of the dynamic order parameters (E/N and n_e/N) in dissipative structures of a gas-discharge plasma is simulated using the model of ambipolar drift and ionization. The model is applicable for explaining the onset of electric energy cumulation (for $E^2/8\pi \gg H^2/8\pi$) at the periphery of electric arcs; beaded, ball, and streak lightning; cathode spots; and other spherical, cylindrical, conical, and planar dissipative plasma structures, viz., the plasmoids previously observed in "enigmatic" phenomena and in experiments with nonequilibrium gas-discharge plasmas. It is shown that, in contrast to Turing regular dissipative structures (1952), the nonlinear profiles of the dynamic order parameters in cumulative–dissipative structures (Vysikaĭlo, 1996) are described not by diffusion processes, but by convective processes of ambipolar drift focusing the electric field density ($E^2/8\pi$). © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The methods for the obtaining and prolonged confinement of extremal states of matter are being intensely studied. The knowledge of such states is essential, for example, for developing new technologies and for synthesizing new materials. The study of Coulomb cumulators (i.e., dissipative plasma structures in which cumulation of energy, mass, and electric field occurs in charged focused mass flows) as the most powerful attractors in connection with focusing electric forces for $E^2/8\pi \gg H^2/8\pi$ is essential.

The role of cumulation is apparently decisive in processes of energy and mass transfer through the interface between different media (with different phase and other states). It is for this reason that cathode and anode spots are formed at the metal–gas, metal–liquid, and other interfaces. In dissipative structures self-formed at the interfaces between different media, the corresponding cumulative processes responsible for a peculiar geometry of transition layers take place. These layers focus the energy and mass flows. Such a structural (anisotropic) activation of media renders the characteristics (dynamic order parameter) closer. This constitutes the goal of formation of cumulative-dissipative structures (plasma electrode spots) in the media. Dynamic plasma structures (cathode spots, lightning, and arcs) are plasmadynamic analogs of hydro- and gasdynamic regular systems such as a sandstorm, tornado, hurricane, drain, gully, swirl, etc. (see Figs. 1–3).

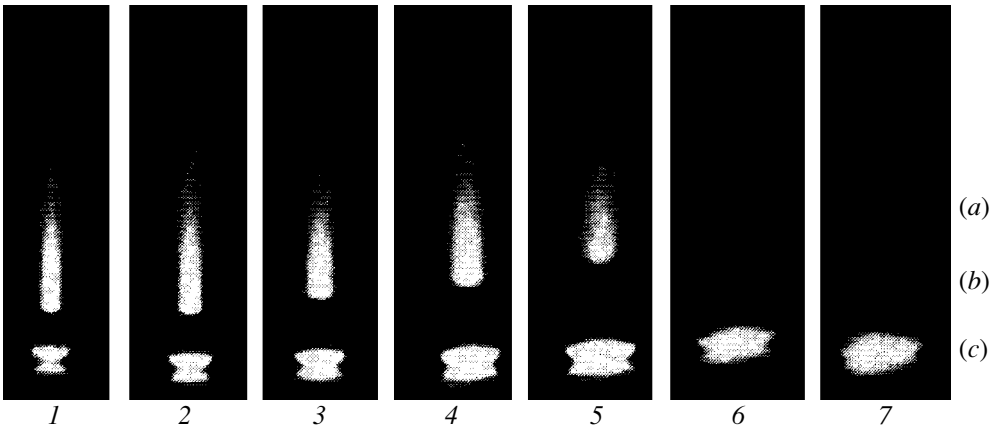


Fig. 1. Evolution of the width of a cylindrical self-focused discharge with increasing current (from left to right); positive column (*a*), Faraday dark space (*b*), and cathode glow and its reflection in a mirror-polished copper cathode (*c*).

1063-7761/04/9805-0936$26.00 © 2004 MAIK "Nauka/Interperiodica"

In the absence of a cathode spot, the conductivity of the medium is negligibly small (for a current density of $\sim 10^{-10}$ A cm^{-2}), while the density of the strongest current in the spot can attain a value of 10^8 A cm^{-2} [1]. The cumulation coefficient λ for current density (j) in the spot relative to a nonactivated and nonstructured medium can be as high as $\lambda_j = j/j_0 = 10^{18}$. What processes are responsible for such cumulation? It appeared that this effect confirms the decisive role of microexplosions of cathode regions, viz., explosive emission of electrons [1]. In accordance with the theory of ectons (portions of electrons from cathode spots), the reason for microexplosions is Joule heating of microtips (inhomogeneities of the cathode). Then why is a cathode spot with a normal current density formed on a mirror-finished electrode (see Fig. 1 from [2])? The effect apparently cannot be explained only by the tip geometry and Joule heating. Joule heating and the explosion of a cathode region are rather consequences of current density cumulation that the reason for this effect. How can we explain the processes of temporal and spatial proportionality in all dissipative plasma structures? The proportionality for cathode spots in space and time has been established experimentally by Kesaev [3]. The "reciprocal" displacement of a cathode spot in a transverse magnetic field (Stark, 1903) remains enigmatic both in the one-dimensional classical model and in the ecton theory [1]. This effect discovered by Stark clearly showed that the cathode spot plasma contains electrons moving not only to the positive, but also to the negative electrode (cathode). (It is these electrons that determine the displacement of the cathode spot as an integral dissipative plasma structure (plasmoid) and not a cumulative jet (high-energy electron beam directed to the anode).) The nearly spherical geometry of the cathode spot and the geometry of the Faraday dark space, coinciding with the hourglass geometry or with the classical Laval nozzle geometry remain unclear (see Figs. 1, 3).

The electric field cumulation and the corresponding focusing of energy and mass flows in dissipative plasma structures (including cathode spots) have been investigated insufficiently. In accordance with the propositions formulated by the author in [4], all dissipative structures cumulate energy–mass flows, forming the regions of focusing or attraction (attractors) in the bulk. In these extended cumulative structures, the type of elements of flows being focused changes under the action of the self-focusing field. The specific energy received from an external force or energy field is enhanced and redistributed in attractors and new degrees of freedom are excited in the medium. Only after this do the energy–mass flows dissipate to the surroundings in a new form. The flows activated in cumulators of dissipative structures are anisotropic and ordered.

In simulating experimental observations for elevated gas pressures under the condition $\mathbf{E}^2/8\pi \gg \mathbf{H}^2/8\pi$, where $\mathbf{E}$ and $\mathbf{H}$ are the electric and magnetic field strength, transitions layers or lenses focusing the electric field and the corresponding flows of charged particles will appear only if the ambipolar drift is taken into consideration. The ambipolar drift leading to focusing effects of this type under elevated gas pressures, when the role of diffusion is small, was disregarded in the models considered earlier [5, 6].

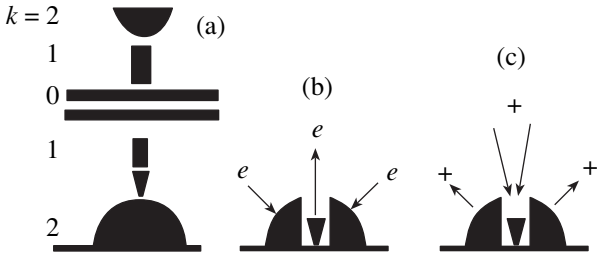


Fig. 2. Cathode spots. (a) Example of a possible arrangement of structured plasmoids with different symmetries ($k = 0$ correspond to planar, 1 to cylindrical, and 2 to spherical symmetry). (b) Arrows indicate the directions of cumulation of electron flows and of the reduced electric field, which explain the reciprocal motion of the cathode spot in a transverse magnetic field as a result of cumulation and formation of a cumulative electron jet responsible for the Faraday dark space. Electrons appear in the bulk in the spot region, for example, due to UV preionization. (c) Corresponding diagram of cumulation of ion flows to the cathode spot.

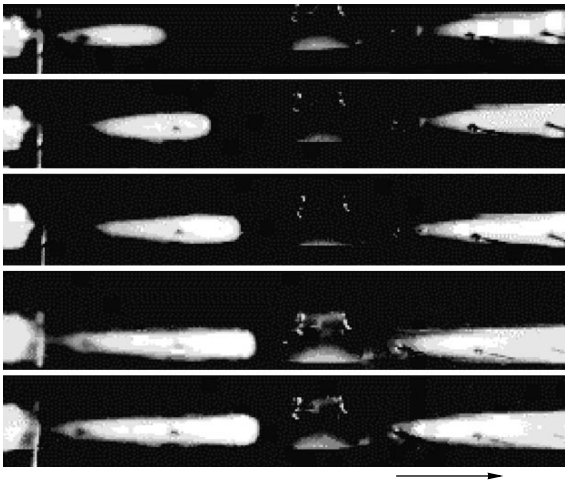


Fig. 3. Discharge in a tube with increasing discharge current. Comparison of the photographs shows the evolution of the jump width with increasing current (from top to bottom) [15, 16]. The arrow marks the region of self-formation of a conical discharge. The gas flow velocity is directed oppositely and amounted to $U = 50$ m/s.

In 1970s, Prigogine used the term dissipative structures as applied to the energy–mass flows ordered by an external force into structures. At present, brilliant examples of chemical dissipative structures breaking spatial symmetry are known. These are Turing structures, in memory of A. Turing, who was the first to put forth (in 1952) the hypothesis that the interaction

between nonlinear chemical reactions and diffusion may lead to the formation of ordered spatial structures differing in the concentrations of reagents [7]. Numerical simulations performed on the basis of a brusselator (a very simple model of chemical reactions, which satisfies all the requirements necessary to the formation of dissipative structures) makes it possible to trace the transition from a spatially homogeneous system to spatially structured states upon a deviation of the system from thermodynamic equilibrium and the emergence of various alternating geometric structures, each of which possesses a quite definite dynamic stability and spatial symmetry.

Cathode spots, as well as Turing structures, may organize regular structures [3, 8]. In regular dissipative systems of cathode spots, elements of dissipative structures (individual spots) are arranged at equal or nonlinearly ordered distances [8]. Nonlinearities in transfer processes, as well as in sources and sinks of charged particles (in particular, the difference in the electric field dependences of the mobilities of electrons and ions and the internal electric field which prevents charge separation), lead to ambipolar drift via which various phases are formed and matched in a regular system. Transient profiles between different phase states of the medium under elevated pressures $P > 10$ torr and $n_e \sim 10^{10}$ cm^{-3} are described by ambipolar drift. Its velocity is similar to that of the velocity of sound.

It was proposed in [4, 9–11] that systems of dissipative structures forming a single entity be referred to as dissipative crystals. Publication [12], in which the concepts of quasi-crystals and turbulent crystals were introduced for ordered flows in liquids, is also worth mentioning. The structurization of the medium in such systems ensures geometrical focusing (cumulation) of energy–mass flows in definite directions. This process is associated with self-formation of transition layers filtering, discriminating, selecting, ordering, and focusing the energy, momentum, and mass flows (such layers play the role of semi-penetrable membranes, walls, etc.) as well as cumulative jets (channels subdivided into stings and jets). In these layers and cumulative jets, cumulation of not only flows, but also of reduced force field takes place (in a gas-discharge plasma, this is the parameter E/N, where N is the density of the gas). In real focusing processes, cumulation with the spherical, cylindrical, and planar symmetry can take place.

Accordingly, several types of symmetry can be distinguished in attractive structures and their cumulative–dissipative regular systems formed by dissipative structures in a plasma (see Fig. 2). Several varieties of structural or cumulative thermal conductivity, electrical conductivity, etc., are also possible. For example, a layered structural cumulative conductivity arises in plasmas, which is manifested in the form of glowing strata parallel to plane electrodes. An arc, lightning, or a linear crack in a metal appears in the case of cylindrical cumulation. Cylindrical cumulation in lightning facilitates the formation of a cumulative "enigmatic" jet of electrons with an energy of several megaelectronvolts, which propagates continuously or in pulses in front of a glowing cumulative-dissipative regular linear system with $L \gg r$, where L and r are the longitudinal and radial dimensions [13]. In the approach proposed here, a cathode spot is a spherical cumulator. In a hydrogen discharge plasma, the author and Shashkov observed cumulation of the cup-in-cup type, where bright cylindrical layers were parallel to one another with the axis perpendicular to plane electrodes.

Focusing (attractive) structures with various types of symmetry form complex geometrical cascades (see Fig. 3).

In plasma-type (charged) cumulators, the collapse of energy–mass flows may result in the penetration of the Coulomb barrier by nuclei; i.e., "cold" transmutation of nuclei as a result of globally local resonances can take place in cumulative-dissipative systems.

The importance of studying extended cumulative Coulomb dissipative structures (collapsing energy–mass flows) and their cumulative–dissipative regular systems, as well as profiles of dynamic order parameters and their discontinuities playing the role of channels, membranes, and potential walls in such systems, becomes obvious and the practical value of investigations in this field may turn out to be unlimited.

2. FORMULATION OF THE PROBLEM

Comprehensive numerical simulation of dynamic self-consistency of all processes associated with passage of current in the formation of the geometry of a cathode spot, Faraday dark space, anode region, and cumulative jets from the cathode, which participate in the formation of these objects, remains an extremely cumbersome and complicated three-dimensional nonstationary problem that has not been correctly formulated and solved so far. For this reason, it is expedient to study simple one-dimensional models of discharge elements possessing a certain spatial symmetry according to the results of experimental observations, which, however, help to clarify the role of ambipolar drift in cumulation of the electric field.

Here, we will study planar as well as spherically and cylindrically symmetric stationary cumulation of the dynamic order parameters (E/N and n_e/N) in plasmoids, which is associated with ambipolar drift (convective ambipolar flows arising due to nonlinearities in transport processes) [2, 14–22].

The model proposed here can be used for describing the processes of the formation of transition layers or focusing membranes in a cathode spot having the visually observed shape of a hemisphere or cone. The electron and ion current are focussed precisely on this spot in structurally different ways (since electrons and ions

move in opposite directions). As current collapse progresses, a new state of the plasma with a normal current density is formed, which is visualized differently (in the form of a bright glow) [4]. Formally, it is assumed that the electron current from the cumulation center is connected by a narrow high-conductivity cumulative jet (defocusing cumulative electron jet) to the positive column (see the diagram in Fig. 2 and Fig. 1). The geometrical size r of the jet is related to the part R of the cathode spot accumulating electrons via the current conservation law. The jet, together with recombination processes and ambipolar drift, determines the geometry of the Faraday dark space [4]. The cathode spot openness parameter for electrons is $\chi \sim (r/R)^2$. In solving problems with radial cumulation, the size of the sting (transformed into a jet) can be disregarded as compared to the radial size of the self-formed plasmoid ($\chi \ll 1$). In the region of the electron jet at the cathode spot, an ion flow focusing element is formed. The characteristic size of this quasi-neutral region will be obtained analytically in this study. The process of mutual focusing of the electron and ion flows in this region leads to the formation of a compact (self-focusing) cathode spot.

The production of electrons and ions in the region with appreciable radial fields is ensured by direct ionization; at the periphery of the structure, the UV preionization of impurities probably plays an important role. Cumulation of profiles in the region of radial fields in the proposed model is determined by ambipolar drift of the plasma due to different electric field dependences of the electron and ion mobilities [2, 14–22] and a sharp (exponential) electric field dependence of the ionization rate. The model of a high-conductivity cumulative sting was formulated for the first time in 1996 in [17], where the corresponding estimates were given in the ambipolar drift-diffusion approximation. This model differs from conventional 2D models in the boundary conditions at the center of the structure being simulated. (In conventional models, e.g., with cylindrical symmetry, the condition leading to the Bessel or diffusion profiles of the dynamic order parameter at the center of the structure is imposed, while the model proposed here accounts for possible unlimited cumulation of energy and mass flows and the dynamic order parameter to the cumulation center.) Let us consider the emergence of cumulation of the dynamic order parameter in detail. We proceed from the basic kinetic concepts using the Boltzmann equations for a system of charged particles in a gas-discharge plasma.

The complete system of Maxwell equations and the kinetic transport equations for electrons and ions in the bulk of a simple gas-discharge plasma at a large distance from the electrodes (and, hence, without a number of familiar problems associated with the boundary conditions at the cathode) are usually solved in perturbation theory disregarding displacement currents (see, for example, [14]). The following ratios play the role of small dimensionless parameters: l_u/L, l_E/L, $\omega\tau_M$, μ_i/μ_e, and $j_i/j_e \ll 1$. Here, L is the characteristic size and ω is the characteristic frequency of inhomogeneities under investigation; $l_u = \varepsilon/eE$ is the energy length over which electrons acquire the characteristic energy ε; $\mathbf{l}_E = \mathbf{E}/(4\pi e n_e)$ is the vectored length of variation of the electric field strength $\mathbf{E}$; e is the electron charge, $\tau_M = 1/(4\pi e n_e \mu_e)$ is the Maxwell time; and μ_e, j_e, μ_i and j_i are the mobilities and current densities for electrons and ions, respectively. The vectored characteristic length of variation of the electric field strength is determined by the concentration of electrons [18] and not ions, as erroneously assumed in [19]. Under definite conditions (in the regions where the neutrality is noticeably violated), this inaccuracy leads to an analytic errors of several order of magnitude in the size of the Coulomb structure or its transient profiles. We apply the terms "Coulomb structure" or "Coulomb cumulator" to a dissipative structure in a gas-discharge plasma, in which the electric field is focused and the role of the magnetic field in this case is insignificant.

3. ANALYTIC MODELS AND NUMERICAL CALCULATIONS

In stationary and quasi-stationary plasmoids (in which displacement currents can be neglected and the plasma consists of electrons and one species of ions), the profile dynamics of the dynamic order parameters (E/N and n_e/N) can be reduced to a single nonlinear Burgers-type equation [14, 18, 20]

$$\frac{\partial n_e}{\partial t} - \frac{\partial \mathbf{l}_E}{\partial t} \cdot \nabla n_e + \mathbf{\Gamma}_0 \cdot \nabla\left(\frac{\mu_i}{\mu_e}\right) - \nabla((D_E + D_a)\nabla n_e) = Q_k, \quad (1)$$

where

$$D_a = \frac{\mu_e D_i + \mu_i D_e}{\mu_e + \mu_i}, \quad D_E = \mu_i \mathbf{E} \cdot \mathbf{l}_E,$$

D_i, D_e, μ_i, and μ_e are the diffusion coefficients and mobilities of ions and electrons, $\mathbf{\Gamma}_j$ and $\mathbf{\Gamma}_e$ are the fluxes of these quantities, and $\mathbf{\Gamma}_0 = j_0/e$, j_0 being the preset current density. The condition $\mathbf{\Gamma}_0 = \mathbf{\Gamma}_i + \mathbf{\Gamma}_e = \text{const}$ determines the relation between the electric field strength and the electron concentration, while the quantity Q_k describes the creation and annihilation of electrons. In the proposed model, it is the product $\mathbf{\Gamma}_0 \cdot \nabla(\mu_i/\mu_e)$ that determines the ambipolar drift convective focusing of the electric field or the entrainment of the plasma profiles by ambipolar drift. This term differs from zero if the field dependences of the mobilities of electrons and ions are different; in this way, nonlinearity in transport processes occurring in a heterogeneous plasmoid is manifested.

Ambipolar drift (and, hence, convective focusing) may be due to nonuniform and nonstationary form of the electron distribution function in sources and sinks

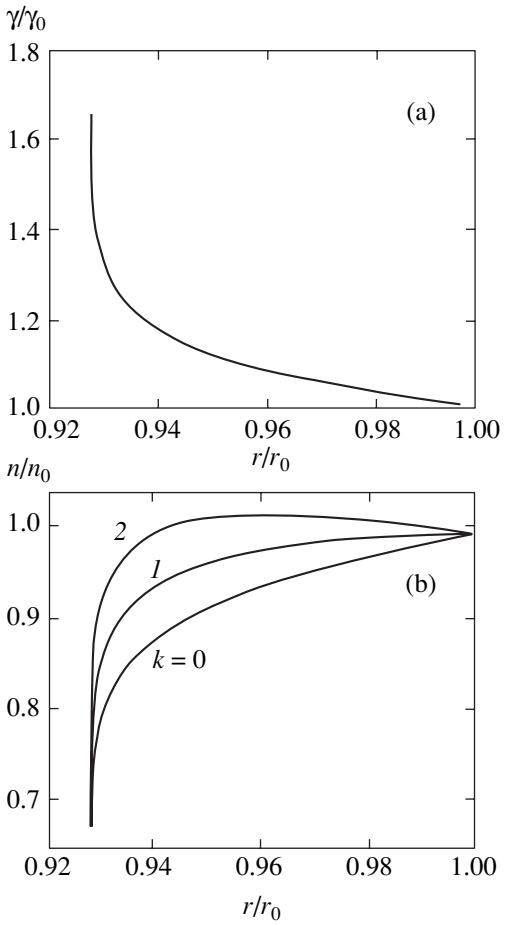


Fig. 4. Profiles of (a) electric field strength $E/E_0 = \gamma/\gamma_0$ and (b) reduced electron concentration as functions of r/r_0 in the ambipolar drift–ionization approximation for various values of k (profiles for $k = 0, 1, 2$ in (a) coincide).

of ions (in Q_k) [20], plasmachemical reactions involving ions [18], violation of neutrality in an inhomogeneous and nonstationary plasma (second term in Eq. (1)) [14], the presence of gas circulation, external magnetic field [21], etc. The quantity $\nabla\mathbf{E}$ can be expressed in terms of ∇n_e only in the 1D approximation in x [14, 18, 19] or in r [4, 9–11, 17].

In the approximation of reaction–ambipolar drift transport, a stationary quasi-neutral profile is described by the differential equations following from Eq. (1):

$$\frac{d(Br^k n_e\gamma)}{dr} = -r^k n_e(\nu - \beta n_e), \qquad (2)$$

where $k = 0$, 1, and 2 describes the planar, cylindrical, and spherical symmetry, respectively. The reaction is accounted for in Eq. (2), first, in the form of direct ionization by an electron impact with frequency ν and, second, under the assumption that the destruction of the plasma is determined by dissociative recombination with effective coefficient β.

Equation (2) makes it possible to numerically calculate the coefficients of quasi-neutral cumulation of the electron (and ion) concentration and the reduced electric field strength ($\gamma(r) = E/N \times 10^{17}$ Td^{-1}) as functions of r. It should be noted that the effective coefficient β of recombination of electrons with ions sharply decreases for $\gamma(r) > 40$, which enables us to disregard recombination in the region of stationary quasi-neutral cumulation of parameter γ. If we represent the drift velocities of electrons and ions in the form $\mu_e E = C\gamma^\alpha$ and $\mu_i E = B\gamma$ (where C = const and B = const), we can analytically derive from Eq. (2) the reduced field profile $\gamma(r)$ in the quasi-neutral approximation:

$$\gamma(r) = -\frac{1}{A}\ln\left[\exp(-A\gamma(0)) + \frac{A\nu_0(r - r_0)}{B(1-\alpha)}\right]. \qquad (3)$$

It can be seen from this relation that, for

$$r_1 = r_0 - \frac{B(1-\alpha)}{A\nu_0\exp(A\gamma(0))},$$

the value of $\gamma(r_1) = \infty$. The size of the region of transition from $\gamma(0)$ to $\gamma(r_1) = \infty$,

$$r_\nu = \frac{B(1-\alpha)}{A\nu_0\exp(A\gamma(0))},$$

is determined by the ion mobility (parameter B), the index $(1 - \alpha)$ of nonlinearity of drift flows of electrons and ions relative to each other, the strength of the electric field dependence of the direct ionization rate on the (parameter A), and the value of the ionization frequency $\nu_0\exp(A\gamma(0))$ for r_0.

Consequently, if the nonlinear process of ambipolar drift–ionization determines the cathode spot radius, the size of the cathode spot must vary in accordance with the above-mentioned parameters; with increasing pressure, the spot radius must sharply decrease. Cumulation coefficient λ_n of plasma concentration is determined in accordance with Eq. (3) and follows from the condition of conservation for the electron current (in the approximation $j_i/j_e \ll 1$):

$$\lambda_n = \frac{n_e}{n_e(0)} \propto -A\left(\ln\left[\exp(-A\gamma(0)) + \frac{A\nu_0(r - r_0)}{B(1-\alpha)}\right]\right)^{-\alpha} r^{-k}. \qquad (4)$$

For $\alpha = 1$, ambipolar drift becomes equal to zero and cumulation profiles of the order parameters disappear. In this case, other transport processes must be taken into account for describing experimental profiles.

In accordance with Eq. (3), the reduced field cumulation ($\lambda_\gamma = \gamma/\gamma(0)$) is independent of the symmetry type or the geometry of the Coulomb ($E^2/8\pi \gg H^2/8\pi$) plasma cumulator ($k = 0, 1, 2$). Figure 4a shows the profile of the reduced field cumulation coefficient λ_γ, which qualitatively reflects dependence (3). In accor-

dance with relation (4), the cumulation of concentration or the degree of ionization of the gas is determined to a considerable extent by the symmetry type and depends on k (see Fig. 4b). For the three one-dimensional cases strongly differing in the electron concentration profiles, we analytically derived the same reduced field profile, which is close to the linear profile (in space) at the boundary of the cathode layer. Thus, relation (3) is based on the assumption made in the Engel–Steenbeck model concerning the linearity of the reduced electric force profile (E/N) in the region of its cumulation at the cathode (see Fig. 4a).

It should be noted that the analytic solution to Eq. (2) in the ambipolar drift–recombination approximation, which describes the Faraday dark space, were obtained earlier in [22] for $k = 0$.

Here, we obtain analytic and numerical solutions for $k = 1$ and 2. The solution can also be sought in the form of a profile $\gamma \sim r^{\eta}$; in this case,

$$\eta = \frac{1-k}{1+\alpha}. \quad (5)$$

In the nitrogen gas-discharge plasma, $\alpha = 3/4$, and in the planar symmetry ($k = 0$), $\eta = 4/7$ in accordance with relation (5). The solution in [22] was obtained for $\eta = 1$. This is associated with a less accurate approximation of ambipolar drift as a function of γ, which was chosen by the author in [22]. For $k = 1$, we have $\eta = 0$, while for $k = 2$, the value of η becomes negative, but cumulation of γ does not arise; in all these cases, only the derivatives of γ and n_e increase (see Fig. 5).

Model (2) is applicable for describing nonuniform profiles of the dynamic order parameter and current collapse in the range $10^{-2} < \gamma < 200$ and encompasses a wide spectrum of nonlinear dynamic effects, including electric field cumulation in a nonequilibrium gas-discharge plasma.

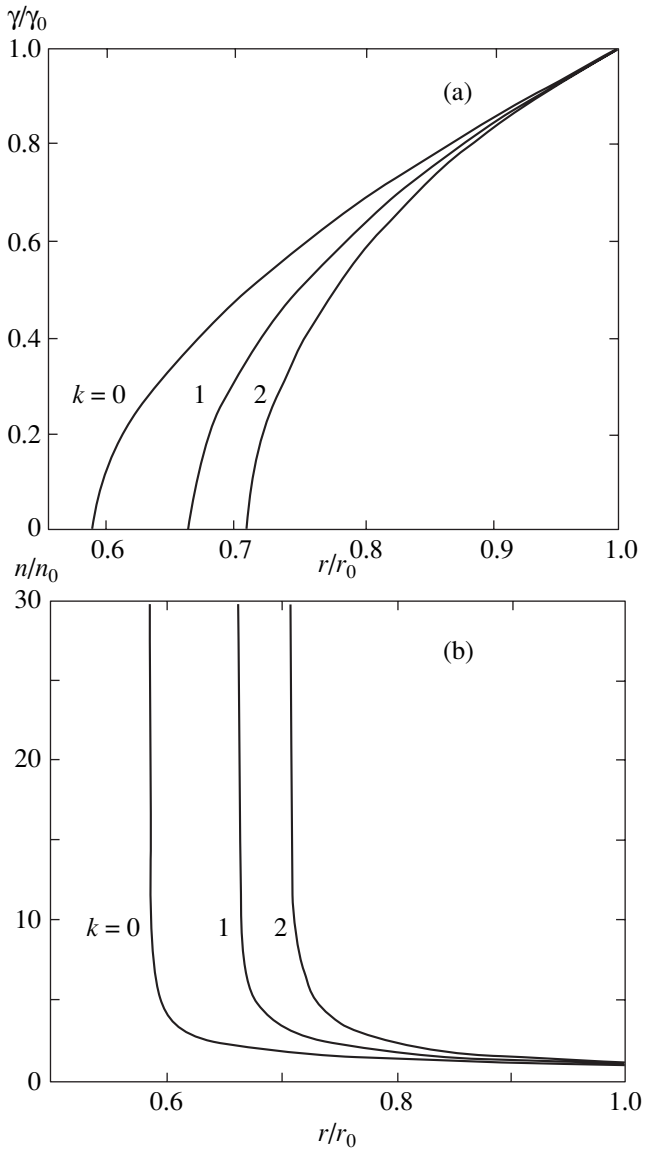


Fig. 5. Profiles of (a) electric field strength $E/E_0 = \gamma/\gamma_0$ and (b) reduced electron concentration as functions of r/r_0 in the ambipolar drift–recombination approximation for various values of k.

4. DISCUSSION

In analytic and numerical calculations, it was assumed that the model operates in the case of cylindrical symmetry also and can be used for describing processes of radial cumulation in a nonequilibrium plasma at the periphery of lightning or an extended arc. In this case, a high-conductivity cumulative jet must be formed in a real physical object; radial electron flows collapse to this jet in the region of the cumulator. The model proposed here precisely describes the characteristic radial profiles in such an attractor.

We have demonstrated here that cumulative processes or Coulomb ($E^2/8\pi \gg H^2/8\pi$) self-focusing of the electric field in plasmoids not only exist, but can also be simulated analytically and numerically in the 1D approximation. Qualitative dependences of cumulating profiles of the dynamic order parameters (γ/γ_0 and n/n_0 on r/r_0) are shown in Figs. 4 and 5. It can be seen that, in the given formulation of the problem, simultaneous cumulation of plasma concentration n and reduced field γ in a Coulomb cumulator is possible only in the case of spherical symmetry ($k = 2$).

The established electric field cumulation in a plasma accompanied by the simultaneous mutual organization of substructures with different symmetries is of special practical importance. Even Eq. (5) and Figs. 1 and 3 demonstrate that a complex play of self-organizing structures and their sizes with various cumulation profiles and defocusing of the dynamic order parameters is possible. The mode of formation of cumulative jets in plasmoids in the dynamics of media self-organization needs to be determined. Nature optimally decides itself which bricks, cylinders, cones, or balls are more suit-

able to be assembled into structures and their ordered systems, viz., cumulative–dissipative "crystals" with elements possessing different types of symmetry (see Figs. 1–3). The problems of self-consistency of characteristic sizes, frequencies, nonlinear resonances, pulsation, and other dynamic order parameters have to be solved by researchers. However, even analysis of simple 1D models shows that, for different values of $k = 0$, 1, 2, an activated medium may change the profiles of dynamic order parameters (including characteristic sizes and frequencies) in the bulk of a gas-discharge plasma (see Figs. 3, 4b, and 5).

In accordance with the generalized Gauss theorem, characteristic sizes r, fields ($F_m \sim 1/r^\eta$), forces ($F \sim 1/r^\eta$), and, by virtue of Newton's second law ($\omega(r) = (F_m/r)^{0.5}$), the characteristic frequencies of processes, $\omega \sim 1/(r^{\eta+1})^{0.5}$, cumulate in cumulators (in the regions of attraction and focussing) [10, 11]. The law of cumulation (η) changes with the excitation of new degrees of freedom (and in the course of electric field cumulation).

In the general case, the presence of ambipolar diffusion, as well as diffusion caused by neutrality violation (see Eq. (1)), does not eliminate the electric field cumulation altogether [11]. However, the length of the attractor region changes accordingly. From the balance of gradients (in the model of ambipolar drift and ambipolar diffusion associated with violation of neutrality), at the boundary of the cathode spot, we can derive the equations of topoenergy and tempoenergy states of a plasmoid element (or elements of a cumulative–dissipative regular system):

$$r_E \sim \frac{D_E}{V_{ar}} = 5.6 \times 10^5 \frac{E/N}{n_e/N},$$

$$\omega_E \sim \frac{V_{ar}^2}{D_E} \sim \omega_E \left(\frac{n_e}{N}\right).$$

Here, the cathode spot radius r_E is measured in cm, E in V cm^{-1}, N and n_e in cm^{-3}, and D_E is the diffusion coefficient associated with violation of neutrality [14]. In this limit, a cathode spot radius differing from r_v was obtained in [4, 17]. When the dynamic order parameters attain their critical values, nonlinear media generally exhibit complication of the geometry of flows (with new degrees of freedom) in space and time as well as stratification with a constant characteristic size or interval (macroquantization in space and time) accompanied by the formation of cumulative–dissipative regular systems.

The presence of several types of diffusion processes (classical ambipolar diffusion and the diffusion associated with neutrality violation) in a plasma may lead, in view of dispersion, to structurization of space-time structures in a gas-discharge plasma as in the case of plankton distribution in the presence of two types of diffusion [23]. For this reason, structured energy and mass transfer and consecutive modification of energy and mass flows can take place in cumulative–dissipative regular systems. In alternating attractors (regions of focusing, or attraction, of flows), cumulative jets (stings and jets) possessing cumulative properties are formed in this case. It is cumulative jets (high-energy beams) and the geometry of structural cumulation of dynamic order parameters that are responsible for the Faraday dark space and its geometry similar to that of the Laval nozzle (hourglass). Cumulative–dissipative processes may form a pulse-periodic, coherent, stationary, or chaotic pattern in the phase space-time with a symmetry of self-organizing structures, which varies in space and time (see Fig. 2).

Thus, we return to the ideas formulated by Kesaev, according to which the cathode spot (as a new dynamic state of plasma cumulating energy and mass flows) is discrete in space and time [1–3] and is characterized by resonance frequencies, nonlinear resonances in space and time, and a hierarchy of internal sizes.

In accordance with the model proposed here, it is the electric field cumulation in the spot associated with ambipolar drift that leads to the explosion of cathode inhomogeneities; this in turn ensures electron emission. Cathode spots and their systems form a multihierarchic cumulative–dissipative regular system with nonlinear resonances associated with cumulative processes in Coulomb attractors.

Analysis of experimental observations and numerical calculations leads to the conclusion that radial ambipolar drift (V_{ar}) may be responsible for the initial stage of cumulation of the dynamic order parameter in the cathode spot plasma [4, 9–11, 17, 20].

The cumulative–dissipative model describes the normal current density at the cathode spot ($j^* = en_e^* V_e (E/N^*)$). The model enables us to estimate the total current in lightning or arcs ($I = Sj^*$, where S is the area of the lightning surface) from the visual characteristics. Using this model, we can explain the reciprocal motion of the cathode spot in a transverse magnetic field (see Fig. 2), estimate its velocity ($\mathbf{V}_a \approx \mu_i[\mathbf{V}_e^* \times \mathbf{H}]/c$), and calculate the spectral power ($W = F^*V \sim 1/\omega^\alpha$, $\alpha = -(3\eta - 1)/(\eta + 1)$ for $F \sim 1/r^\eta$) of flicker noise from the operation of cathode spots and other Coulomb attractors (plasmoids) with a focused electric field and charge particle flows [4, 10]. In the framework of cumulative–dissipative dynamics, we can consider the cumulation of electrons into a cumulative jet (see Fig. 2).

The cumulative–dissipative approach developed here can be used for studying a number of enigmatic phenomena with cylindrical symmetry existing in gas-discharge plasmas. For example, the concentration of energy in the bulk of lightning, which obviously corresponds to its luminous intensity, can be attributed to cumulative radial processes induced by radial ambipolar drift. The existence of radial cumulation of the electric field also explains the emergence at the center of a

lightning of electrons having energies on the order of megaelectronvolts, causing a 10-m long air breakdown under the atmospheric pressure [13] (for a high-energy electron beam degradation on the order of 1 MeV m^{-1}). There are no exploding microtips in the sky, which would ensure the emission of such electrons. However, such energy may appear as a result of geometrical cumulation of the electric field in extended Coulomb attractors. Processes of radial electric field cumulation at the periphery of lightning also explain the fact that a 1-km long flash of lightning preserves its transverse size.

From the standpoint of cumulative dynamics, the Lawrence and Dennington experiments (1930–1931), as well as the Nagaoka and Siguira experiments (1924), become less enigmatic (see [13], pp. 514–518). In these experiments, the broadening of the Zn line was discovered, which corresponds to electric field strengths on the order of 10^5 V cm^{-1} (if we attribute this broadening to the Stark effect). Such field strengths were observed in a spark channel with a width approximately equal to $r = 0.02$ cm for a discharge gap of $L = 1$ cm at a voltage of a few kilovolts ($U = 2 \times 10^3$ V) applied to the electrodes. This can be transverse electric fields ($E_r \sim U/r \sim 10^5$ V cm^{-1}) rather than longitudinal fields. Both transverse and longitudinal fields lead to the Stark effect, but transverse fields (focused to the center of the regular cylindrical system, viz., electric arc) should not be integrated over the gap as in the case of longitudinal fields $\left[\int_0^L E_x dx = U(L)\right]$, as was done by Nagaoka and Siguira [13]. Such errors are obvious in the formation of transverse pressure gradients or transverse hydrodynamic fields in a drain (cylindrical hydrodynamic attractor or cumulator of energy and mass flows). The hydrodynamic analogy suggests that radial cumulation of electric fields in plasmoids (visualized in space and time in the form of dissipative–cumulative regular systems ordered at large distances) is possible in an electric arc as well (see Figs. 1 and 3).

Self-organization of the attractor described by Eq. (2) occurs under the action of an external electric force or an energy difference. The external generalized (electric in the present case) force sets the activated elements of the medium in motion and orders their convective (ambipolar in the case of a plasma) flows. The activated elements are ordered by this force to such an extent that the reduced force itself is cumulated in the attractor or cumulator formed by it. The cumulative or explosive profile of the electric field is described by Eq. (3). Such organization–self-organization processes lead to complex cumulation of dynamic order parameters of external factor (the reduced external force E/N in our case) in an activated nonlinear medium as well as of dynamic order parameters determining the extent of activation of medium elements (the temperature of various degrees of freedom, the degree of ionization, etc.). Convective processes of ambipolar drift in a nonequilibrium gas-discharge plasma ensure the cumulative–dissipative self-organization of the medium and cumulation of the electric field strength. The remaining processes (including the formation of tips at the subsequently exploding electrodes) may be secondary processes relative to self-cumulation of the electric field strength.

Thus, in the case of self-organization of a continuous medium (including nonequilibrium gas-discharge plasmas), external forces induce in the medium not only opposing forces (in accordance with the Le Chatelier principle), but also orthogonal forces, which can locally exceed considerably the mean values of forces distributed over the entire activated volume of the medium in the attractor or cumulator formed by these forces.

On the basis of available experimental observations in gas-discharge plasmas, it can be concluded that regions with longitudinal electric field, as well as with transverse or radial (relative to the total current vector) fields, can appear during the formation of plasma structures. After the attainment of critical values, cumulation of the reduced force excites new degrees of freedom in elements of the medium being activated and forms dissipative structures as well as their systems, viz., "dissipative channels" with cumulative jets (or stings). The fields collapsing towards the center in the radial directions form cumulative focused stings. In regular cumulative–dissipative systems, the elements of the activated medium (especially in stings) acquire cumulative (hyper) properties [4, 9–11].

In this study, a simple ambipolar drift mechanism has been proposed and tested. The mechanism describes self-cumulation of the electric field strength or volume charge in a gas-discharge current-carrying plasma for low concentrations of plasma particles. Such an ambipolar focusing mechanism can be associated, for example, with different electric field dependences of mobilities of electrons and ions. The author believes that this mechanism operates in the formation of the periphery of cathode spots, arcs, lightnings, etc. It is well known that the formation of a cathode spot is accompanied by pulsed (cumulative) transport of the cathode material to the interelectrode gap [3, 8]. This effect is associated with electric energy cumulation in dissipative plasma structures and is successfully explained by the model proposed here. In contrast to the Turing model, the transport is executed not by diffusion, but by ambipolar drift.

However, further studies of electric field cumulation in a medium activated by the field and of the formation of extended Coulomb cumulators and their dissipative crystals, viz., regular systems in a gas-discharge plasma, are required.

Arcs and other cumulative–dissipative crystalline systems exist in a medium, while a medium with cumu-

lative stings and with boundaries cumulating energy and mass flows is self-formed even in vacuum.

The formation of attractors (or regions of focusing of energy and mass flows) is a multidimensional property of any continuous nonlinear medium. A more detailed study of volume charge cumulation in cathode spots will make it possible to clearly visualize and explain the effects of convective Coulomb focusing of electric field as a form of energy in remote stars, molecular and atomic clusters, and even in nuclei of microscopic structures.

The analytic treatment of static cumulation of the dynamic order parameters in a gas-discharge plasma has been carried out in complete accordance with [24] and can be generalized to the processes of electric field dynamic cumulation. The Coulomb self-focusing exists and cannot be absent since the existence of Coulomb forces is evident.

REFERENCES

1. G. A. Mesyats, *Ectons in Vacuum Discharge* (Nauka, Moscow, 2000).
2. Ph. I. Vysikaĭlo, A. F. Glova, and M. M. Smakotin, Fiz. Plazmy (Moscow) **14**, 734 (1988) [Sov. J. Plasma Phys. **14**, 434 (1988)].
3. I. G. Kesaev, *Cathode Processes in Electric Arc* (Nauka, Moscow, 1968).
4. Ph. I. Vysikaĭlo, in *Proceedings of Seminar on Synergetics* (Mosk. Gos. Univ., Moscow, 2001), Vol. 4, p. 106.
5. Yu. P. Raĭzer, *Modern Physical Foundations of Gas Discharge* (Nauka, Moscow, 1980).
6. Yu. P. Raĭzer, Teplofiz. Vys. Temp. **24**, 984 (1986).
7. A. Turing, Philos. Trans. R. Soc. London, Ser. B **237**, 37 (1952).
8. V. A. Granovskiĭ, *Electric Current in a Gas. Steady Current* (Nauka, Moscow, 1971).
9. Ph. I. Vysikaĭlo, in *Proceedings of International Conference: the VI Zababakhin Scientific Readings* (RFYaTs–VNIITF, Snezhinsk, Chelyabinsk oblast, 2001), p. 31.
10. Ph. I. Vysikaĭlo, in *Abstracts of X Scientific Seminar: the Earth Planet System* (Mosk. Gos. Univ., Moscow, 2002), p. 268.
11. Ph. I. Vysikaĭlo, I. V. Sharov, *et al.*, in *Abstracts of XI Conference on Physics of Gas Discharge* (RGRTA, Ryazan, 2002), Part 1, p. 14.
12. M. I. Rabinovich and A. B. Ezerskiĭ, *Dynamical Theory of Forming* (Yanus-K, Moscow, 1998).
13. L. B. Loed, *Fundamental Processes of Electrical Discharges in Gases* (Wiley, New York, 1939; Gostekhizdat, Moscow, 1950).
14. Ph. I. Vysikaĭlo, Fiz. Plazmy **11**, 1256 (1985) [Sov. J. Plasma Phys. **11**, 720 (1985)].
15. V. N. Babichev, Ph. I. Vysikaĭlo, and S. A. Golubev, Pis'ma Zh. Tekh. Fiz. **12**, 992 (1986) [Sov. Tech. Phys. Lett. **12**, 409 (1986)].
16. V. N. Babichev, Ph. I. Vysikaĭlo, S. A. Golubev, *et al.*, Dokl. Akad. Nauk SSSR **297**, 833 (1987) [Sov. Phys. Dokl. **32**, 979 (1987)].
17. Ph. I. Vysikaĭlo, Preprint No. 0025-A, TRINITI (1996); Preprint No. 0051-A, TRINITI (1999); Preprint No. 0074-A, TRINITI (2000); Preprint No. 0090-A, TRINITI (2002).
18. Ph. I. Vysikaĭlo, Fiz. Plazmy (Moscow) **16**, 1268 (1990) [Sov. J. Plasma Phys. **16**, 731 (1990)].
19. A. V. Dem'yanov, D. A. Mazalov, A. P. Napartovich, *et al.*, Zh. Éksp. Teor. Fiz. **110**, 1266 (1996) [JETP **83**, 697 (1996)].
20. Ph. I. Vysikaĭlo, Fiz. Plazmy (Moscow) **13**, 216 (1987) [Sov. J. Plasma Phys. **13**, 122 (1987)].
21. A. V. Gurevich and E. E. Tsedilina, Usp. Fiz. Nauk **91**, 609 (1967) [Sov. Phys. Usp. **10**, 214 (1967)].
22. Yu. S. Akishev, Ph. I. Vysikaĭlo, A. P. Napartovich, and V. V. Ponomarenko, Teplofiz. Vys. Temp. **18**, 266 (1980).
23. A. B. Medvinskiĭ, S. V. Petrovskiĭ, I. A. Tikhonova, *et al.*, Usp. Fiz. Nauk **172**, 31 (2002) [Phys. Usp. **45**, 27 (2002)].
24. E. I. Zababakhin and I. E. Zababakhin, *Inorganic Cumulation Phenomena* (Nauka, Moscow, 1988).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 945–952.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1082–1091.

PLASMA, GASES

Anomalous Transport Regimes in a Stochastic Advection–Diffusion Model

I. L. Dranikov, P. S. Kondratenko, and L. V. Matveev
Nuclear Safety Institute, Russian Academy of Sciences, Moscow, 113191 Russia
e-mail: kondrat@ibrae.ac.ru
Received November 3, 2003

Abstract—A general solution to the stochastic advection–diffusion problem is obtained for a fractal medium with long-range correlated spatial fluctuations. A particular transport regime is determined by two basic parameters: the exponent $2h$ of power-law decay of the two-point velocity correlation function and the mean advection velocity u. The values of these parameters corresponding to anomalous diffusion are determined, and anomalous behavior of the tracer distribution is analyzed for various combinations of u and h. The tracer concentration is shown to decrease exponentially at large distances, whereas power-law decay is predicted by fractional differential equations. Equations that describe the essential characteristics of the solution are written in terms of coupled space-time fractional differential operators. The analysis relies on a diagrammatic technique and makes use of scale-invariant properties of the medium. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In many situations, transport in highly disordered media cannot be described by the classical diffusion equation [1]. In particular, this is the case with fractal media characterized by long-range spatial correlations of fluctuating characteristics. One approach to problems of this kind is based on the use of fractional differential equations (e.g., see [2–4]). They admit solutions characterized by anomalous time dependence of the tracer-cloud radius R at long times ($R \propto t^\gamma$ with $\gamma \neq 1/2$) and power-law (instead of Gaussian) decay of tracer concentration at long distances (at $r \gg R$). "Heavy" (power-law) tails in tracer distributions are extremely important for applications (e.g., for substantiating the reliability of radioactive waste disposal sites) in view of the huge difference between power-law and Gaussian decay. However, the standard fractional-diffusion approach relies on a formal analysis and requires both further mathematical substantiation (e.g., see [5, 6]) and generalization (see [6–12]). Therefore, the ensuing conclusions must be tested against specific physical models.

The stochastic advection–diffusion model with a slow (power-law) long-distance decay of correlations is one such model. The results of its analysis performed in [12, 13] under simplifying assumptions are consistent with those obtained by solving fractional-diffusion equations. (A review of the relationship between these approaches and relevant bibliography can be found in [14]). However, one may question the sensitivity of these results to the approximations employed and their agreement with the exact solution to the advection–diffusion problem.

In this paper, this solution is obtained in a general form by analyzing a group of scaling transformations [15] with the use of the Feynman diagrammatic technique [16], which was applied to transport in random media in [17–19].

In Section 2, the problem is stated and a diagrammatic representation of the Green function is constructed. In Section 3, the behavior of the tracer concentration at long and short distances is analyzed in the case of zero mean advection velocity. In Section 4, we analyze the case when this velocity is finite. The main results are summarized and discussed in the final section.

2. STATEMENT OF THE PROBLEM

The model is based on an equation describing advection and molecular diffusion of concentration $c(\mathbf{r}, t)$:

$$\frac{\partial c}{\partial t} + \nabla(\mathbf{v} - D\nabla)c = 0, \qquad (1)$$

where both advection velocity $\mathbf{v} = \mathbf{v}(\mathbf{r})$ and diffusivity $D(\mathbf{r})$ are random functions of coordinates. Hereinafter, we consider the problem with a tracer distribution prescribed at the initial moment $t = 0$ in the absence of sources (sources can readily be taken into account).

The medium is statistically homogeneous and isotropic, and both $\mathbf{v}$ and D are represented as

$$\mathbf{v}(\mathbf{r}) = \mathbf{u} + \mathbf{v}'(\mathbf{r}), \quad D(\mathbf{r}) = \bar{D} + D'(\mathbf{r}), \qquad (2)$$

where $\mathbf{u} = \langle \mathbf{v}(\mathbf{r}) \rangle$ and $\bar{D} = \langle D(\mathbf{r}) \rangle$ are ensemble-averaged quantities independent of coordinates, and $\mathbf{v}'(\mathbf{r})$ and $\mathbf{D}'(\mathbf{r})$ are the fluctuating velocity and diffusivity, respectively ($\langle \mathbf{v}'(\mathbf{r}) \rangle = 0$, $\langle D'(\mathbf{r}) \rangle = 0$).

1063-7761/04/9805-0945$26.00

The velocity field satisfies the incompressibility condition:

$$\mathrm{div}\mathbf{v} = \mathrm{div}\mathbf{v}' = 0. \quad (3)$$

Since a fractal medium is considered, fluctuations of its characteristics are long-range correlated; i.e., the long-distance decay of the correlation functions of the random $\mathbf{v}'$ and D' follows power laws. In particular, the two-point velocity correlation function,

$$K_{ij}^{(2)}(\mathbf{r}_1 - \mathbf{r}_2) = \langle v_i'(\mathbf{r}_1) v_j'(\mathbf{r}_2)\rangle, \quad (4)$$

behaves as follows when $|\mathbf{r}_1 - \mathbf{r}_2| \gg a$:

$$K_{ii}^{(2)}(\mathbf{r}_1 - \mathbf{r}_2) \propto V^2\left(\frac{a}{|\mathbf{r}_1 - \mathbf{r}_2|}\right)^{2h}, \quad (5)$$

where $h > 0$, a is the short-range cutoff radius, and V^2 is the characteristic value of $K_{ij}^{(2)}(\mathbf{r})$ for $|\mathbf{r}| \lesssim a$. Thus, $K_{ij}^{(2)}(\mathbf{r})$ is a homogeneous function of degree $-2h$ for $|\mathbf{r}| \gg a$. Similarly, the n-point velocity correlation function,

$$K_{i_1 i_2 \ldots i_n}^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) = \langle v_{i_1}'(\mathbf{r}_1) v_{i_2}'(\mathbf{r}_2) \ldots v_{i_n}'(\mathbf{r}_n)\rangle, \quad (6)$$

satisfies the following relation when $|\mathbf{r}_i - \mathbf{r}_j| \gg a$ for any pair of $\mathbf{r}_i$ and $\mathbf{r}_j$:

$$K_{i_1 i_2 \ldots i_n}^{(n)}(\lambda\mathbf{r}_1, \lambda\mathbf{r}_2, \ldots, \lambda\mathbf{r}_n) = \lambda^{-nh} K_{i_1 i_2 \ldots i_n}^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n). \quad (7)$$

By analogy with the theory of critical phenomena [15], the parameter h is called the scaling dimension of velocity fluctuation $\mathbf{v}'$. According to (7), the Fourier transform of the n-point correlation function,

$$\begin{aligned} K_{i_1 i_2 \ldots i_n}^{(n)}\{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n\} &= \int d\mathbf{r}_1 d\mathbf{r}_2 \ldots d\mathbf{r}_n \\ &\times \exp[-i(\mathbf{k}_1 \cdot \mathbf{r}_1 + \mathbf{k}_2 \cdot \mathbf{r}_2 + \ldots + \mathbf{k}_n \cdot \mathbf{r}_n)] \\ &\times K_{i_1 i_2 \ldots i_n}^{(n)}(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n) \\ &\equiv (2\pi)^3 \delta(\mathbf{k}_1 + \mathbf{k}_2 + \ldots + \mathbf{k}_n) \\ &\times \tilde{K}_{i_1 i_2 \ldots i_n}^{(n)}(\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{n-1}), \end{aligned} \quad (8)$$

satisfies the scaling relation

$$K_{i_1 i_2 \ldots i_n}^{(n)}\{\lambda\mathbf{k}_1, \lambda\mathbf{k}_2, \ldots, \lambda\mathbf{k}_n\} = \lambda^{n(h-3)} K_{i_1 i_2 \ldots i_n}^{(n)}\{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n\}. \quad (9)$$

The power exponent $n(h - 3)$ on the right-hand side of (9) is called the scaling exponent of $K_{i_1 i_2 \ldots i_n}^{(n)}\{\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_n\}$.

Combining (5) with (9), we find that the Fourier transform of the two-point correlation function behaves as follows when $ka \ll 1$ ($k \equiv |\mathbf{k}|$):

$$\tilde{K}_{ii}^{(2)}(\mathbf{k}) \sim V^2 a^{2h} k^{2h-3}. \quad (10)$$

Relations analogous to (4)–(10) can also be written out for the correlation functions of diffusivity fluctuations.

The concentration satisfying Eq. (1) at an arbitrary instant can be expressed in terms of its initial distribution as

$$c(\mathbf{r}, t) = \int d\mathbf{r}' G(\mathbf{r}, \mathbf{r}'; t) c(\mathbf{r}', 0), \quad (11)$$

where the Green function $G(\mathbf{r}, \mathbf{r}'; t)$ solves the equation

$$\left\{\frac{\partial}{\partial t} + \frac{\partial}{\partial x_i} v_i(\mathbf{r}) - \frac{\partial}{\partial x_i} D \frac{\partial}{\partial x_i}\right\} G(\mathbf{r}, \mathbf{r}'; t) = 0 \quad (12)$$

subject to the initial condition

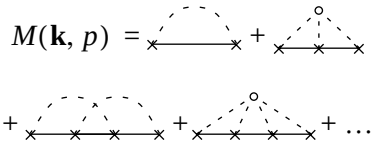
$$G(\mathbf{r}, \mathbf{r}'; 0) = \delta(\mathbf{r} - \mathbf{r}'). \quad (13)$$

Of practical interest is the tracer concentration averaged over the ensemble of realizations of the medium, $\bar{c}(\mathbf{r}, t) \equiv \langle c(\mathbf{r}, t)\rangle$. It satisfies the equation obtained from (11) by replacing c with $\bar{c}$ and G with $\bar{G}$, where $\bar{G}(\mathbf{r} - \mathbf{r}', t) \equiv \langle G(\mathbf{r}, \mathbf{r}'; t)\rangle$ is the ensemble-averaged Green function (henceforth called Green function for simplicity). The calculation of $\bar{G}(\mathbf{r} - \mathbf{r}', t)$ is facilitated by invoking the "cross" diagrammatic technique developed in [17] and applied to transport theory for disordered media in [18, 19].

Performing Fourier and Laplace transforms in space and time, respectively, and using Eq. (12), initial condition (13), and expressions (2), we obtain

$$\bar{G}\{\mathbf{k}, p\} = \frac{1}{p + i\mathbf{k} \cdot \mathbf{u} + \bar{D}k^2 - M(\mathbf{k}, p)}, \quad (14)$$

where $\mathbf{k}$ and p are the Fourier and Laplace variables, respectively, and $M(\mathbf{k}, p)$ is the "self-energy" operator expressed as the sum of irreducible skeleton diagrams [16]:

$M(\mathbf{k}, p) =$  (15)

Here, horizontal lines and crosses represent, respectively, functions $\bar{G}$ and perturbation operators of the form

$$\hat{T} = -v_i'(\mathbf{r})\frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_i} D'(\mathbf{r}) \frac{\partial}{\partial x_i}. \quad (16)$$

Each dashed line joins the crosses relating to a particular cumulant. (Expansion in terms of cumulants plays a

key role in ensemble averaging.) By definition (16), each cross contains gradients, as well as velocity and diffusivity fluctuations. In the Fourier representation, gradients are replaced by wave vectors. In particular, the D' term contains the product of the wave vectors relating to respective G lines adjoining the cross on its right and left, while the $\mathbf{v}'$ term contains a single wave vector associated with either line by virtue of (3). Each dashed line emanating from a cross is associated with a wave vector over which integration is performed. Substituting (14) into diagrammatic expansion (15), one obtains an integral equation for $M(\mathbf{k}, p)$.

The sections that follow present an analysis of its solutions.

3. TRANSPORT WITH ZERO MEAN ADVECTION

First, we analyze the role played by diffusivity fluctuations. The analysis is performed here for the first diagram in self-energy operator (15). By virtue of (16), it can be represented as the sum

$$M_2 = M_2^{(\mathrm{A})} + M_2^{(\mathrm{D})}.$$

The summands represent the contributions due to advection-velocity and diffusivity fluctuations. Up to constant factors, they are expressed as

$$M_2^{(\mathrm{A})} \sim k^2 \int d\mathbf{q} \frac{|\mathbf{k}-\mathbf{q}|^{2h-3}}{p + \bar{D}q^2 - M(\mathbf{q}, p)}, \tag{17}$$

$$M_2^{(\mathrm{D})} \sim k^2 \int d\mathbf{q} \frac{q^2|\mathbf{k}-\mathbf{q}|^{2g-3}}{p + \bar{D}q^2 - M(\mathbf{q}, p)}. \tag{18}$$

Here, two-point velocity and diffusivity correlation functions are replaced with their respective asymptotics corresponding to small wave vectors. Since g in (18) is the scaling dimension of diffusivity, it holds that $g > 0$. If $M(\mathbf{q}, p) \sim q^2$ for small wave vectors, then the integral in (18) is convergent as $\mathbf{q}$ goes to zero even if $\mathbf{k} = 0$ and $p = 0$. The apparent divergence of the integral for large wave vectors is explained by the fact that the long-wavelength asymptotic expression for the two-point correlation function used in the integrand fails in this limit. This means that the integral is constant up to small corrections under the condition k ($k \ll a^{-1}$) of interest here, and (18) implies that $M_2^{(\mathrm{D})}(\mathbf{k}, p) \sim k^2$. The term $M_2^{(\mathrm{D})}(\mathbf{k}, p) \sim k^2$ in expression (14) for the Green function can be combined with $\bar{D}k^2$ to obtain a renormalized mean diffusivity. This can also be done when all higher order diagrams are taken into account. For this reason, we retain only the contributions due to advection-velocity fluctuations to the expansion of the self-energy operator (i.e., only the first term in (16)).

Similarly, expression (17) implies that $M_2^{(\mathrm{A})}(\mathbf{k}, p) \sim k^2$ for small wave vectors when $h > 1$. Since this dependence holds in all higher order diagrams, the stochastic diffusion model can also be reduced to normal diffusion in this case.

Thus, velocity fluctuations with scaling dimension $h > 1$, as well as diffusivity fluctuations with any dimension, are irrelevant with respect to deviation from normal diffusion. When $h \le 1$, an essentially different behavior is observed.

The cases when $h < 1$ and $h = 1$ are analyzed separately below.

3.1. $h < 1$

If $M(\mathbf{q}, p) \sim q^2$ in the integrand of (17) as $p \longrightarrow 0$, then the integral is divergent at the lower wave-vector limit as $\mathbf{k}, p \longrightarrow 0$. Since it is convergent at the higher wave-vector limit, the use of long-wavelength asymptotics of the velocity correlation function is justified, and the divergence indicated above implies that $M(\mathbf{q}, p) \gg \bar{D}k^2$ as $\mathbf{k}, p \longrightarrow 0$. Therefore, further analysis of the case when $\mathbf{u} = 0$ (see (14)) can be performed for

$$\bar{G}\{\mathbf{k}, p\} = \frac{1}{p - M(\mathbf{k}, p)}. \tag{19}$$

Equation (15) is the diagrammatic expansion of the self-energy operator in terms of correlation functions that are scale-invariant by virtue of (9). Accordingly, we may assume (and subsequently prove) that both self-energy operator and Green function have analogous properties. Expression (19) implies that the scaling exponents of the self-energy operator and the Laplace variable p are equal. Therefore, the scaling relations for $M(\mathbf{k}, p)$ and $\bar{G}\{\mathbf{k}, p\}$ must have the form

$$M(\lambda\mathbf{k}, \lambda^{\Delta}p) = \lambda^{\Delta}M(\mathbf{k}, p), \tag{20}$$

$$\bar{G}(\lambda\mathbf{k}, \lambda^{\Delta}p) = \lambda^{-\Delta}\bar{G}(\mathbf{k}, p), \tag{21}$$

where Δ is a scaling exponent to be determined.

To prove (20) and (21), we use diagrammatic representation (15) to write $M(\mathbf{k}, p)$ as a sum in which the nth summand is the contribution of all diagrams containing n crosses:

$$M(\mathbf{k}, p) = k^2 L(\mathbf{k}, p) = k^2 \sum_{n=2}^{\infty} L_n(\mathbf{k}, p). \tag{22}$$

The factor k^2 corresponds to the gradients in the extreme right and left crosses. The scaling exponent of the nth term in (22), Δ_n, is the sum of the exponents of the elements of the corresponding diagram. These elements include a combination of n-point velocity correlators, n gradients, $n - 1$ Green functions, and the $3n$-dimensional differential of wave vectors. The corre-

sponding exponents are $n(h-3)$, n, $-(n-1)\Delta$, and $3n$, respectively. (Since the wave vectors used as integration variables are combined additively with $\mathbf{k}$, the corresponding scaling exponents are equal to that of $\mathbf{k}$.) Equating the sum of the exponents enumerated above to the exponent Δ of the operator M, we obtain

$$nh + n + (n-1)\Delta = \Delta, \tag{23}$$

which yields a self-energy scaling exponent independent of the order of a diagram:

$$\Delta = 1 + h. \tag{24}$$

Since (20) with $\Delta = 1 + h$ is thus valid for each summand in (22), it holds for the series as a whole. By virtue of (24), relation (20) entails the representation

$$M(\mathbf{k}, p) = Va^h k^{1+h} \varphi(\xi), \tag{25}$$

where $\varphi(\xi)$ is a dimensionless function of the dimensionless scaling variable

$$\xi = \frac{p}{Va^h k^{1+h}}. \tag{26}$$

The factor Va^h in (25) and (26) is obtained by using (5).

Next, we analyze the behavior of $\varphi(\xi)$ at large and small values of its argument. It is obvious that $\varphi(\xi)$ must approach a finite limit as $\xi \longrightarrow 0$. Indeed, if $\lim_{\xi \to 0} \varphi(\xi) = \infty$, then (25), (26), and (19) would imply that $M(\mathbf{k}, 0) = \infty$ and $\bar{G}\{\mathbf{k}, 0\} = 0$. However, diagrammatic expansion (15) would then yield $M(\mathbf{k}, 0) = 0$, which contradicts the assumption that $\lim_{\xi \to 0} \varphi(\xi) = \infty$. It can be proved in a similar manner that $\lim_{\xi \to 0} \varphi(\xi) \neq 0$. Thus, we have $\lim_{\xi \to 0} \varphi(\xi) = A \sim 1$ and

$$M(\mathbf{k}, 0) = AVa^h k^{1+h}. \tag{27}$$

Comparing this result with (22), we find that $L_n(\mathbf{k}, 0) \propto k^{-(1-h)}$. The ensuing divergence of the integrals in $L_n(\mathbf{k}, p)$ for $p = 0$ as $\mathbf{k} \longrightarrow 0$ implies that they are convergent at $q_i \ll a^{-1}$ for finite $\mathbf{k}$ and p; i.e., the asymptotic expressions for velocity correlation functions corresponding to this limit are valid. Expression (27) obtained for $\xi = 0$ can be extended to nonzero p such that $\xi \ll 1$. The next term in the expansion of $M(\mathbf{k}, 0)$ in terms of ξ either is p up to a constant factor if $h < 1/2$ (in which case the scaling of the denominator in (19) does not change) or scales with $k^{2h-1}p^{(2-h)/(1+h)}$ if $h > 1/2$, prevailing over p in (19) in order of magnitude. (Recall that p corresponds to the first time derivative in the fractional-diffusion approach.)

In the opposite limit of $\xi \gg 1$ (k^{1+h} goes to zero faster than p), the first approximation of (22) with respect to $\mathbf{k}$ can be used:

$$M(\mathbf{k}, p) \approx k^2 L(0, p). \tag{28}$$

Here, the function $L(0, p)$ can be found by combining scaling relation (20) with (24):

$$L(0, p) \propto p^{-(1-h)/(1+h)}. \tag{29}$$

In this case, the self-energy operator behaves asymptotically as $M \propto k^2 p^{-(1-h)/(1+h)}$ at $p \gg Va^h k^{1+h}$. (However, this clearly does not entail any singularity of M as $p \longrightarrow 0$.) By virtue of (25) and (26), it follows that the first term in the expansion of $\varphi(\xi)$ is

$$\varphi(\xi) \sim \xi^{-(1-h)/(1+h)} \text{ at } \xi \gg 1. \tag{30}$$

Since $\mathbf{k} = 0$ is not a singular point of the integrals in $L_n(\mathbf{k}, p)$ when $p \neq 0$, we have an analytic expansion in integer powers of $\mathbf{k}^2$ as $\mathbf{k} \longrightarrow 0$, and relations (20) and (24) imply that the function

$$f(\xi) \equiv \frac{\varphi(\xi)}{\xi} \equiv \frac{M(\mathbf{k}, p)}{p} \tag{31}$$

can be represented as

$$f(\xi) = \sum_{n=1}^{\infty} b_n \xi^{2n/(1+h)}, \quad \xi \gg 1, \tag{32}$$

where b_n are dimensionless coefficients. We emphasize that (32) corresponds to an analytic expansion of $M(\mathbf{k}, p)$ in terms of the vector $\mathbf{k}$ that holds at $\xi \gg 1$ (see (26)). Comparing this result with expression (27) for $M(\mathbf{k}, p)$ at large k, we conclude that $M(\mathbf{k}, p)$ has a branch point at

$$k \sim \left(\frac{p}{Va^h}\right)^{1/(1+h)}$$

(which corresponds to a branch point of $f(\xi)$ at $|\xi| \sim 1$). In other words, the scaling of propagator (19) can be essentially different, depending on the ratio of p and k^{1+h} even in the Fourier–Laplace domain where both variables are small (this behavior is analogous to the behavior of vertices in the Landau theory of Fermi liquids [20]). To the best of our knowledge, this behavior of a completely scale-invariant propagator is not predicted by any model based on the standard fractional-diffusion equations [2, 3], but is possible in generalizations involving coupled space-time fractional-differential operators (as in [9, 11]).

In the present model, one can easily construct interpolation operators having the asymptotic properties of the propagator considered here. For example, we can use $\partial_t - \mathfrak{D}^{(1+h)/2}$ if $h < 1/2$ and $\partial_t^{(2-h)/(1+h)} \mathfrak{D}^{h-1/2} - \mathfrak{D}^{(1+h)/2}$ if $h > 1/2$, where ∂_t denotes a time derivative and $\mathfrak{D}$ is a linear combination of $\partial_t^{2/(1+h)}$ and a Laplace operator multiplied by a factor of dimension $(Va^h)^{2/(1+h)}$.

Next, we analyze the behavior of the Green function in a space-time representation. The inverse Fourier–Laplace transform of (19) yields

$$\bar{G}(\mathbf{r}, t) = \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{r}} I(Va^h k^{1+h} t), \quad (33)$$

where

$$I(s) = \frac{1}{2\pi i} \int_{\xi_0 - i\infty}^{\xi_0 + i\infty} \frac{d\xi}{\xi} \frac{e^{is\xi}}{1 + f(\xi)}, \quad \xi_0 > 0. \quad (34)$$

Combining (33) and (34), we obtain a general expression for the Green function:

$$\bar{G}(\mathbf{r}, t) = (Va^h t)^{-3/(1+h)} \Phi\left(\frac{r^{1+h}}{Va^h t}\right), \quad (35)$$

where the argument of $\Phi(x)$ scales with unity and $\Phi(0) \sim 1$. In particular, this implies that the tracer cloud radius behaves as

$$R \sim (Va^h t)^{1/(1+h)} \quad (36)$$

at long times. At short distances from the origin such that $r \ll (Va^h t)^{1/(1+h)}$, the exponential in (33) can be ignored to obtain

$$\bar{G}(\mathbf{r}, t) \sim (Va^h t)^{-3/(1+h)}, \quad (37)$$

which can also be derived from (36) by using the conservation of the total number of tracer particles. Expression (37) estimates the long-time decay of concentration at a fixed point in space.

Let us now elucidate the asymptotic form of the Green function at distances much greater than R, i.e., at $r \gg (Va^h t)^{1/(1+h)}$. According to (33), the possibility (or impossibility) of power-law decay of $\bar{G}(\mathbf{r}, t)$ is determined by the behavior of $I(s)$ at $s \ll 1$. By substituting (32) into (34), the function $I(s)$ is rewritten as

$$I(s) = \sum_{n=0}^{\infty} d_n s^{2n/(1+h)} \equiv \sum_{n=0}^{\infty} d_n [(Va^h t)^{2/(1+h)} k^2]^n. \quad (38)$$

Substituting this formula into (33), we express $\bar{G}$ as a sum in which the first term is $\delta(\mathbf{r})$ and the nth term is proportional to the result obtained by applying the Laplace operator to $\delta(\mathbf{r})$ n times. In other words, each summand is zero at $\mathbf{r} \neq 0$. Mathematically, this implies that $r^{1+h}/Va^h t = \infty$ is an essentially singular point of the function $\Phi(x)$ and the Green function cannot be represented as a series expansion in inverse powers of $r^{1+h}/Va^h t$.

The asymptotic form of $\bar{G}$ at $r \gg (Va^h t)^{1/(1+h)}$ can be found by shifting the contour of integration with respect to k in (33) from the real axis into the upper half-plane (after performing the integral over the angular variables). In the limit considered here, this yields

$$\bar{G}(\mathbf{r}, t) \propto \exp(ik_s r), \quad (39)$$

where k_s is the singular point of $I(Va^h k^{1+h} t)$ in the upper half-plane of k that is nearest to the real axis. Since the function $\bar{G}$ associated with concentration by definition (11)–(13) can only be positive, the exponent in (39) must be a real quantity. Therefore, k_s lies on the imaginary axis and $\mathrm{Im}\, k_s \sim (Va^h t)^{-1/(1+h)}$ in accordance with expansion (38). As a result, we have an asymptotic expression for the Green function at large r:

$$\bar{G}(\mathbf{r}, t) \propto \exp\left[-\frac{Br}{(Va^h t)^{1/(1+h)}}\right], \quad B \sim 1. \quad (40)$$

Combined with the normalization factor $(Va^h t)^{-3/(1+h)}$ in (35), this formula determines the behavior of concentration at $r \gg (Va^h t)^{1/(1+h)}$, i.e., its growth at a fixed distant point and its instantaneous profile. Thus, the spatial tracer distribution predicted by the stochastic advection model does not have a "heavy" (power-law) tail even if the power-law decay of the velocity correlator is slow (with $h < 1$).

3.2. $h = 1$

When $h = 1$, expression (17) with $M(\mathbf{q}, p) \sim q^2$ substituted into the integral yields a logarithmically divergent quantity: $M_2^{(A)} \sim Vak^2 \ln\mu$, where

$$\mu = \left(\max\left\{\frac{pa}{V}, (ka)^2\right\}\right)^{-1}. \quad (41)$$

This suggests that a logarithmic approximation of the self-energy operator should be sought in the form

$$M(\mathbf{k}, p) \sim Vak^2 \ln^{\alpha}\mu \quad (42)$$

with $\alpha > 0$. The integral over the wave vectors in an arbitrary diagram of order $n > 2$ also yields a logarithm. However, such a diagram contains the product of $n - 2$ additional functions $\bar{G}$, as compared to the second-order one. According to (42) and (19), this results in a small factor $\ln^{-\alpha(n-2)}\mu$. Therefore, calculation of the self-energy operator can be restricted to the skeleton diagram with two crosses, and Eq. (17) reduces to

$$M(\mathbf{k}, p) \sim k^2 V^2 a^2 \int d\mathbf{q} \frac{|\mathbf{k} - \mathbf{q}|^{-1}}{p - M(\mathbf{q}, p)}. \quad (43)$$

Combined with (42), this yields $\alpha = 1 - \alpha$, i.e., $\alpha = 1/2$. Thus, when the velocity scaling dimension is $h = 1$, the Green function is

$$\bar{G}\{\mathbf{k}, p\} = \frac{1}{p + D_{\mathrm{ef}} k^2 \ln^{1/2}\mu}, \quad (44)$$

where the argument of the logarithm is given by (41) and $D_{ef} \sim Va$. Following the derivation of (36), we find that the tracer cloud radius behaves as

$$R \sim (Vat)^{1/2} \ln^{1/4}\left(\frac{Vt}{a}\right) \tag{45}$$

at long times and conclude that the long-distance expansion of the tracer concentration does not contain any power-law terms.

4. TRANSPORT WITH NONZERO MEAN ADVECTION ($u \neq 0$)

Advection by mean flow is a fortiori much more effective than transport by velocity fluctuations. For this reason, we consider the case when

$$u \ll V. \tag{46}$$

We introduce the new variable

$$p' = p + i\mathbf{k} \cdot \mathbf{u} \tag{47}$$

instead of p for convenience, i.e., use a reference frame moving with a velocity $\mathbf{u}$: $\mathbf{r} \longrightarrow \mathbf{r}' \equiv \mathbf{r} - \mathbf{u}t$. Green function (14) is expressed in terms of p' as follows:

$$\tilde{\tilde{G}}\{\mathbf{k}, p'\} = \frac{1}{p' - \tilde{M}(k, p')}. \tag{48}$$

In the general case, a nonzero mean advection velocity violates the scale invariance expressed by (20) and (21). However, the problem can be analyzed in two important limit cases.

If $\mathbf{k}$ and p' are such that

$$\max\{|p'|, |\tilde{M}(\mathbf{k}, p')|\} \gg uk, \tag{49}$$

then the term $-i\mathbf{q} \cdot \mathbf{u}$ can be neglected in the denominator of a Green function having the form

$$\tilde{\tilde{G}}\{\mathbf{k} - \mathbf{q}, p'\} = [p' - i\mathbf{q} \cdot \mathbf{u} - \tilde{M}(\mathbf{k}, p')]^{-1}.$$

Accordingly, the results obtained above remain valid for $\mathbf{u} \neq 0$ irrespective of the scaling dimension h. Using expression (27) as an estimate for $\tilde{M}(\mathbf{k}, p')$, we can write

$$\max\left\{k, \frac{|p'|}{u}\right\} \gg k_* \tag{50}$$

instead of (49), where

$$k_* \sim \frac{1}{a}\left(\frac{u}{V}\right)^{1/h}. \tag{51}$$

This result means that tracer propagation at $t \ll t_* \equiv (uk_*)^{-1}$ obeys the laws of advection with zero mean velocity. In this case, the tracer cloud radius is greater than the corresponding displacement by the mean flow ($R \gg ut$).

When $\max\{k, |p'|/u\} \ll k_*$, the behavior of the self-energy operator is essentially different for $h > 1/2$, $h < 1/2$, and $h = 1/2$. We analyze each of these cases separately.

Case 1: $\max\{k, |p'|/u\} \ll k_*$, $h > 1/2$.

The integrals in the expressions for $\tilde{M}(\mathbf{k}, p')$ are dominated by the contributions from $|\mathbf{q}_i| > k_*$. Therefore, the quantity $\tilde{L}(\mathbf{k}, p')$ analogous to $L(\mathbf{k}, p')$ in (22) is independent of $\mathbf{k}$ and p' and $\tilde{L} \sim Va^h k_*^{h-1} \sim u^2 t_*$. Accordingly,

$$\tilde{M} \sim u^2 t_* k^2. \tag{52}$$

Thus, advection at $t \gg t_*$ obeys the laws of normal diffusion when $h > 1/2$.

Case 2: $\max\{k, |p'|/u\} \ll k_*$, $h < 1/2$.

An analysis of the diagram with two crosses shows that the integrals over wave vectors are convergent when $|\mathbf{q}_i| \ll k$. Therefore, Green functions of the form

$$\tilde{\tilde{G}}\{\mathbf{k} - \mathbf{q}, p'\} = [p' - i\mathbf{q} \cdot \mathbf{u} - \tilde{M}(\mathbf{k} - \mathbf{q}, p')]^{-1}$$

can be approximated as follows:

$$\tilde{\tilde{G}}\{\mathbf{k} - \mathbf{q}, p'\} \approx [p' - i\mathbf{q} \cdot \mathbf{u} - \tilde{M}(\mathbf{k}, p')]^{-1}. \tag{53}$$

Similarly, the factors $\mathbf{k} - \mathbf{q}$ corresponding to the gradients in (16) can be simplified:

$$\mathbf{k} - \mathbf{q} \longrightarrow \mathbf{k}. \tag{54}$$

As a result, the scale invariance expressed by (20) and (21) is restored, but the scaling exponent Δ is obviously different. To calculate Δ, we must take into account one distinction from the case of $\mathbf{u} = 0$. In view of (53) and (54), the wave vectors used as integration variables should be associated with the scaling exponent Δ. Thus, the scaling exponent Δ of the nth-order diagram in $\tilde{M}(\mathbf{k}, p')$ is the sum of n, $n(h-3)$, $n(h-3)\Delta$, $-(n-1)\Delta$, and $3n\Delta$, which correspond to the n factors $\mathbf{k}$ due to the gradients in crosses, a combination of n-point velocity correlators, the product of $n-1$ Green functions, and the product of differentials, respectively. As a result, we have the equation

$$n + nh\Delta - (n-1)\Delta = \Delta, \tag{55}$$

which yields

$$\Delta = \frac{1}{1-h}. \tag{56}$$

Relations (20) and (56) can be used to represent the self-energy operator, by analogy with (25), as

$$\tilde{M} \sim \frac{1}{t_*}(ut_*k)^{1/(1-h)}\psi(\zeta), \tag{57}$$

where $\psi(\zeta)$ is a dimensionless function of the dimen-

sionless variable

$$\zeta = p' t_* (u t_* k)^{-1/(1-h)}. \quad (58)$$

In the limit case of $\zeta \ll 1$, when $\tilde{M} \gg p'$, we should drop p' in denominator (53). Then, $\tilde{M}(\mathbf{k}, p')$ becomes independent of p' and can be determined up to a numerical factor of order unity by combining (57) with (58):

$$\tilde{M} \sim t_*^{-1} (u t_* k)^{1/(1-h)}. \quad (59)$$

In the opposite limit of $\zeta \gg 1$, when $\tilde{M} \ll p'$, the leading-order contribution to the self-energy operator is due to the diagram with two crosses, where the Green function can be approximated as follows: $\bar{G}(\mathbf{k} - \mathbf{q}, p') \approx (p' - i\mathbf{q} \cdot \mathbf{u})^{-1}$. Calculating this diagram, we obtain

$$\tilde{M} \sim \frac{k^2 u^2}{p'} (p' t_*)^{2h}. \quad (60)$$

The structure of expressions (57)–(60) is analogous to that of (25)–(29). Following the derivation of (35), (36), and (40), we use (47), (57), and (58) to calculate the inverse Fourier–Laplace transform of (48). As a result, we obtain an expression for the Green function that is valid at $t \gg t_*$ when $h < 1/2$:

$$\bar{G}(\mathbf{r}, t) = \frac{1}{(u t_*)^3} \left(\frac{t_*}{t}\right)^{3(1-h)} \Psi\left[\frac{t_*}{t}\left(\frac{r'}{u t_*}\right)^{1/(1-h)}\right], \quad (61)$$

where the argument of $\Psi(x)$ scales with unity and $\Psi(0) \sim 1$. This implies that the tracer cloud radius behaves as

$$R \sim u t_* \left(\frac{t}{t_*}\right)^{1-h} \quad (62)$$

at long times. Within the cloud (at $r' \ll u t_* (t/t_*)^{1-h}$),

$$\bar{G}(\mathbf{r}, t) \sim \left[\frac{1}{u t_*}\left(\frac{t_*}{t}\right)^{1-h}\right]^3. \quad (63)$$

To elucidate the behavior of the Green function at longer distances, we follow the analysis of the case of $\mathbf{u} = 0$. Since the singular point of the Fourier transform of $\bar{G}$ as a function of k nearest to the real axis is

$$k_s \sim i \frac{1}{u t_*}\left(\frac{t_*}{t}\right)^{1-h},$$

the asymptotic form of the Green function at large r' is

$$G(\mathbf{r}, t) \propto \exp\left[-C \frac{r'}{u t_*}\left(\frac{t_*}{t}\right)^{1-h}\right], \quad C \sim 1. \quad (64)$$

Case 3: $\max\{k, |p'|/u\} \ll k_*$, $h = 1/2$.

When $\mathbf{u} \neq 0$ and $h = 1/2$, the equations are similar to those obtained for $\mathbf{u} = 0$ and $h = 1$, except that the integrand is independent of $\tilde{M}(\mathbf{q}, p')$ in the region of logarithmic behavior. Finally, we have

$$\tilde{M} \sim u^2 t_* k^2 \ln \nu,$$

$$\nu = \left(\max\left\{\frac{p'}{u k_*}, \left(\frac{k}{k_*}\right)^2\right\}\right)^{-1}. \quad (65)$$

(Again, this result can be obtained by solving a truncated equation with a two-cross diagram.) Following the derivation of (45), we find an estimate valid at $t \gg t_*$:

$$R \sim \left[u^2 t_* t \ln\left(\frac{t_*}{t}\right)\right]^{1/2}. \quad (66)$$

5. CONCLUSIONS

We made use of a diagrammatic technique and scale invariance to obtain a number of important results concerning passive-scalar transport in long-range correlated highly disordered media without invoking any simplifying assumptions.

The predicted behavior strongly depends on the scaling dimension of spatial velocity fluctuations, which is determined by the power exponents in velocity correlation functions. Velocity fluctuations with a scaling dimension $h > 1$, as well as diffusivity fluctuations of any dimension, do not cause any deviation from normal diffusion.

For a flow with velocity correlations characterized by a dimension $h < 1$, a superdiffusive regime is predicted. In particular, when the advection velocity is zero, the tracer cloud radius increases as $R \propto t^{1/(1+h)}$, i.e., faster than in the case of normal diffusion. This scaling is qualitatively similar to that predicted by the fractional-diffusion model, where the transport equation contains a time derivative and a derivative of order $(1 + h)$ with respect to coordinates (e.g., in contrast to [21]). However, both models lead to essentially different results in the long-distance limit ($r \gg R$). Whereas the fractional-diffusion model predicts a power-law of concentration decay in this limit (e.g., see [2, 14]), our analysis shows that the concentration decreases exponentially with an exponent $\sim -(r/R)$. In other words, the stochastic advection model does not admit any power-law tail in the tracer distribution even when velocity is long-range correlated.

We believe that this discrepancy is not accidental. The physically well-motivated stochastic advection model takes into account the very low probability of high velocities, which rules out slow (power-law) decay of concentration at long distances. Both the standard fractional-diffusion model [2, 3] and its generalizations with arbitrary parameters [11] (but not the model proposed in [9]), being purely mathematical, ignore this physical observation and impose no restrictions on the

law of concentration decay at long distances. In the theory of Lévy processes, the problem analyzed here corresponds to a class of jump probabilities for which dependence on coupled space-time variables cannot be expressed in terms of delta functions.

The results obtained here for nonzero mean advection velocity ($\mathbf{u} \neq 0$) are summarized as follows. At short times ($t \ll t_*$), when anomalous diffusion prevails over advection ($R \gg ut$), transport is similar to the case of $\mathbf{u} = 0$, whereas different behavior is predicted in the long-time limit ($t \gg t_*$), when advection plays a dominant role ($R \ll ut$). In the latter case, transport in the moving reference frame depends on the relation between h and 1/2. If $h > 1/2$, then the tracer cloud spreads according to the law of normal diffusion, with an effective diffusivity depending on u. If $h < 1/2$, then a superdiffusive regime is obtained, with $R \propto t^{1-h}$.

In the transport regimes separating those described above ($h = 1$ and $h = 1/2$ when $\mathbf{u} = 0$ and $\mathbf{u} \neq 0$, respectively), the power laws for $R(t)$ are modified by logarithmic factors.

ACKNOWLEDGMENTS

We thank A.M. Dykhne for helpful remarks and S.A. Rybak for fruitful discussions.

REFERENCES

1. M. B. Isichenko, Rev. Mod. Phys. **64**, 961 (1992); S.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
2. G. M. Zaslavsky, in *Levy Flights and Related Topics in Physics*, Ed. by M. F. Schlesinger, G. M. Zaslavsky, and U. Frisch (Springer, Berlin, 1995), Lect. Notes Phys. **450**; V. V. Uchaĭkin, Usp. Fiz. Nauk **173**, 847 (2003) [Phys. Usp. **46**, 821 (2003)].
3. A. Compte, Phys. Rev. E **53**, 4191 (1996).
4. M. M. Meerschaert, D. A. Benson, H.-P. Scheffler, and B. Baeumer, Phys. Rev. E **65**, 041103 (2002).
5. M. Bologna, P. Grigolini, and J. Riccardi, Phys. Rev. E **60**, 6435 (1999); R. Metzler, J. Klafter, and I. M. Sokolov, Phys. Rev. E **58**, 1621 (1998).
6. I. M. Sokolov and R. Metzler, Phys. Rev. E **67**, 010101(R) (2003).
7. R. Metzler and T. F. Nonnenmacher, Phys. Rev. E **57**, 6409 (1998).
8. I. M. Sokolov, Phys. Rev. E **63**, 011104 (2001).
9. V. Yu. Zaburdaev and K. V. Chukbar, Zh. Éksp. Teor. Fiz. **121**, 299 (2002) [JETP **94**, 252 (2002)].
10. A. V. Chechkin, R. Gorenflo, and I. M. Sokolov, Phys. Rev. E **66**, 046129 (2002).
11. M. M. Meerschaert, D. A. Benson, H.-P. Scheffler, and P. Becker-Kern, Phys. Rev. E **66**, 060102(R) (2002).
12. K. V. Chukbar and V. Yu. Zaburdaev, Phys. Rev. E **68**, 033101 (2003).
13. D. L. Koch and J. F. Brady, Phys. Fluids **31**, 965 (1988); Phys. Fluids A **1**, 47 (1989).
14. O. G. Bakunin, Usp. Fiz. Nauk **173**, 757 (2003) [Phys. Usp. **46**, 733 (2003)].
15. A. Z. Patashinskiĭ and V. L. Pokrovskiĭ, *Fluctuation Theory of Phase Transitions* (Nauka, Moscow, 1975; Pergamon Press, Oxford, 1979).
16. A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Physics* (Fizmatgiz, Moscow, 1962; Prentice Hall, Englewood Cliffs, N.J., 1963).
17. A. A. Abrikosov and L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **35**, 1158 (1958) [Sov. Phys. JETP **8**, 1090 (1958)]; Zh. Éksp. Teor. Fiz. **36**, 319 (1959) [Sov. Phys. JETP **9**, 220 (1959)].
18. V. M. Finkel'berg, Zh. Éksp. Teor. Fiz. **53**, 40 (1967) [Sov. Phys. JETP **26**, 268 (1967)].
19. Yu. A. Dreĭzin and A. M. Dykhne, Zh. Éksp. Teor. Fiz. **63**, 242 (1972) [Sov. Phys. JETP **36**, 127 (1972)].
20. L. D. Landau, Zh. Éksp. Teor. Fiz. **35**, 97 (1958) [Sov. Phys. JETP **8**, 70 (1958)].
21. K. V. Chukbar, Zh. Éksp. Teor. Fiz. **109**, 1335 (1996) [JETP **82**, 719 (1996)].
22. G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Process: Stochastic Models with Infinite Variance* (Chapman and Hall, New York, 1994).

Translated by A. Betev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 953–959.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1092–1099.

PLASMA, GASES

Coulomb Broadening of the Peak of Electromagnetically Induced Transparency in Plasma

S. A. Babin, M. G. Stepanov, D. V. Churkin*, and D. A. Shapiro

Institute of Automation and Electrometry, Siberian Division, Russian Academy of Sciences, Universitetskiĭ pr. 1, Novosibirsk, 630090 Russia

**e-mail: dimkins@yandex.ru*

Received November 27, 2003

Abstract—We have measured the shape of the Autler–Townes doublet and the peak of electromagnetically induced transparency (EIT) under plasma conditions. We compare the experimental results with the calculated spectrum of the probe field of a three-level ArII Λ-scheme by taking into account Coulomb collisions. We show that the Coulomb broadening of the EIT peak is small (less than 40%), while the saturation resonance is broadened under the experimental conditions by a factor of 3. In contrast to the saturation resonance attributable to the Bennett dip in the velocity distribution of the population, the EIT peak is a coherent effect and is broadened mainly through Coulomb dephasing. *© 2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

A strong resonant monochromatic wave can split the energy levels and, accordingly, the emission (absorption) spectrum at transitions involving these states—an effect that has long been known and that was initially called the dynamic Stark effect [1]. Subsequently, the splitting of the spectrum into two components was called the Autler–Townes splitting (doublet). In nonlinear gas spectroscopy, field splitting is a basic effect in the classification of perturbation theory [2]; allowance for the thermal motion of particles significantly changes its spectral manifestations. For example, in a Raman scattering scheme (see inset to Fig. 1) with large Doppler broadening, the field splitting of the probe-field spectrum manifests itself only for coaxial waves in the Stokes case, i.e., at $k_\mu < k$. If only the probe level l is populated, then other nonlinear effects induced by a strong field, in particular, the saturation effect and the nonlinear interference effect (NIEF) do not show up. In this case, the absorption spectrum of the probe field describes the field splitting in pure form.

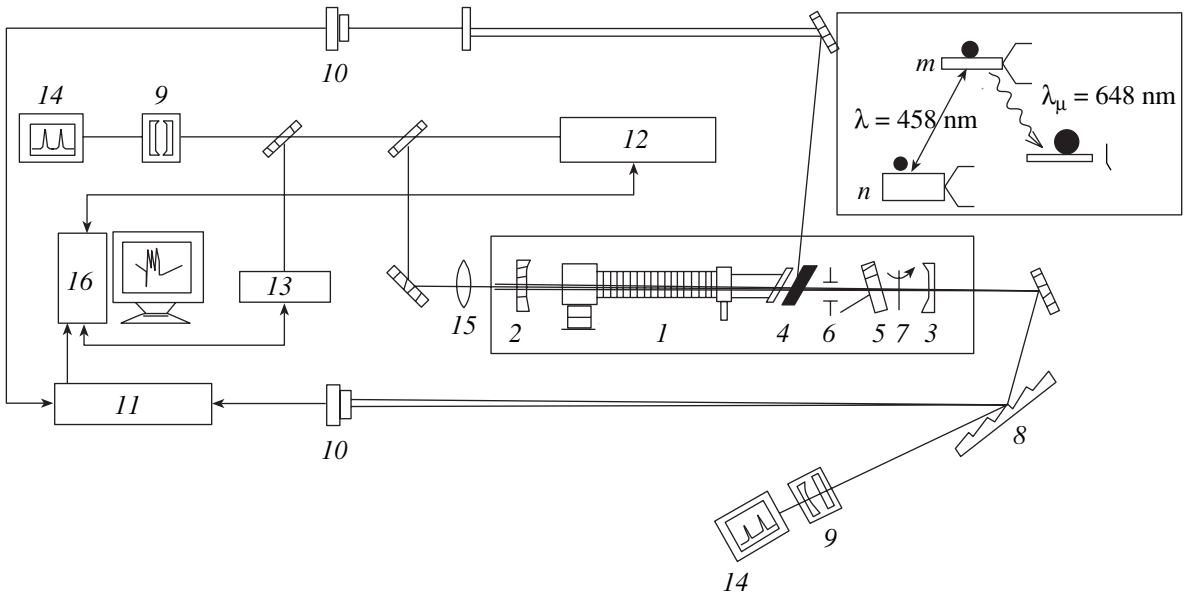


Fig. 1. Experimental setup to study field splitting: *1*—discharge tube, *2* and *3*—mirrors, *4*—Brewster plate, *5*—etalon, *6*—diaphragm, *7*—obturator, *8*—diffraction grating, *9*—scanning interferometer, *10*—photodetector, *11*—synchronous detector, *12*—tunable dye laser, *13*—wavelength meter, *14*—oscillograph, *15*—lens, and *16*—computer.

1063-7761/04/9805-0953$26.00 © 2004 MAIK "Nauka/Interperiodica"

Due to quantum interference, the field splitting of the absorption spectrum induced by a strong wave at the adjacent transition is accompanied by a significant reduction or, ideally, complete elimination of the probe-wave absorption at the frequencies corresponding to the spectral region between the split components. This effect, which was called electromagnetically induced transparency (EIT), has been actively studied in recent years (see, e.g., [3, 4] and references therein). The possibility of the elimination of light absorption under resonance conditions via the EIT effect is being actively used in various problems. In particular, the efficiency of the resonant laser frequency transformation by nonlinear optical methods [5] increases significantly, which makes it possible to use not only pulsed, but also relatively weak continuous laser emission (see, e.g., [6, 7]). In many applications, it is important to obtain the narrowest possible EIT peak, but level relaxation, and field and Doppler broadenings affect the shape of the EIT resonance [8].

In this work, we observed the Autler–Townes doublet and the EIT peak at ionic transitions in a low-temperature plasma for the first time. Ion–ion Coulomb scattering was found to also affect the shape of the peak under plasma conditions. The influence of Coulomb scattering on the shape of the resonances due to the saturation and NIEF effects under ion laser plasma conditions was studied in detail previously [9–11]. The Coulomb broadening of the EIT peak measured in this work proved to be much smaller than the Coulomb broadening of the saturation resonances. Our experimental and theoretical studies allowed us to quantitatively describe the influence of Coulomb ion–ion interaction on the field splitting and to explain the observed features.

2. EXPERIMENT

Previously, the nonlinear resonances in a Λ-scheme attributable to field splitting were experimentally studied mainly in molecular spectra (see, e.g., [12]) by using molecular Raman lasers. In this case, the observed and calculated resonance shapes are difficult to compare, because, apart from field splitting, other nonlinear effects (saturation and NIEF) contribute significantly to the total profile due to the large population of the lower level *n*. However, the measured spectrum splitting, which is proportional to the strong-field Rabi frequency

$$G = |E|d_{mn}/2\hbar$$

($|E|$ is the amplitude of the electric field, and $\hbar$ is the Planck constant), allowed the dipole moment d_{mn} of the *m*–*n* transition to be determined directly. The measurements were carried out at *G* values much larger than the relaxation constants Γ_{ij} ($i, j = m, n, l$).

As we noted above, for the profile of the Autler–Townes doublet in the Λ-scheme (see Fig. 1) to be recorded free from other nonlinear effects, the probe wave must propagate coaxially with the strong wave, and its frequency must be lower than the strong-field frequency (the Stokes case); at the same time, only the probe level *l* must be populated, while the levels *m* and *n* must be ideally empty. In contrast to thermally populated rovibrational molecular levels, ionic levels in a plasma allow the required conditions to be realized. In particular, a similar case is realized when the strong and probe fields are resonant, respectively, to the ArII laser transition and the transition to a metastable state with a large population. We chose a scheme with the following levels:

$$|n\rangle = 4s^2P_{1/2}, \quad |m\rangle = 4p^2S_{1/2}, \quad |l\rangle = 3d^2P_{3/2}.$$

The corresponding relaxation constants and Einstein coefficients (in units of 10^7 s^{-1}) are

$$\Gamma_n = 300, \ \Gamma_m = 15, \ \Gamma_l = 8, \ A_{mn} = 9, \ A_{ml} = 1.$$

The characteristic level populations in an argon laser plasma are $N_n \sim 1$, $N_m \sim 5$, and $N_l \sim 100$ (in units of 10^9 cm^{-3}) [9, 11]. Thus, the following relations hold for the level scheme chosen:

$$\Gamma_l \lesssim \Gamma_m \ll \Gamma_n \ll kv_T$$

for the relaxation constants and

$$N_l \gg N_m \gtrsim N_n$$

for the level populations.

Under argon laser plasma conditions, some of the manifestations of field splitting have been observed previously when studying the generation at coupled laser transitions in a V-scheme: when the frequency of the Stokes radiation was detuned, a decoupled resonance was observed for large detunings of the high-frequency laser field, and complex resonance structures were observed near the exact resonance for the Stokes radiation when it was tuned to the line center [13, 14]. The shape of the Autler–Townes doublet and the influence of Coulomb diffusion on it have not been investigated.

In our experiments, we studied the spectrum of a Stokes probe field in the presence of a strong field at the adjacent transition (the Λ-scheme, Fig. 1). We measured the difference between the absorption coefficients for the probe field in the absence and in the presence of a strong field that corresponded to the nonlinear correction to the probe-field work $\Delta\mathcal{P}_\mu$. For a signal of sufficient amplitude to be produced, the intensity of the strong field must be large ($G \gtrsim 100$ MHz). The intracavity field of a single-frequency 457.9-nm line ($4p^2S_{1/2}$–$4s^2P_{1/2}$) ion laser with minimum angular momentum ($j_m = j_n = 1/2$) was used to achieve these values. As a result, we measured the Autler–Townes doublet shape under argon laser plasma conditions without

the Doppler base with an accuracy high enough to make a comparison with the theory.

The experimental setup is shown in Fig. 1. An argon laser discharge tube *1* (length $l = 50$ cm, channel diameter $d = 7$ mm, and working current $I \sim 100$ A) was placed in a cavity with entrance (*2*) and exit (*3*) mirrors opaque for the generated emission, but transparent for the probe emission. Thus, there were two waves in the cavity: a standing linearly polarized generated wave and a traveling linearly polarized probe wave. Etalon *5* provided the selection of one longitudinal mode and smooth tuning of the generation frequency, while diaphragm *6* separated out the TEM_{00} mode. The cavity mirrors selected the line with a wavelength $\lambda \approx 458$ nm; the transmission losses in the cavity at this line were ≈0.3%, which provided a high field intensity inside the cavity. The emergent emission from the argon laser was directed by the mirror to a diffraction grating *8*. One order from this grating was entered into a scanning Fabry–Perot interferometer *9*, which was used to control the mode composition of the emission and to determine the strong-field detuning from the resonance, while the other order was diverted to photodetector *10*, the signal from which was the reference one for a synchronous detector *11*.

A dye laser *12* whose wavelength ($\lambda_\mu \approx 648$ nm) was recorded by a wavelength meter *13* was used as the probe-field source. The automatic frequency control (AFC) system [15] allowed us to tie the cavity mode to the selector peak and to smoothly change the probe-field frequency over a range up to 4.5 GHz. The frequency was tuned with computer *16* at discrete steps of less than 20 MHz; the step approached the emission line width (about 10 MHz). The dye-laser spectrum was recorded by the scanning interferometer *9* with a free dispersion range of 5 GHz connected to an oscillograph, which was used to control the mode composition of the emission. Before being entered into the discharge tube, the probe field was prefocused by lens *15* to provide the maximum possible field uniformity in the cavity. After the passage through the discharge tune, the probe-field beam was reflected from an additional plate *4* and diverted by the mirrors to photodetector *10* connected to the synchronous detector *11*. The angle between the beams of the probe field and the generated emission was $\sim 10^{-3}$ rad, which allowed the feedback to be avoided. The strong field was modulated at a frequency of ~1 kHz with obturator *7*; the synchronous detection at the modulation frequency allowed us to automatically subtract the Doppler base and to separate out the nonlinear corrections induced by the strong field. The personal computer *16*, to which all of the measuring instruments were connected through an ADC, was used for controlling the experiment and for synchronous data acquisition and recording.

To separate in frequency the resonances from the oppositely directed standing-wave components and to observe the field splitting induced by the traveling

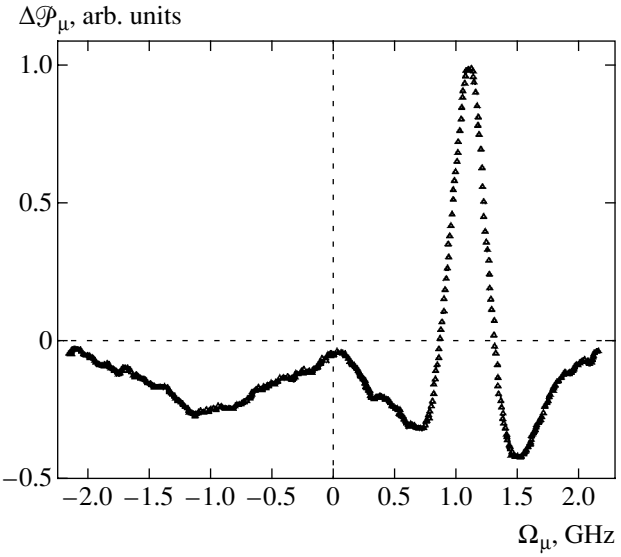


Fig. 2. Experimental profile for the nonlinear correction in the probe-field spectrum for the strong-field parameters $G \approx 100$ MHz and $\Omega \approx 1.6$ GHz.

(coaxial) wave in pure form, the strong field must be detuned from the resonance by a value larger than the population resonance width. An experimental frequency profile of the nonlinear correction to the probe-field absorption coefficient for a strong-field detuning $\Omega = \omega - \omega_{mn} \approx 1.6$ GHz is shown in Fig. 2. The negative values on the plot correspond to an increase in probe-field absorption induced by the strong field, while the positive values correspond to a decrease in absorption, which is equivalent to electromagnetically induced transparency. A sharp structure attributable to field splitting is observed for the coaxial component: the split low-amplitude absorption profile with a splitting of about 0.5 GHz and the high-amplitude EIT peak centered at a frequency $\Omega_\mu = \Omega k/k_\mu \approx 1.1$ GHz between the split components. A wide (with a FWHM of about 1.3 GHz) population resonance with a low amplitude in accordance with the level population ratio is seen symmetric about the line center at a frequency $\Omega_\mu = -\Omega k/k_\mu \approx -1.1$ GHz. The small peak at $\Omega_\mu = 0$ corresponds to the effect of higher order spatial harmonics, which is most pronounced at the exact resonance for the strong field ($\Omega = 0$) [16]; it is not considered here.

Since the contribution of the saturation effect that forms the population resonance is the same for the oppositely directed and coaxial components, we subtracted the left part of the plot ($\Omega_\mu < 0$) from its right part ($\Omega_\mu > 0$) to separate out the field splitting effect in pure form. The Autler–Townes doublet profile corrected in this way is shown in Fig. 3 together with theoretical curves computed without and with Coulomb ion–ion interaction. Since the amplitude of the doublet components for our parameters is small compared to the amplitude of the peak, it would be more precise to

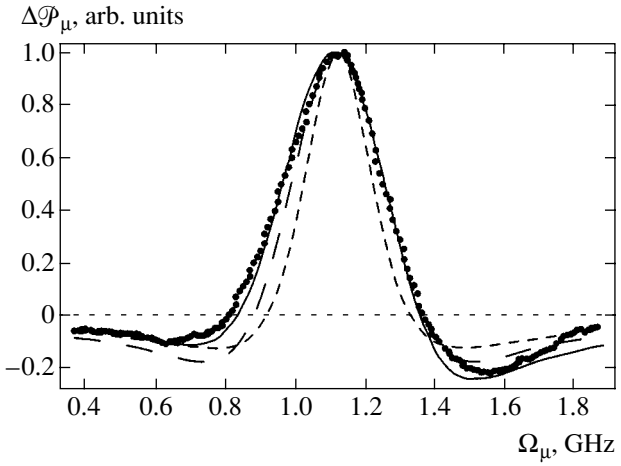


Fig. 3. Nonlinear correction in the probe-field spectrum $\Delta\mathcal{P}(\Omega_\mu)$ corresponding to the field splitting effect for the following strong-field parameters: G = 100 MHz, Ω = 1.59 GHz, and kv_T = 4.9 GHz; the short and long dashes represent the calculations using the perturbation theory without diffusion (4) and with diffusion ($\nu = 2 \times 10^7$ s^{-1}), respectively; the solid line and the dots represent the numerically calculated and experimental values, respectively.

use the term "EIT peak profile." Below, we compare the experimental and theoretical results.

3. THEORY

The nonlinear correction to the probe-field work for the Λ-scheme (Fig. 1) with $N_m = N_n \ll N_l$ calculated in the model of relaxation constants using the perturbation theory ($G \ll \Gamma_{ij}$) is [2]

$$\Delta\mathcal{P}_\mu^{(1)} = \mathcal{P}_\mu^{(1)}(0) - \mathcal{P}_\mu^{(1)}(G)$$

$$= 4\hbar\omega_\mu|G_\mu|^2|G|^2\frac{\sqrt{\pi}N_l\exp(-\Omega_\mu^2/k_\mu^2 v_T^2)(k-k_\mu)}{k^2 v_T} \quad (1)$$

$$\times \mathrm{Re}\frac{1}{(\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k))^2},$$

where $|G|$ and $|G_\mu|$ are the Rabi frequencies of the strong and probe fields, k and k_μ are their wave vectors, $\Omega = \omega - \omega_{mn}$ and $\Omega_\mu = \omega_\mu - \omega_{ml}$ are the field frequency detunings relative to the corresponding resonance, $v_T = \sqrt{2T/M}$ is the thermal velocity, and N_l is the population of level l.

This formula describes the Autler–Townes doublet in absorption with the EIT peak centered at the probe-field frequency $\Omega_\mu = k_\mu\Omega/k$ with the width

$$\Gamma_p = (k_\mu\Gamma_{nl} + (k-k_\mu)\Gamma_{ml})/k. \quad (2)$$

In this approximation, the splitting $\Delta_{AT} \approx 2\Gamma_p$ does not depend on the strong-wave intensity and is determined by the relaxation constant of the forbidden transition Γ_{nl} for close magnitudes of the wave vectors. On the other hand, it is well known that in a strong field $|G| \gg \Gamma_{ij}$ (when the contribution of the relaxation constants may be ignored), the splitting for stationary atoms is determined by its Rabi frequency $|G|$, while allowance for the thermal motion leads to the addition of a scaling factor that depends on the relation between the wave vectors of the probe and strong fields (see, e.g., 12, 17]):

$$\Delta_{AT} = 4|G|\sqrt{(1-k_\mu/k)k_\mu/k}. \quad (3)$$

For an arbitrary relation between $|G|$ and Γ_{ij}, the expression for the nonlinear correction calculated in the Doppler limit ($|G|, \Gamma_{ij} \ll kv_T$) is [17]

$$\Delta\mathcal{P}_\mu = 2\hbar\omega_\mu|G_\mu|^2\frac{\sqrt{\pi}N_l\exp(-\Omega_\mu^2/k_\mu^2 v_T^2)}{k_\mu v_T} \quad (4)$$

$$\times\left(1 - \mathrm{Re}\frac{\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k)}{\sqrt{(\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k))^2 + \dfrac{4k_\mu(k-k_\mu)|G|^2}{k^2}}}\right).$$

In the limit $|G| \ll \Gamma_{ij}$, this expression reduces to (1), a result of perturbation theory. As the field amplitude increases, the splitting increases and is described by Eq. (3) in the limit $|G| \gg \Gamma_{ij}$.

Under the experimental conditions ($|G| \approx 100$ MHz, $\Gamma_{mn} \approx \Gamma_{nl} \approx 280$ MHz, $\Gamma_{ml} \approx 25$ MHz), the approximation of perturbation theory ($|G| \ll \Gamma_{mn}, \Gamma_{nl}$) holds well. The profile calculated using formula (4) with the field splitting (determined by the Rabi frequency $|G|$) is virtually identical to result (1) of perturbation theory—the latter curve is indicated by short dashes in Fig. 3. The FWHM of the EIT peak (and, accordingly, the splitting) in this approximation is determined by the relaxation constant of the forbidden transition

$$\Delta_{AT} \sim 2\Gamma_p \approx 2\Gamma_{nl}k_\mu/k \approx 400 \text{ MHz}.$$

The calculated curve qualitatively agrees with the experimental curve, but the width of the EIT peak in the experiment is appreciably larger (by about 40%); allowance for the field broadening yields no such broadening.

Coulomb ion scattering [9] is known to be mainly responsible for the broadening of nonlinear resonances in an ion laser plasma. This scattering is satisfactorily

described by a model of diffusion in velocity space with a velocity-independent coefficient [18, 19]:

$$D = \nu v_T^2/2, \quad \nu = \frac{16\sqrt{\pi}NZ^2e^4\Lambda}{3M^2v_T^3}, \tag{5}$$

where ν is the effective ion–ion collision frequency; $v_T = \sqrt{2T_i/M}$ is the thermal velocity; Ze and M are the charge and mass of the active ions, respectively; N is the effective number density of the perturbing ions; and Λ is the Coulomb logarithm.

The Coulomb broadening of the resonances due to saturation was studied in detail in an experiment; in particular, it was shown that the Coulomb broadening could reach a factor of 100 with respect to the radiative width for long-lived metastable levels [11]. The pattern of the diffusive broadening of population resonances is fairly easy to understand: a strong monochromatic wave produces Bennett structures with a width $\Gamma_{mn}/k \ll v_T$ against the background of a Maxwellian velocity distribution for the population of level $j = m, n$ with width v_T. Diffusion in velocity space tends to level off the nonequilibrium, causing the resonant structure to be broadened. The characteristic change in velocity increases with time t following the diffusion law

$$\Delta v_j \sim \sqrt{Dt}.$$

Over the level lifetime Γ_j^{-1}, diffusion in velocities causes the saturation resonance in the spectrum to be broadened by

$$\Delta_j = k\Delta v_j \approx \frac{kv_T}{2}\sqrt{\nu/\Gamma_j}, \quad j = m, n; \tag{6}$$

i.e., the longer the level lifetime, the larger the broadening of the saturation resonance, as distinct from the model of relaxation constants. For laser transitions, the Bennett dip is broadened predominantly at the relatively long-lived upper level. For our level scheme, the characteristic broadening of the saturation resonance is $\Delta_j/\Gamma_{mn} \sim 3$; accordingly, the width of the population resonance observed for the oppositely directed strong and probe waves centered at $\Omega_\mu \approx -1.1$ GHz (see Fig. 2) is larger by a factor of about 3 than the width of the EIT peak.

Since the EIT peak is produced by coherent effects, the Coulomb broadening mechanism in this case differs fundamentally from the broadening mechanism of population resonances. Apart from a change in the population distribution, diffusion in the velocity space also leads to dephasing (phase diffusion) of the nondiagonal density (coherence) matrix element through a random change in the ion coordinate:

$$\langle \Delta r^2 \rangle \sim \Delta v^2 t^2 \sim Dt^3,$$

which corresponds to the change in phase

$$\langle \Delta\varphi^2 \rangle = k^2 \langle \Delta r^2 \rangle \sim Dk^2 t^3.$$

The dephasing is significant when $\Delta\varphi \sim 1$. Hence, we can estimate the dephasing time scale τ_D and the related correction to the homogeneous transition width:

$$\tau_D^{-1} \sim (Dk^2)^{-1/3} \approx (\nu(kv_T)^2)^{-1/3}. \tag{7}$$

Accordingly, the correction to the width of the EIT peak (2) under experimental conditions is estimated as $k_\mu \tau_D^{-1}/k \approx 300$ MHz, which is appreciably larger than $\Gamma_p \approx 200$ MHz. This value is in conflict with the experiment, in which the observed broadening is appreciably smaller than Γ_p.

Since perturbation theory may be used to describe the experiment, the effect can be analyzed in more detail. For coaxial strong and probe waves in the Stokes case ($k_\mu < k$), we may use the nonlinear correction in the probe-field spectrum calculated using perturbation theory up to the second order in $|G|$ with diffusion in velocities [20]. Reducing the expression to a more familiar form for nonlinear spectroscopy, we obtain the line profile that corresponds to the field splitting:

$$\Delta\mathcal{P}(\Omega_\mu) = \frac{4\sqrt{\pi}\hbar\omega_\mu|G_\mu|^2|G|^2N_l(k-k_\mu)}{k^2v_T}$$

$$\times \mathrm{Re}\left\{\int_0^\infty dt \exp(i\Omega_\mu t)\Phi(t)\right\}^2, \tag{8}$$

$$\Phi(t) = \exp\{-(\Gamma_p + i\Omega k_\mu/k)t$$

$$- D(k-k_\mu)^2(k_\mu/k)^2t^3/3\}.$$

Here, we ignore the force of friction, because the resonant velocity is less than $0.4v_T$. The phase diffusion is determined not by the factor Dk^2, as suggested by estimate (7), but by a factor of $(k-k_\mu)^2k_\mu^2/k^4$ smaller quantity. Accordingly, the diffusion width of the field-splitting resonance may be expressed as

$$\Gamma_D \approx [D(k-k_\mu)^2k_\mu^2/k^2]^{-1/3}. \tag{9}$$

In the experiment,

$$k_\mu/k \approx 0.7, \quad (k-k_\mu)^2k_\mu^2/k^4 \approx 0.04.$$

Thus, the diffusion width of the field-splitting resonance, $\Gamma_D \approx 100$ MHz, is by a factor of about 3 smaller than $(Dk^2)^{1/3}$; the diffusive broadening is small, $\Gamma_D < \Gamma_p$. In the limit $D \longrightarrow 0$, the expression for the profile

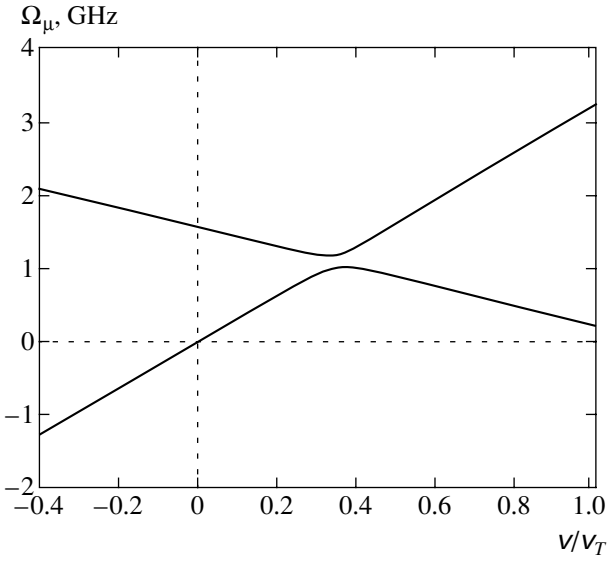


Fig. 4. Frequency branches calculated using formula (10) for the following experimental conditions: $\Omega \approx 100$ MHz and $\Omega = 1.59$ GHz.

shape reduces to (1). The result of our calculation using formula (8) of perturbation theory with Coulomb diffusion for experimental conditions is indicated by long dashes in Fig. 3. The curve satisfactorily describes the experiment; the slight deviations are attributable mainly to asymmetry in the experimental profile. The curve (solid line) calculated numerically from the system of equations for the density matrix with the diffusive and field broadenings and with the force of Coulomb friction is also shown in the figure. This curve is in even better agreement with the experiment; it also describes the asymmetry. On the right slope of the EIT peak, the results of our numerical calculations and calculations using perturbation theory with diffusion are in close agreement; deviations are clearly seen only on the left slope.

4. DISCUSSION

Our comparison of the experimental and calculated curves shows that, in contrast to the saturation resonances, population diffusion in velocities does not lead to any significant broadening of the resonance due to field splitting. A good approximation to describe the experiment is perturbation theory with diffusion; in this case, the field broadening is negligible. The opposite limiting case was analyzed in [21]: the diffusion shape of the Autler–Townes doublet components was calculated for a field splitting $|G|$ much larger than the resonance width. In this case, the diffusion width of the split components was found to be $\sim\sqrt{Dk^2/|G|}$, i.e., it decreases with increasing $|G|$, and no appreciable broadening of the resonances was observed in an experiment with strong fields in the *V*-scheme [13, 14].

Since the amplitude of the Autler–Townes doublet components in our case is small compared to the amplitude of the EIT peak, it is of considerable interest to discuss the influence of Coulomb diffusion on the shape of the EIT peak—its relative broadening under experimental conditions did not exceed 40%. Qualitatively, such a weak influence can be understood by using the pattern of frequency branches: the velocity dependence of the resonant frequencies of the split components (see, e.g., [13]). The resonant frequencies when the homogeneous width is ignored are described in our case by the expression

$$\Omega_\mu(v) = k_\mu v + (\Omega - kv)/2 \pm \sqrt{(\Omega - kv)^2/4 + |G|^2}. \quad (10)$$

The results of our calculation using formula (10) for experimental parameters are shown in Fig. 4. When averaged over velocities, the integral is accumulated in the vicinity of the extrema of the function $\Omega_\mu(v)$, called turning frequencies [13]. The size of the vicinity that gives a significant contribution is determined by the slope of the function, $d\Omega_\mu/dv$—the asymptotic behavior at large velocities is determined by the coefficients k_μ and $(k_\mu - k)$, which differ greatly in our case. Therefore, the integral is accumulated at velocities $v > 0.4v_T$ for the resonance $\Omega_\mu^1 < 1$ GHz and at $v < 0.4v_T$ for the second resonance $\Omega_\mu^2 < 1$ GHz. Given the Maxwellian distribution function, this leads to a larger amplitude of the resonance that is farther from the center of the line, as confirmed by the experiment (see Fig. 2). Formulas (1), (4), and (8) derived in the Doppler limit do not describe the asymmetry. The role of Coulomb population diffusion under these conditions reduces to the walk of particles on the frequency branch along the velocity axis in a vicinity of the order

$$\Delta v_j \sim v_T\sqrt{\nu_{ii}/2\Gamma_j}, \quad j = m, n,$$

whose size does not exceed the size of the region that contributes to the integral; therefore, this effect is weak, with the influence of diffusion on the wing shape being stronger than on the width of the EIT peak. The asymmetry in the doublet components is the result of averaging over velocities with allowance made for the finite Doppler width and is virtually independent of diffusion. Thus, the main broadening mechanism in this case is Coulomb dephasing (phase diffusion), whose effect, in turn, is weakened by a factor of $[(k - k_\mu)^2 k_\mu^2/k^4]^{-1/3} \sim 3$ under experimental conditions. As a result, the Coulomb broadening of the EIT peak does not exceed 40%.

The experiment also allows us to determine the width of the population resonance (attributable to the Bennett dip at the upper level *m*), which is observed for oppositely directed probe and strong fields—the negative detuning range in Fig. 2. It is of considerable interest to compare our result with the data of previous

experiments on the Lamb dip and the spontaneous emission spectrum (see, e.g., [9]). The FWHM of the saturation resonance in our case is $\Delta \approx 1.3$ GHz, which corresponds to a relative broadening of the Bennett dip by a factor of $\gamma = (\Delta/2\Gamma_{mn})(k/k_\mu) \approx 3.7$. This value is slightly larger than that yielded by measurements of the spontaneous emission spectrum under the same conditions (see [9]). In contrast to previous measurements, the field broadening of the population resonance in our experimental conditions was large and it could not be ignored. Our estimation of the characteristic values yields the following Coulomb, homogeneous, and field widths of the saturation resonance under experimental conditions:

$$\Delta_m = k_\mu v_T \sqrt{\nu_{ii}/2\Gamma_m} \approx 0.7 \text{ GHz}$$

($\Delta_D = 2\ln 2\Delta_m \approx 1$ GHz is the FWHM),

$$2\Gamma_{mn} k_\mu/k \approx 0.35 \text{ GHz},$$

$$\Delta_G \approx 2|G|\sqrt{2\Gamma_{mn}/\Gamma_m} \approx 0.8 \text{ GHz}.$$

Consequently, the diffusion and field widths are comparable in magnitude and are appreciably larger than the homogeneous width: $\Delta_D \gtrsim \Delta_G > 2\Gamma_{mn}k_\mu/k$. As was shown in [9, 22], the squares of the field and diffusion widths are added in these conditions; i.e., the total width may be expressed as

$$\Delta = \sqrt{\Delta_D^2 + \Delta_G^2} \approx 1.3 \text{ GHz}, \quad (11)$$

in good agreement with the measurements.

5. CONCLUSIONS

Thus, we have measured for the first time the shape of the nonlinear resonance due to field splitting under plasma conditions. Our experimental and theoretical studies of this effect at relatively low field intensities ($|G| < \Gamma_{nl}$) show that Coulomb ion–ion scattering, which leads to ion diffusion in the velocity space, affects the Autler–Townes doublet profile and the width of the EIT peak only slightly. This influence reduces to a small (about 40%) broadening of the peak and to a change in the wing shape of the split components. The main broadening mechanism is Coulomb dephasing (phase diffusion), which causes an effective increase in the homogeneous width by Γ_D described by (9). We have shown that the broadening of the saturation resonance via Coulomb population diffusion under these conditions is almost an order of magnitude larger; the field broadening in this case also gives a significant contribution.

ACKNOWLEDGMENTS

We would like to thank S.I. Kablukov, E.V. Podivilov, and S.G. Rautian for helpful discussions; O.V. Belai for help with numerical calculations; and V.V. Potapov who designed the experimental setup. This work was supported by the Russian Foundation for Basic Research (project no. 02-02-39025) and the Program of State Support for Leading Scientific Schools (project no. NSh-439.2003.2).

REFERENCES

1. S. H. Autler and C. H. Townes, Phys. Rev. **100**, 703 (1955).
2. S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in Atomic and Molecular Spectra* (Nauka, Novosibirsk, 1979).
3. S. E. Harris, Phys. Today **50**, 36 (1997).
4. J. P. Marangos, J. Mod. Opt. **45**, 471 (1998).
5. S. E. Harris, J. E. Field, and A. Imamoğlu, Phys. Rev. Lett. **64**, 1107 (1990).
6. S. Babin, U. Hinze, E. Tiemann, and B. Wellegehausen, Opt. Lett. **21**, 1186 (1996).
7. A. S. Zibrov, M. D. Lukin, and M. O. Scully, Phys. Rev. Lett. **83**, 4049 (1999).
8. C. Y. Ye and A. S. Zibrov, Phys. Rev. A **65**, 023806 (2002).
9. S. A. Babin and D. A. Shapiro, Phys. Rep. **241**, 119 (1994).
10. S. A. Babin, S. I. Kablukov, M. A. Kondratenko, and D. A. Shapiro, Pis'ma Zh. Éksp. Teor. Fiz. **64**, 241 (1996) [JETP Lett. **64**, 263 (1996)].
11. A. A. Apolonsky, S. A. Babin, A. I. Chernykh, *et al.*, Phys. Rev. A **55**, 661 (1997).
12. B. Wellegehausen, IEEE J. Quantum Electron. **15**, 1108 (1979).
13. O. G. Bykova, V. V. Lebedeva, N. G. Bykova, and A. V. Petukhov, Opt. Spektrosk. **53**, 171 (1982) [Opt. Spectrosc. **53**, 101 (1982)].
14. O. G. Bykova, L. E. Grin', V. V. Lebedeva, and A. É. Sedel'nikova, Opt. Spektrosk. **64**, 1216 (1988) [Opt. Spectrosc. **64**, 725 (1988)].
15. B. V. Bondarev, S. M. Kobtsev, A. V. Karablev, and V. M. Lunin, Opt. Atmos. **2**, 1319 (1989).
16. S. A. Babin, E. V. Podivilov, V. V. Potapov, *et al.*, Zh. Éksp. Teor. Fiz. **121**, 807 (2002) [JETP **94**, 694 (2002)].
17. B. J. Feldman and M. S. Feld, Phys. Rev. A **5**, 899 (1972).
18. S. G. Rautian, Zh. Éksp. Teor. Fiz. **51**, 1176 (1966) [Sov. Phys. JETP **24**, 788 (1966)].
19. G. I. Smirnov and D. A. Shapiro, Zh. Éksp. Teor. Fiz. **76**, 2084 (1979) [Sov. Phys. JETP **49**, 1054 (1979)].
20. S. G. Rautian and D. A. Shapiro, Zh. Éksp. Teor. Fiz. **94** (10), 110 (1988) [Sov. Phys. JETP **67**, 2018 (1988)].
21. M. G. Stepanov and D. A. Shapiro, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 27 (1998) [JETP Lett. **68**, 29 (1998)].
22. M. G. Stepanov and D. A. Shapiro, Zh. Éksp. Teor. Fiz. **113**, 1632 (1998) [JETP **86**, 888 (1998)].

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 960–973.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1100–1114.

FLUIDS

Rayleigh Instability in Liquid-Crystal Jets¶

L. G. Fel and Y. Zimmels

Department of Civil and Environmental Engineering Technion 32000, Haifa, Israel
e-mail: lfel@techunix.technion.ac.il

Received November 14, 2003

Abstract—Capillary instability of isothermal incompressible liquid-crystal (LC) jets is considered within the linear hydrodynamics of uniaxial nematic LCs. Free boundary conditions with strong tangential anchoring of director **n** at the surface are formulated in terms of the mean surface curvature $\mathcal{H}$ and the Gaussian surface curvature $\mathcal{G}$. The static version of the capillary instability is shown to depend on the elasticity modulus K, the surface tension σ_0, and the radius r_0 of the LC jet, expressed in terms of the characteristic parameter $\kappa = K/\sigma_0 r_0$. The problem of the capillary instability in LC jets is solved exactly, and a dispersion relation that reflects the effect of elasticity is derived. It is shown that increase in the elasticity modulus results in decrease in both the cut-off wavenumber *k* and the disturbance growth rate *s*. This implies an enhanced stability of LC jets in comparison to ordinary liquids. In the specific case where the hydrodynamic and orientational LC modes can be decoupled, the dispersion equation is given in a closed form. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The breakup of liquid jets injected via a circular nozzle into stagnant fluids has been the subject of widespread research over the years. Previous studies that followed the seminal works of Lord Rayleigh have established that a complex jet flow is influenced by a large number of parameters. These include nozzle internal flow effects, the jet velocity profile **V**(**r**), and the physical state of both liquid and gas. Although the hydrodynamic equations are nonlinear, the linear stability theory can provide qualitative descriptions of the breakup phenomena and predict the existence of different breakup regimes.

Using a linear theory, Rayleigh showed [1] that the jet breakup is a consequence of the hydrodynamic instability, or more exactly, the capillary instability. Neglecting the effect of the ambient fluid, the viscosity of the jet liquid, and gravity, he demonstrated that a cylindrical liquid jet is unstable with respect to disturbances characterized by wavelengths larger than the jet circumference. Rayleigh also considered the case of a viscous jet in an inviscid gas and an inviscid gas jet in an inviscid liquid [2]. Weber [3] extended Rayleigh's result to Newtonian viscous fluids and showed that viscosity tends to reduce the breakage rate and increase the drop size. Chandrasekhar [4] considered the effect of a uniform magnetic field on the capillary instability of a liquid jet. A mechanism of bending disturbances and of buckling, slowly moving, highly viscous jets was presented by Taylor [5]. Further developments of the theory in Newtonian liquids was concerned with additional factors such as the dynamic action of the ambient gas (leading to atomization of the jet), the nonlinear interaction of growing modes that lead to satellite drop formation, and the spatial character of instability [6, 7].

The capillary instability in jets of non-Newtonian suspensions and emulsions presents a different category of cases governed by power-law (pseudoplastic and dilatant) liquids. The effective viscosity of the pseudoplastic liquid decreases with increasing strain rate, whereas in dilatant liquids, it increases [7]. The behavior of capillary jets of dilute and concentrated polymer solutions suggests a strong influence of the macromolecular coils on their flow patterns [7]. Free jets of polymeric liquids that exhibit oscillations are reported in [8].

The idea of the Rayleigh instability was applied to tubular membranes in dilute lyotropic phases [9]. Their relaxation, following optical excitation, is characterized by a long time and can be described by means of the hydrodynamic approach [10]. Bending deformations of such membranes are governed by the Helfrich energy [11], which depends on the curvature of the tube. Thus, competition between the surface tension and curvature energy of the membrane that has been immersed in water renders the initial shape of the tube unstable. The hydrodynamic formalism used in [10] and the hydrodynamics of fluids with an inner order such as liquid crystals (LCs) [12] have similar features. In [10], the order parameter is a unit vector normal to the membrane surface. In contrast, the order parameter *Q* of an LC fluid is defined throughout the space it occupies.

The continuum theory of LC phases has emerged as a rigorous part of condensed matter theory. The hydrodynamics of the nematic LC phases was developed during the 1960s–1970s in the pioneering works of Erick-

¶This article was submitted by the authors in English.

Table 1. The basic physical parameters α_i, ρ, K, and σ_0 and their derivatives η_i, β_i, γ_i, B_i, μ_i, λ, and ν_i for nematic LC 4-methoxybenzylidene-4-butylaniline (MBBA) at 25°C taken from [23, 30]

α_1, mPa s	α_2, mPa s	α_3, mPa s	α_4, mPa s	α_5, mPa s	α_6, mPa s
7	–78	–1	84	46	–33
η_1, mPa s	η_3, mPa s	η_5, mPa s	λ	μ_1	μ_2
42	50	104	1.026	1.013	0.013
β_1, mPa s	β_2, mPa s	β_3, mPa s	β_4, mPa s	γ_1, mPa s	γ_2, mPa s
42	25	79	59	77	–79
B_1, mPa s	B_2, mPa s	B_3, mPa s	B_4, mPa s	$\mathcal{B}$	ϑ, m^2/s
58	104	25	78	5.92	1.2×10^{-10}
ρ, kg/m^3	K, N	σ_0, N/m	$l = K/\sigma_0$, m	ν_i, m^2/s	ϑ/ν_i
1.2×10^3	9×10^{-12}	38×10^{-3}	2.4×10^{-10}	10^{-5}–10^{-4}	10^{-6}–10^{-5}

Table 2. The basic physical parameters α_i, ρ, K, and σ_0 and their derivatives η_i, β_i, γ_i, B_i, μ_i, λ, and ν_i for LC paraazoxyanisole (PAA) at 122°C taken from [23, 30]

α_1, mPa s	α_2, mPa s	α_3, mPa s	α_4, mPa s	α_5, mPa s	α_6, mPa s
4	–6.9	–0.2	6.8	5	–2.1
η_1, mPa s	η_3, mPa s	η_5, mPa s	λ	μ_1	μ_2
3.4	4.5	13.7	1.06	1.03	0.03
β_1, mPa s	β_2, mPa s	β_3, mPa s	β_4, mPa s	γ_1, mPa s	γ_2, mPa s
3.4	2.25	11.45	4.55	6.7	–7.1
B_1, mPa s	B_2, mPa s	B_3, mPa s	B_4, mPa s	$\mathcal{B}$	ϑ, m^2/s
4.34	9.36	2.26	11.24	7.11	1.8×10^{-9}
ρ, kg/m^3	K, N	σ_0, N/m	$l = K/\sigma_0$, m	ν_i, m^2/s	ϑ/ν_i
1.4×10^3	11.9×10^{-12}	40×10^{-3}	3×10^{-10}	10^{-6}–10^{-5}	10^{-4}–10^{-3}

sen [13, 14], Leslie [15 16], Parodi [17], and the Harvard Group [18],[1] and its predictions were successfully confirmed in many experimental observations. The combination of viscous and elastic properties is likely to produce new evolution patterns of hydrodynamic instabilities, in terms of Benard–Rayleigh, Marangoni, and electrohydrodynamic effects [19], which cannot occur in ordinary liquids. In particular, we refer to nonsteady-state (oscillatory) evolution of the instability that appears via the Hopf bifurcation [20]. The instability of an LC jet poses an additional challenge with respect to the effects listed above. This already is already applicable in the framework of the linear stability theory.

Linear analysis of the capillary instability in a thin nematic LC fiber was recently performed in [21] under the assumption that the director field $\mathbf{n}(\mathbf{r})$ is fixed and does not change even if the fiber shape evolves via the linear instability process. In this analysis, the only influence of the LC nematicity is due to the anisotropy of the elastic surface energy and the anisotropy of viscous LC moduli. The above assumption stipulates the predominance of elastic forces over the surface tension ($l \gg r_0$) and over hydrodynamic forces ($Er \ll 1$), where $l = K/\sigma$ and $Er = \eta V r_0/K$ denote the anchoring extrapolation length [22] and the Ericksen number [19], respectively; η and K are viscous and elastic moduli; V is the LC velocity; and r_0 stands for the geometric length scale, i.e., the radius of the LC jet. The first condition ($l \gg r_0$) is difficult to implement for most known nematic LCs with well-studied physical parameters. Indeed, classical nematic LCs, also known as MBBA and PAA, have an anchoring extrapolation length of $l \approx 3 \times 10^{-10}$ m (Tables 1 and 2). This value indicates strong

[1] The name Harvard Group was proposed by De Gennes [19] and denotes five authors (see [18]).

anchoring at the surface.[2] Otherwise, the radius of the jet must be decreased to the molecular scale. In the case of strong director anchoring at the surface, the second requirement, $Er \ll 1$, does not allow a continuous transition to ordinary liquids (the classical Rayleigh–Weber theory) which is an important benchmark in the theory. We note that as the elasticity tends to zero, $K \longrightarrow 0$ then $Er \longrightarrow \infty$. Moreover, disregarding the bulk elasticity effects in LCs excludes the competition between the bulk forces and surface tension that is crucial for the physical picture of thin LC films (see Section 3). In this context [21], the Leslie–Ericksen equation of angular motion of the director $\mathbf{n}(\mathbf{r}, t)$ was neglected and the elastically induced nondissipative contributions to the Navier–Stokes equation were not included in the LC hydrodynamics.

A more realistic setup of the problem consists of a rigid boundary condition of strong director anchoring at the free surface of LC jets. The simplest case constitutes a tangential orientation of the director at the surface, $\mathbf{n} \cdot \mathbf{e} = 0$, where $\mathbf{e}$ is a unit vector normal to the jet surface. Such orientation, having strong anchoring and temperature independence, is observed at the free surface of the classical nematic PAA mentioned above [24–26]. Assuming that the scale of deformation of the initial surface is much larger than the molecular length of LCs, we conclude that if the orientation of the director $\mathbf{n}$ is set tangential to the undisturbed surface, then it must also remain tangential when the surface is smoothly disturbed.

The elastic properties of LCs are expected to change the evolution patterns of jets that are made from them. In this paper, we derive a rigorous mathematical model of capillary instability for isothermal incompressible nematic LC jets in the single elastic approximation. This model shows how the combined viscous and elastic properties of LC fluids determine the boundary conditions at the free surface with strong tangential anchoring of the director and the range where instability prevails.

2. HYDRODYNAMICS OF AN LC JET

In this section, we first formulate the problem of capillary instability and then derive the basic equations that govern the linear hydrodynamics of an LC jet. The incompressible flow of a nematic LC is described by a set of differential equations: the continuity equation, the Nayier–Stokes equation for viscoelastic LCs, and the Leslie–Ericksen equation of angular motion of the director $\mathbf{n}(\mathbf{r}, t)$. They are supplemented by boundary conditions on the LC free surface with strong tangential anchoring of the director.

The basic notation and linear hydrodynamic equations for uniaxial nematic LCs follow the theory given in [18] (the so-called Harvard Group approach), which has become standard in many monographs, e.g., [12, 27]. We note that the Harvard Group and Ericksen–Leslie–Parodi approaches are in full agreement (a detailed discussion is given in [19]).

2.1. Basic Notation and Variables

The following basic variables describe the nematic LC medium: velocity $\mathbf{V}(\mathbf{r}, t)$, pressure $P(\mathbf{r}, t)$, and director $\mathbf{n}(\mathbf{r}, t)$, $\mathbf{n}^2 = 1$. The initial values of the functions are denoted by "0," either as a subscript or superscript. The following notation, which is commonly accepted in the theory of LCs, is used henceforth:

1. The free energy density E_d of a deformed nonchiral uniaxial nematic LC, given in the quadratic approximation in terms of the derivatives $\partial\mathbf{n}/\partial x_j$ and in the single elastic approximation, has the form

$$E_d = \frac{K}{2}(\mathrm{div}^2\mathbf{n} + \mathrm{rot}^2\mathbf{n}), \tag{1}$$

where $K \geq 0$ is known as the Frank elasticity modulus. In the vicinity of a phase transition, $K \propto Q^2$ [19], and in the isotropic phase, it vanishes.

2. The bulk molecular field $\mathbf{F}$ and the Ericksen elastic stress tensor τ_{ki}, which set the equilibrium distribution of the $\mathbf{n}$ field in an LC, are determined by the variational derivatives:[3]

$$\mathbf{F} = \mathbf{M} - \mathbf{n}\langle\mathbf{n}, \mathbf{M}\rangle, \text{ or } F_i = (\delta_{ij} - n_i n_j)M_j, \tag{2}$$

where

$$M_i = \frac{\partial}{\partial x_k}\frac{\partial E_d}{\partial(\partial_k n_i)} - \frac{\partial E_d}{\partial n_i},$$
$$\tau_{ki} = \frac{\partial E_d}{\partial(\partial_k n_i)}, \quad \partial_k = \frac{\partial}{\partial x_k}, \tag{3}$$

i.e.,

$$\mathbf{M} = K\Delta_3\mathbf{n}, \quad \tau_{ki} = K(\delta_{ki}\mathrm{div}\mathbf{n} + (\mathbf{n}\cdot\mathrm{rot}\mathbf{n})n_m\epsilon_{mki} + [[\mathbf{n}\times\mathrm{rot}\mathbf{n}]\times\mathbf{n}]_m\epsilon_{mki}), \tag{4}$$

where ϵ_{mki} is a completely antisymmetric unit tensor of the third rank (the Levi-Civita tensor).

[2] On the basis of an heuristic argument, De Gennes [19] noted that if the anisotropic interaction at a nematic-substrate interface is as large as that acting between nematic molecules, the anchoring energy σ can be roughly estimated as $\sigma \sim K/a$, where K is the Frank modulus and a is the molecular dimension; hence, taking $K \approx 8 \times 10^{-12}$ N and $a \approx 5 \times 10^{-10}$ m, we find $\sigma \approx 1.6 \times 10^{-2}$ N/m, which corresponds to the strong anchoring in virtually all practical cases. An extensive review by Cognard [23] lists sixteen of the most studied nematic LCs with corresponding σ measured at equilibrium with air (see Table 9 in [23]). All values lie in the range between 2.45×10^{-2} N/m for MPPB and 4×10^{-2} N/m for 5CB, which well supports the qualitative consideration of De Gennes.

[3] Here and throughout, unless noted otherwise, we apply the summation rule over indices repeated in a tensor product, e.g., $a_{ij}b_{jk} = \sum_j a_{ij}b_{jk}$.

3. If the deviations of the director $\mathbf{n} = \mathbf{n}^0 + \mathbf{n}^1$ from its initial orientation $\mathbf{n}^0$ along the z direction are small, then

$$\begin{gathered} n_x^0 = n_y^0 = 0, \quad n_z^0 = 1, \\ 1 \gg n_x^1, n_y^1 \gg n_z^1 \sim (n_x^1)^2, (n_y^1)^2, \end{gathered} \tag{5}$$

and simple algebra yields the linear approximation

$$F_x = K\Delta_3 n_x^1, \quad F_y = K\Delta_3 n_y^1, \quad F_z = 0, \tag{6}$$

where Δ_3 is the three-dimensional Laplacian. Similar considerations regarding the Ericksen stress tensor τ_{ki} give

$$\begin{gathered} \tau_{xx} = \tau_{yy} = \tau_{zz} = K\mathrm{div}\mathbf{n}^1, \\ \tau_{xy} = -\tau_{yx} = K\left(\frac{\partial n_y^1}{\partial x} - \frac{\partial n_x^1}{\partial y}\right), \\ \tau_{yz} = -\tau_{zy} = K\left(\frac{\partial n_z^1}{\partial y} - \frac{\partial n_y^1}{\partial z}\right), \\ \tau_{zx} = -\tau_{xz} = K\left(\frac{\partial n_x^1}{\partial z} - \frac{\partial n_z^1}{\partial x}\right). \end{gathered} \tag{7}$$

The stresses given by Eqs. (7) do not contribute to the nondissipative stress tensor $T_{ik}^{(r)}$ used in the linear hydrodynamics of LCs (see Eq. (8) below).

4. The reactive (nondissipative) $T_{ik}^{(r)}$ and dissipative $T_{ik}^{(d)}$ stress tensors are defined as

$$\begin{aligned} T_{ik}^{(r)} = &-P\delta_{ik} - \tau_{kj}\frac{\partial n_j}{\partial x_i} - \frac{\lambda}{2}(n_i F_k + n_k F_i) \\ &+ \frac{1}{2}(n_i F_k - n_k F_i), \end{aligned} \tag{8}$$

$$\begin{aligned} T_{ik}^{(d)} = &2\eta_1 \Upsilon_{ik} + (\eta_3 - 2\eta_1)(n_i \Upsilon_{kj} n_j + n_k \Upsilon_{ij} n_j) \\ &+ (2\eta_1 + \eta_5 - 2\eta_3) n_i n_k n_j n_m \Upsilon_{jm}, \end{aligned} \tag{9}$$

where the antisymmetric Ω_{ik} (vorticity) and symmetric Υ_{ik} parts of the derivative $\partial_k V_i$ are given by

$$\begin{gathered} \Omega_{ik} = \frac{1}{2}\left(\frac{\partial V_k}{\partial x_i} - \frac{\partial V_i}{\partial x_k}\right), \\ \Upsilon_{ik} = \frac{1}{2}\left(\frac{\partial V_k}{\partial x_i} + \frac{\partial V_i}{\partial x_k}\right). \end{gathered} \tag{10}$$

Three independent viscous moduli η_j, the kinetic coefficient λ, and the rotational viscosity γ_1 determine the dissipative stress tensor $T_{ik}^{(d)}$, the fourth-rank viscosity tensor η_{ikjm}, and the dissipative function D in the absence of heat fluxes,

$$\begin{gathered} D = \eta_{ikjm}\Upsilon_{ik}\Upsilon_{jm} + \frac{1}{\gamma_1}\mathbf{F}^2, \quad T_{ik}^{(d)} = \eta_{ikjm}\Upsilon_{jm}, \\ \eta_{ikjm} = \eta_1(\xi_{ij}\xi_{km} + \xi_{kj}\xi_{im}) \\ + \frac{\eta_3}{2}(n_i n_j \xi_{km} + n_k n_j \xi_{im} + n_i n_m \xi_{kj} + n_k n_m \xi_{ij}) \\ + \eta_5 n_i n_k n_j n_m. \end{gathered} \tag{11}$$

The tensor η_{ikjm} consists of three independent uniaxial invariants [12] and is highly symmetric, $\eta_{ikjm} = \eta_{kimj} = \eta_{jmik}$. The requirement that D is positive becomes

$$\eta_1 \geq 0, \quad \eta_3 \geq 0, \quad \eta_5 \geq 0, \quad \gamma_1 \geq 0. \tag{12}$$

The parameter λ is close to +1 or –1 for rodlike or disklike molecules, respectively. If the liquid is viscoisotropic, then $\lambda = 0$.

5. The hydrodynamic reactive (nondissipative) $\mathbf{m}^{(r)}$ and dissipative $\mathbf{m}^{(d)}$ fields are defined as

$$\begin{gathered} m_i^{(r)} = -(\mathbf{V}\cdot\nabla_3)n_i + n_k\Omega_{ki} + \lambda\xi_{ij}\Upsilon_{jk}n_k, \\ \mathbf{m}^{(d)} = \frac{1}{\gamma_1}\mathbf{F}, \end{gathered} \tag{13}$$

where ∇_3 is the three-dimensional gradient operator, $(\nabla_3)^2 = \Delta_3$.

6. The surface tension σ of a nematic LC is given by [28]

$$\sigma = \sigma_0 + \sigma_1\langle\mathbf{n}, \mathbf{e}\rangle^2, \tag{14}$$

where σ_0 and σ_1 are isotropic and anisotropic surface tension moduli respectively, and $\mathbf{e}$ is a unit vector normal to the LC surface.

7. Another system of viscous moduli α_i (called the Leslie viscosities) relate the dissipative and kinetic moduli as[4]

$$\begin{gathered} \eta_1 = \alpha_4/2, \quad \lambda = -\gamma_2/\gamma_1, \\ \eta_5 = \alpha_1 + \alpha_4 + \alpha_5 + \alpha_6, \\ \gamma_1 = \alpha_3 - \alpha_2, \quad \gamma_2 = \alpha_3 + \alpha_2, \\ \eta_3 - 2\eta_1 = \alpha_5 + \alpha_2\lambda, \\ 2\eta_1 + \eta_5 - 2\eta_3 = \alpha_1 + \gamma_2^2/\gamma_1, \end{gathered} \tag{15}$$

[4] The correct expression for η_5 is given in [18].

with the support of the Onsager–Parodi relation [17] as $\alpha_3 + \alpha_2 = \alpha_6 - \alpha_5$. In the vicinity of a phase transition, the viscous moduli α_i have different dependences on order parameter Q: $\alpha_1 \propto Q^2$, α_2, α_3, α_5, $\alpha_6 \propto Q$, and $\alpha_4 \propto Q^0$ [19].

Tables 1 and 2 (see above) summarize viscosities and other physical parameters that characterize the most frequently used and well-studied nematic LCs, also known as MBBA and PAA.

2.2. Basic Equations

The complete system of hydrodynamic equations for the isothermal incompressible nematic LC reflects the conservation laws of mass and of linear and angular momenta.

1. The continuity equation

$$\mathrm{div}\mathbf{V} = 0. \tag{16}$$

2. The Navier–Stokes equation for viscoelastic LC,

$$\rho\frac{\partial V_i}{\partial t} + \rho(\mathbf{V}\cdot\nabla_3)V_i = \frac{\partial}{\partial x_k}(T_{ik}^{(r)} + T_{ik}^{(d)}). \tag{17}$$

3. The Leslie–Ericksen equation of angular motion of the director $\mathbf{n}(\mathbf{r}, t)$,

$$\frac{\partial \mathbf{n}}{\partial t} = \mathbf{m}^{(r)} + \mathbf{m}^{(d)}. \tag{18}$$

The last equation is written for a negligible specific angular moment of inertia $\mathscr{L}_{LC}$ of the LC, namely, $\mathscr{L}_{LC} \ll \rho r_0^2$, where r_0 is the characteristic size of the system. This is true in our case, where r_0 is the radius of the jet.

We consider an isothermal incompressible jet flowing along the z axis out of a nozzle at velocity $\mathbf{V}$. The initial orientation of director $\mathbf{n}^0$ is assumed to be collinear with $\mathbf{V}$. Deviations from the initial values of the director and pressure are defined as $\mathbf{n}^1 = \mathbf{n} - \mathbf{n}^0$ and $P_1 = P - P_0$, respectively, where $P_0 = \sigma_0/r_0$ is the unperturbed pressure within the cylindrical jet. In the linear approximation, $|\mathbf{n}^1| \ll 1$, Eqs. (16)–(18) are simplified as

$$\begin{aligned}
&\mathrm{div}\mathbf{V} = 0, \quad \rho\frac{\partial V_i}{\partial t} = -\frac{\partial P_1}{\partial x_i} + \frac{\partial T_{ik}^{(d)}}{\partial x_k} \\
&+ \frac{1-\lambda}{2}n_i^0\mathrm{div}\mathbf{F} - \frac{1+\lambda}{2}(\mathbf{n}^0\cdot\nabla_3)F_i, \\
&\frac{\partial n_i^1}{\partial t} = n_k^0\Omega_{ki} + \lambda\xi_{ij}^0\Upsilon_{jk}n_k^0 + \frac{1}{\gamma_1}F_i, \\
&\xi_{ij}^0 = \delta_{ij} - n_i^0 n_j^0, \quad i, j, k = x, y, z.
\end{aligned} \tag{19}$$

Choosing $n_z^0 = 1$ gives $F_z = 0$, and hence,

$$\frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} = 0, \tag{20}$$

$$\begin{aligned}
\rho\frac{\partial V_x}{\partial t} &= -\frac{\partial P_1}{\partial x} + \left(\beta_1\Delta_2 + \beta_2\frac{\partial^2}{\partial z^2}\right)V_x \\
&+ (\beta_2 - \beta_1)\frac{\partial^2 V_z}{\partial x\partial z} - \frac{\lambda+1}{2}\frac{\partial F_x}{\partial z},
\end{aligned}$$

$$\begin{aligned}
\rho\frac{\partial V_y}{\partial t} &= -\frac{\partial P_1}{\partial y} + \left(\beta_1\Delta_2 + \beta_2\frac{\partial^2}{\partial z^2}\right)V_y \\
&+ (\beta_2 - \beta_1)\frac{\partial^2 V_z}{\partial y\partial z} - \frac{\lambda+1}{2}\frac{\partial F_y}{\partial z},
\end{aligned} \tag{21}$$

$$\begin{aligned}
\rho\frac{\partial V_z}{\partial t} &= -\frac{\partial P_1}{\partial z} + \left(\beta_2\Delta_2 + \beta_3\frac{\partial^2}{\partial z^2}\right)V_z \\
&- \frac{\lambda-1}{2}\left(\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y}\right),
\end{aligned}$$

$$\begin{aligned}
\frac{\partial n_x^1}{\partial t} &= \frac{\lambda+1}{2}\frac{\partial V_x}{\partial z} + \frac{\lambda-1}{2}\frac{\partial V_z}{\partial x} + \frac{F_x}{\gamma_1}, \\
\frac{\partial n_y^1}{\partial t} &= \frac{\lambda+1}{2}\frac{\partial V_y}{\partial z} + \frac{\lambda-1}{2}\frac{\partial V_z}{\partial y} + \frac{F_y}{\gamma_1},
\end{aligned} \tag{22}$$

$$\frac{\partial n_z^1}{\partial t} = 0,$$

where $\Delta_2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the two-dimensional Laplacian, $\beta_1 = \eta_1$, $\beta_2 = \eta_3/2$, $\beta_3 = \eta_5 - \eta_3/2$, and F_x and F_y are given in (6). Because isotropic viscosity implies that $\beta_i = \beta$, the LCs MBBA and PAA mentioned above are clearly far from isotropic (see Tables 1 and 2).

To make the problem more specific and easier to solve, we consider axisymmetric disturbances in the system of a cylindrical LC jet with radius r_0. In this case,

$$\frac{\partial V_z}{\partial z} + \frac{\partial V_r}{\partial r} + \frac{V_r}{r} = 0, \tag{23}$$

$$\begin{aligned}
\rho\frac{\partial V_r}{\partial t} &= -\frac{\partial P_1}{\partial r} + \left[\beta_1\left(\Delta_{2c} - \frac{1}{r^2}\right) + \beta_2\frac{\partial^2}{\partial z^2}\right]V_r \\
&+ (\beta_2 - \beta_1)\frac{\partial^2 V_z}{\partial r\partial z} - \mu_1\frac{\partial F_r}{\partial z},
\end{aligned} \tag{24}$$

$$\begin{aligned}
\rho\frac{\partial V_z}{\partial t} &= -\frac{\partial P_1}{\partial z} + \left[\beta_2\Delta_{2c} + \beta_3\frac{\partial^2}{\partial z^2}\right]V_z \\
&- \mu_2\left(\frac{\partial F_r}{\partial r} + \frac{F_r}{r}\right),
\end{aligned} \tag{25}$$

$$\gamma_1\frac{\partial n_r^1}{\partial t} = \gamma_1\mu_1\frac{\partial V_r}{\partial z} + \gamma_1\mu_2\frac{\partial V_z}{\partial r} + F_r, \quad n_z^1 = 0, \quad (26)$$

where

$$\Delta_{2c} = \frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r},$$

$$F_r = K\left(\Delta_{2c} - \frac{1}{r^2} + \frac{\partial^2}{\partial z^2}\right)n_r^1, \quad (27)$$

$$\mu_1 = \frac{\lambda+1}{2}, \quad \mu_2 = \frac{\lambda-1}{2}.$$

Equations (23)–(26) describe the ordinary linear hydrodynamic behavior of isotropic incompressible liquids if the LC properties vanish: K, $\gamma_1 \longrightarrow 0$ and $\beta_i = \beta$. The result is the well-known continuity equation and the linearized Navier–Stokes equation,

$$\mathrm{div}\mathbf{V} = 0, \quad \rho\frac{\partial \mathbf{V}}{\partial t} = -\nabla P_1 + \beta\Delta_3\mathbf{V}. \quad (28)$$

2.3. Boundary Conditions at the Free Surface

Boundary conditions at the free surface of an LC state that the jump in normal stress consists of two parts: one depends on the surface tension σ and the other on the elastic disturbance W_{elast} of the uniform director field $\mathbf{n}_0(\mathbf{r})$. Assuming that no tangential stresses exist at the free surface, we can express the boundary conditions at $r = r_0$ as

$$(T_{ik}^{(r)} + T_{ik}^{(d)})e_k + (2\sigma\mathcal{H} + W_{\mathrm{elast}})e_i + \frac{\partial\sigma}{\partial x_i} = 0, \quad (29)$$

where e_i are the components of the normal unit vector $\mathbf{e}$ in the reference frame of the LC cylinder and $\mathcal{H} = (1/R_1 + 1/R_2)/2$ denotes the mean surface curvature with the principal radii R_1 and R_2.

The nonhydrodynamic part of the boundary conditions with strong tangential anchoring of the director at the free surface holds if the scale of deformation of the initial surface is much larger than the molecular length of LCs.[5] This determines the tangential behavior of a smoothly disturbed director $\mathbf{n}$ at the free surface, $e_z \ll e_r \sim 1$:

$$\mathbf{e}\cdot\mathbf{n} = 0 \longrightarrow e_z + n_r^1 = 0 \text{ at } r = r_0. \quad (30)$$

The last constraint cancels the gradient term in Eq. (29). We finally obtain the boundary conditions in the linear approximation of the variables n_r^1, V_r, V_z, and P_1,

$$T_{rr}^{(r)} + T_{rr}^{(d)} + 2\sigma\mathcal{H} + W_{\mathrm{elast}} = 0, \quad T_{zr}^{(r)} + T_{zr}^{(d)} = 0. \quad (31)$$

Substitution of the expressions for the reactive and dissipative stress tensors gives

$$2\beta_1\Upsilon_{rr} - P_1 = 2\sigma_0(\mathcal{H}_0 - \mathcal{H}) - W_{\mathrm{elast}},$$
$$2\beta_2\Upsilon_{zr} = \mu_2 F_r, \text{ at } r = r_0, \quad (32)$$

where $\mathcal{H}_0 = (2r_0)^{-1}$ is the initial mean curvature of the LC cylinder. The equations for the jet surface—disturbed by a wave $\zeta(z, t)$—and its radial velocity $\partial\zeta/\partial t$ are given by

$$r(z, t) = r_0 + \zeta(z, t), \quad V_r = \frac{\partial\zeta}{\partial t} \text{ at } r = r_0, \quad (33)$$

where $\zeta \ll r_0$ is the radial displacement of a surface point. The principal radii of the surface curvature, in the linear approximation with respect to ζ, and its derivatives can be expressed as

$$\frac{1}{R_1} = \frac{1}{r_0+\zeta} \approx \frac{1}{r_0} - \frac{\zeta}{r_0^2}, \quad \frac{1}{R_2} \approx -\frac{\partial^2\zeta}{\partial z^2}. \quad (34)$$

This transforms the boundary conditions given by Eqs. (30) and (32) into

$$n_r^1 = \frac{\partial\zeta}{\partial z}, \quad V_r = \frac{\partial\zeta}{\partial t}, \quad (35)$$

$$2\beta_2\Upsilon_{zr} = \mu_2 F_r, \quad (36)$$

$$P_1 - 2\beta_1\Upsilon_{rr} = -\sigma_0\left(\frac{\zeta}{r_0^2} + \frac{\partial^2\zeta}{\partial z^2}\right) + W_{\mathrm{elast}}. \quad (37)$$

The term W_{elast} deserves further discussion. It reflects the existence of normal stresses at the surface, which arise due to the resistance of the uniformly orientated continuous LC media to the surface disturbance. The term W_{elast} vanishes in undisturbed LC jets and depends linearly on the elastic modulus K, the radius r_0, and the derivatives of ζ. Moreover, the invariance of the problem under inversion of the z axis requires dependence on derivatives of only even orders. An explicit expression for W_{elast} is derived in Section 3.1.

3. PLATEAU INSTABILITY IN AN LC CYLINDER

Before considering the sophisticated mathematics of Eqs. (23)–(26) supplemented by boundary conditions (35)–(37), we discuss capillary instability of the LC cylinder. This is done by applying the Plateau considerations [29] on the shape of a liquid mass withdrawn from the action of gravity.

[5] Strictly speaking, this assumption is correct when the equilibrium distribution of the director field $\mathbf{n}(\mathbf{r})$ is free of singularities. The problem of the minimal surface of an LC drop presents another situation where a substantial rearrangement of the field $\mathbf{n}(\mathbf{r})$ at the surface can decrease the total energy by destroying the disclination core within the drop.

We consider an LC cylinder with surface disturbed in accordance with (33), where $\zeta = \zeta_0 \cos kz$, ζ_0 is small compared to r_0, and $k = 2\pi/\Lambda$, with Λ being the disturbance wavelength. The idea of Plateau, applied here, is to find the cut-off wavelength Λ_s of the disturbance that determines breakage of the cylinder into droplets with due decrease in the total energy.

The average volume v over one wavelength Λ in the z direction is given by

$$v = \frac{1}{\Lambda}\int_0^\Lambda dz \int_s ds = \pi\left(r_0^2 + \frac{1}{2}\zeta_0^2\right) \longrightarrow r_0 = \sqrt{\frac{v}{\pi}}\left(1 - \frac{1}{4}\frac{\pi\zeta_0^2}{v}\right), \tag{38}$$

where r_0 in the right-hand side is given as a second-order expansion in ζ_0. The total energy $\mathscr{E}$ of the LC cylinder per unit wavelength with a disturbed director field $\mathbf{n}(\mathbf{r})$ is given by

$$\mathscr{E} = \sigma_0 \int_s ds + \frac{K}{2}\int_0^\Lambda dz \int_s (\mathrm{div}^2\mathbf{n} + \mathrm{rot}^2\mathbf{n})\, ds. \tag{39}$$

The static director field $\mathbf{n}(\mathbf{r})$ can be found from Eq. (27) and the associated boundary condition (35),

$$n_z^0 = 1, \quad F_r = 0 \longrightarrow \left(\Delta_{2c} - \frac{1}{r^2} + \frac{\partial^2}{\partial z^2}\right) n_r^1 = 0, \quad n_r^1 = \frac{\partial \zeta}{\partial z} \text{ at } r = r_0. \tag{40}$$

Equation (40) has the solution

$$n_r^1(r, z) = -\frac{k\zeta_0}{I_1(kr_0)} I_1(kr) \sin kz, \tag{41}$$

which is finite at $r = 0$, where $I_m(x)$ is the modified Bessel function of order m. The contribution of elastic forces is determined by

$$\mathrm{div}^2\mathbf{n} + \mathrm{rot}^2\mathbf{n} = k^2\left[\frac{k\zeta_0}{I_1(kr_0)}\right]^2 \times [A_1^2(kr)\sin^2 kz + A_2^2(kr)\cos^2 kz], \tag{42}$$

where

$$A_1(q) = \frac{dI_1(q)}{dq} + \frac{1}{q} I_1(q), \quad A_2(q) = I_1(q).$$

Simple integration of Eq. (39) gives

$$\mathscr{E} = 2\pi\sigma_0 r_0\left(1 + \frac{1}{4}k^2\zeta_0^2\right) + \frac{\pi}{2}K\left[\frac{k\zeta_0}{I_1(kr_0)}\right]^2 \times \int_0^{kr_0} [A_1^2(q) + A_2^2(q)]\, q\, dq. \tag{43}$$

Inserting r_0 from Eq. (38) into the first term above, we obtain

$$\mathscr{E} - 2\sigma_0\sqrt{\pi v} = \sigma_0\frac{\pi\zeta_0^2}{2r_0}(\varpi^2 - 1) + \frac{\pi}{2}K\left[\frac{\zeta_0\varpi}{r_0 I_1(\varpi)}\right] \times \int_0^{\varpi} [A_1^2(q) + A_2^2(q)]\, q\, dq, \quad \varpi = kr_0. \tag{44}$$

The positive root $\varpi_s = k_s r_0$ of the expression in the right-hand side of Eq. (44) determines the cut-off wavelength Λ_s of capillary disturbances that renders the LC cylinder unstable.

The quadratic approximation with respect to the derivatives $\partial\mathbf{n}/\partial x_j$ in Eq. (1), which provides the basis for the Frank theory, makes expression (44) correct only in terms of the ϖ^2 approximation. Indeed, the power of ϖ in Eq. (44) must not exceed 2, otherwise calculation becomes inconsistent. We thus obtain

$$\mathscr{E} - 2\sigma_0\sqrt{\pi v} = \sigma_0\frac{\pi\zeta_0^2}{2r_0}(\varpi^2 - 1) + \pi K k^2\zeta_0^2, \quad \varpi_s = \frac{1}{\sqrt{1 + 2\kappa}}, \quad \kappa = \frac{K}{\sigma_0 r_0}, \tag{45}$$

where the subscript s denotes the static nature of the Plateau instability. The asymptotic behavior of $\varpi_s(\kappa)$ shows two important limits:

$$\varpi_s = 1 - \kappa \text{ if } \kappa \ll 1, \quad \varpi_s = \frac{1}{\sqrt{2\kappa}}\left(1 - \frac{1}{4\kappa}\right) \text{ if } \kappa \gg 1. \tag{46}$$

Figure 1 shows a plot of $k_s r_0$ versus κ for the Plateau instabilities in the LC and in ordinary liquid.

The corresponding asymptotic cut-off wavelength Λ_s is obtained as

$$\Lambda_s = 2\pi r_0(1 + \kappa) \text{ if } \kappa \ll 1, \quad \Lambda_s = 2\pi\sqrt{\frac{2K}{\sigma_0}}\sqrt{r_0}\left(1 + \frac{1}{4\kappa}\right) \text{ if } \kappa \gg 1. \tag{47}$$

This result shows that $k \geq k_s$ increases the total energy $\mathscr{E}$ of the disturbed system, whereas $k < k_s$ decreases it. According to (46), there are two marginal regimes of instability.

1. The capillary regime $r_0 \gg K/\sigma_0$. Here, Λ_s is close to the circumference of the cylinder and the elastic deformation contribution $\int E_d dV$ to the total energy $\mathscr{E}$ is negligible. This regime should apply to a wide range of nematic LCs, because the common values of $K \sim 10^{-11}$ J/m [19] and $\sigma_0 \sim 10^{-2}$ J/m^2 [23] lead to $K/\sigma_0 \sim 10^{-9}$ m. This value is evidently smaller than the presently attainable radii of the jet.

2. The elastic regime $r_0 \ll K/\sigma_0$. This case reflects the dominance of elastic deformation and predicts an unusual behavior for $\Lambda_s \propto \sqrt{r_0}$. This regime cannot be reached by a simple increase of the elastic moduli, because their magnitude is determined by $K \sim \kappa_B T/a$, where $\kappa_B T \approx 4 \times 10^{-21}$ J is the Boltzmann thermal energy at room temperature and $a \approx 5 \times 10^{-10}$ m is the molecular length of the LC. In contrast, the effect of surface tension can be diminished by surfactants or by charging the surface of the liquid. In the latter case, the charge can virtually eliminate the effect of surface tension and provide conditions under which elastic forces predominate.

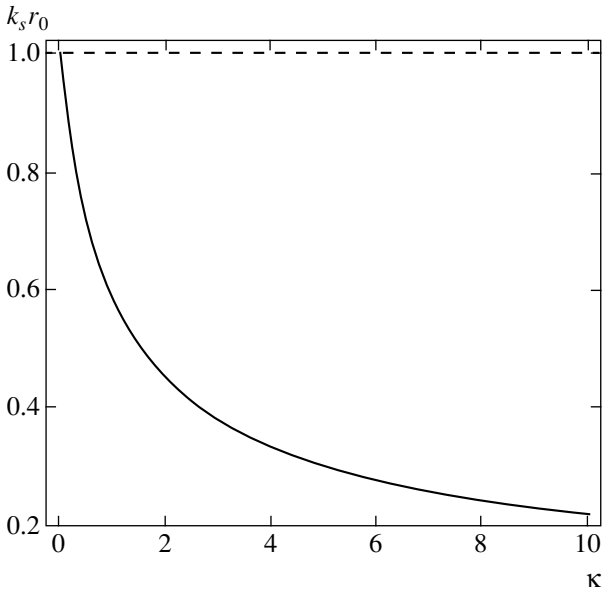


Fig. 1. Universal plots of $k_s r_0$ versus κ for the Plateau instabilities in an LC cylinder (solid line), and in ordinary liquid, $k_s r_0 = 1$ (dashed line).

3.1. W_{elast} and the Gaussian Surface Curvature

A straightforward way to derive an expression for W_{elast} is to solve the elastic problem for the stresses existing on a deformed axisymmetric surface of an LC cylinder. This is related to the Plateau instability, which obviates the need to repeat the entire procedure.

When we turn from Plateau considerations regarding the static instability of LC cylinders to the capillary instability of LC jets, the question is whether the cut-off wavelengths of the static (Λ_s) and hydrodynamic (Λ_d) problems coincide. This question was neglected by Rayleigh in his studies on isotropic viscous liquids, because the cut-off wavelengths always coincide for ordinary liquids, $\Lambda_s \equiv \Lambda_d$. This identity reflects a deep equivalence principle of the bifurcation point for a nontrivial steady state of a dynamic system and the threshold of static instability related to the minimum of its free energy $\mathscr{E}$ [4].

Using this $\Lambda_s \equiv \Lambda_d$, we construct the term W_{elast}, which enters boundary condition (37). For this, we examine and represent the total energy (45) as

$$\mathscr{E} - 2\sigma_0\sqrt{\pi V} = \frac{\pi\zeta_0 r_0}{2}\left[-\sigma_0\left(\frac{\zeta_0}{r_0^2} - \zeta_0 k^2\right) + 2K\frac{\zeta_0}{r_0}k^2\right]. \quad (48)$$

Next, we compare the expression in the brackets with the right-hand side of Eq. (37). This gives W_{elast}, which generates the elastic contribution in (48),

$$W_{\text{elast}} = 2K\mathscr{G}, \quad \mathscr{G} = \frac{1}{R_1 R_2} = -\frac{1}{r_0}\frac{\partial^2\zeta}{\partial z^2}, \quad (49)$$

where $\mathscr{G}$ is the Gaussian surface curvature in accordance with (34). Thus, the final expression for boundary conditions (29) is based on two fundamental invariants of the surface curvature, the mean surface curvature $\mathscr{H}$ and the Gaussian surface curvature $\mathscr{G}$.

4. DISPERSION RELATION

Rayleigh was the first to observe [1] that the instability problem is not so definite, contrary to the Plateau theory. The mode whereby a system deviates from unstable equilibrium must depend on the nature and characteristics of small displacements to which this system is subjected. In the absence of such displacement, any system, however unstable, cannot depart from equilibrium. These characteristics, being hydrodynamic, reflect the effect of viscosity, which predominates over inertia. For ordinary liquids, the mode of the maximum instability, which corresponds to the wavelength $\Lambda_R = 4.508 \times 2r_0$, exceeds the circumference of the liquid cylinder. We anticipate that the instability of LC jets has similar features.

The fact that the velocity potential does not exist in an anisotropic viscoelastic liquid dictates a standard approach to this problem that was first elaborated by Rayleigh [2]. We define the Stokes stream function $\Psi(\mathbf{r}, t)$ and the director potential $\Theta(\mathbf{r}, t)$ as

$$V_r = -\frac{1}{r}\frac{\partial\Psi}{\partial z}, \quad V_z = \frac{1}{r}\frac{\partial\Psi}{\partial r}, \quad n_r^1 = \frac{\partial\Theta}{\partial r}, \quad (50)$$

such that continuity equation (23) holds. From the other

three equations, (24)–(26), we have

$$\frac{\partial P_1}{\partial r} = (\beta_2 - \beta_1)\frac{\partial^2}{\partial r \partial z}\left(\frac{1}{r}\frac{\partial \Psi}{\partial r}\right) - \frac{1}{r} \times \frac{\partial}{\partial z}\left[\beta_1 r \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \Psi}{\partial r}\right) + \beta_2 \frac{\partial^2 \Psi}{\partial z^2} - \rho \frac{\partial \Psi}{\partial t} + \mu_1 r F_r\right], \quad (51)$$

$$\frac{\partial P_1}{\partial z} = \frac{1}{r} \times \frac{\partial}{\partial r}\left[\beta_2 r \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \Psi}{\partial r}\right) + \beta_3 \frac{\partial^2 \Psi}{\partial z^2} - \rho \frac{\partial \Psi}{\partial t} - \mu_2 r F_r\right], \quad (52)$$

$$\frac{\partial^2 \Theta}{\partial r \partial t} = \frac{1}{r}\left[\mu_2 r \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \Psi}{\partial r}\right) - \mu_1 \frac{\partial^2 \Psi}{\partial z^2}\right] + \frac{1}{\gamma_1} F_r,$$
$$F_r = K\left(\Delta_{2c} + \frac{\partial^2}{\partial z^2} - \frac{1}{r^2}\right)\frac{\partial \Theta}{\partial r}. \quad (53)$$

Applying the commutation rules gives

$$\left(\Delta_{2c} - \frac{1}{r^2}\right)\frac{\partial \Theta}{\partial r} = \frac{\partial}{\partial r}\Delta_{2c}\Theta \longrightarrow F_r = K\frac{\partial}{\partial r}\left(\Delta_{2c} + \frac{\partial^2}{\partial z^2}\right)\Theta,$$

which facilitates simplification of the above equations. Assuming that an axisymmetric disturbance characterized by the wavelength $2\pi/k$ increases exponentially in time with the growth rate s gives

$$\{\Psi, \Theta, \zeta, P_1, F_r\} = \{i\psi(r), i\theta(r), \varsigma(r), p(r), if(r)\}e^{st + ikz}. \quad (54)$$

Inserting (54) in (51)–(53) leads to the amplitude equations

$$\frac{1}{k}\frac{\partial p}{\partial r} = \beta_4 \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \psi}{\partial r}\right) - (\beta_2 k^2 + s\rho)\frac{\psi}{r} + \mu_1 f, \quad \beta_4 = 2\beta_1 - \beta_2, \quad (55)$$

$$kp = \frac{1}{r} \times \frac{\partial}{\partial r}\left\{r\left[\beta_2 \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \psi}{\partial r}\right) - (\beta_3 k^2 + s\rho)\frac{\psi}{r} - \mu_2 f\right]\right\}, \quad (56)$$

$$s\frac{\partial \theta}{\partial r} = \mu_2 \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \psi}{\partial r}\right) + \mu_1 k^2 \frac{\psi}{r} + \frac{1}{\gamma_1} f, \quad f = K\frac{\partial}{\partial r}(\Delta_{2c} - k^2)\theta. \quad (57)$$

The new variables in (54) require reformulating boundary conditions (35)–(37) as

$$k\varsigma = \frac{\partial \theta}{\partial r}, \quad s\varsigma = k\frac{\psi}{r},$$
$$\frac{\mu_2}{\beta_2} f = \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \psi}{\partial r}\right) + k^2 \frac{\psi}{r}, \quad (58)$$
$$p = 2\beta_1 k \frac{\partial}{\partial r}\left(\frac{\psi}{r}\right) + \varsigma\Gamma,$$

where

$$\Gamma = \sigma_0\left(k^2 - \frac{1}{r_0^2}\right) + 2K\frac{1}{r_0}k^2.$$

The real forms of amplitude equations (55)–(57) and boundary conditions (58) imply that expression (54) divides the five variables into two groups: P_1, ζ, and Ψ, Θ, F_r. These groups are shifted with respect to each other by the phase angle $\pi/2$.

4.1. Reduction of the Amplitude Equations

In this section, we perform the standard procedure for the simplification of amplitude equations (55)–(57). Substituting f from (57) in the other amplitude equations, we obtain

$$\frac{1}{k}\frac{\partial p}{\partial r} = B_1 \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \psi}{\partial r}\right) - (B_2 k^2 + s\rho)\frac{\psi}{r} + s\gamma_1\mu_1 \frac{\partial \theta}{\partial r}, \quad (59)$$

$$kp = \frac{1}{r}\frac{\partial}{\partial r}\left\{r\left[B_3 \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \psi}{\partial r}\right) - (B_4 k^2 + s\rho)\frac{\psi}{r}\right]\right\} - s\gamma_1\mu_2 \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \theta}{\partial r}\right), \quad (60)$$

$$0 = \mu_2 \frac{\partial}{\partial r}\left(\frac{1}{r}\frac{\partial \psi}{\partial r}\right) + \mu_1 k^2 \frac{\psi}{r} + \frac{K}{\gamma_1}\frac{\partial}{\partial r} \times \left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \theta}{\partial r}\right) - \left(k^2 + \frac{s\gamma_1}{K}\right)\theta\right], \quad (61)$$

where

$$B_1 = \beta_4 - \gamma_1\mu_1\mu_2, \quad B_2 = \beta_2 + \gamma_1\mu_1^2, \quad B_3 = \beta_2 + \gamma_1\mu_2^2, \quad B_4 = \beta_3 - \gamma_1\mu_1\mu_2, \quad (62)$$

and $B_2 > 0$, $B_3 > 0$ by virtue of conditions (12). Let a new stream function χ be defined as $\psi = r\partial\chi/\partial r$. The orientational ($\vartheta$) and kinematic ($\nu_i$) viscosities, as well

as the other auxiliary functions, are defined by the relations

$$\vartheta = \frac{K}{\gamma_1}, \quad \nu_i = \frac{B_i}{\rho}, \quad u_i^2 = k^2 + \frac{s}{\nu_i},$$
$$w^2 = k^2 + \frac{s}{\vartheta}, \quad \frac{\vartheta}{\nu_i} \ll 1 \longrightarrow u_i^2 \le w^2, \tag{63}$$

where the first inequality in (63) applies to the known nematic LC fluids (see Tables 1 and 2). Using the new notation, we find the first integrals of the amplitude equations,

$$\frac{p}{k} = (B_1\Delta_{2c} - B_2u_2^2)\chi + s\gamma_1\mu_1\theta, \tag{64}$$

$$kp = (B_3\Delta_{2c} - B_4u_4^2)\Delta_{2c}\chi - s\gamma_1\mu_2\Delta_{2c}\theta, \tag{65}$$

$$0 = (\mu_2\Delta_{2c} + \mu_1k^2)\chi + \vartheta(\Delta_{2c} - w^2)\theta. \tag{66}$$

Next, we eliminate the pressure amplitude p from Eqs. (64) and (65). This gives

$$[B_3\Delta_{2c}^2 - (B_1k^2 + B_4u_4^2)\Delta_{2c} + B_2u_2^2k^2]\chi - s\gamma_1(\mu_2\Delta_{2c} + \mu_1k^2)\theta = 0, \tag{67}$$

$$(\mu_2\Delta_{2c} + \mu_1k^2)\chi + \vartheta(\Delta_{2c} - w^2)\theta = 0. \tag{68}$$

Diagonalizing the matrix of operators in (67) and (68), we obtain homogeneous equations for the functions $\chi(r)$ and $\theta(r)$,

$$[D_3\Delta_{2c}^3 - D_2\Delta_{2c}^2 + D_1\Delta_{2c} - D_0]\begin{pmatrix}\chi\\ \theta\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix}, \tag{69}$$

where

$$D_0 = k^2(\vartheta B_2u_2^2w^2 - s\gamma_1\mu_1^2k^2),$$
$$D_1 = \vartheta(B_1k^2w^2 + B_2k^2u_2^2 + B_4w^2u_4^2) + 2s\gamma_1\mu_1\mu_2k^2,$$
$$D_2 = \vartheta(B_1k^2 + B_3w^2 + B_4u_4^2) - s\gamma_1\mu_2^2, \tag{70}$$
$$D_3 = \vartheta B_3.$$

It is easy to verify that all coefficients D_j are positive if the conditions $B_i > 0$ and $\mu_2 \ll 1$, $\vartheta/\nu_i \ll 1$ are satisfied (for all i). The latter conditions are in good agreement with numerous observations in nematic LCs [19].

Factoring the polynomial differential operator further (recalling that $D_3 > 0$) gives

$$D_3\Delta_{2c}^3 - D_2\Delta_{2c}^2 + D_1\Delta_{2c} - D_0 = D_3(\Delta_{2c} - m_1^2)(\Delta_{2c} - m_2^2)(\Delta_{2c} - m_3^2). \tag{71}$$

Equation (71) facilitates finding the finite solutions of Eq. (69),

$$\chi(r) = \sum_{j=1}^{3}\frac{C_j}{m_j}I_0(m_jr),$$
$$\theta(r) = \sum_{j=1}^{3}\frac{G_j}{m_j}I_0(m_jr), \tag{72}$$

where the second fundamental solutions that diverge at $r = 0$ have been excluded, C_j and G_j are indeterminate coefficients, and m_j^2 are three generic[6] roots of the cubic equation

$$D_3m^6 - D_2m^4 + D_1m^2 - D_0 = 0 \longrightarrow \sum_{j=1}^{3}m_j^2 = \frac{D_2}{D_3}, \tag{73}$$
$$\sum_{j\ne k}^{3}m_j^2m_k^2 = \frac{D_1}{D_3}, \quad \prod_{j=1}^{3}m_j^2 = \frac{D_0}{D_3}.$$

The coefficients G_j can be expressed via C_j after inserting solutions (72) into Eq. (68):

$$G_j = \frac{1}{\vartheta}g_jC_j,$$
$$g_j = \frac{\mu_1k^2 + \mu_2m_j^2}{w^2 - m_j^2}, \quad j = 1, 2, 3. \tag{74}$$

The amplitude of the pressure $p(r)$, the stream function $\psi(r)$, and the displacement $\varsigma(r_0)$ of a point on the surface are easily found from Eqs. (57), (64), (68), and (74) as

$$p(r) = k\sum_{j=1}^{3}\frac{l_i}{m_j}C_jI_0(m_jr),$$
$$l_j = B_1m_j^2 - B_2u_2^2 + \frac{s}{\vartheta}\gamma_1\mu_1g_j,$$
$$\psi(r) = r\sum_{j=1}^{3}C_jI_1(m_jr), \tag{75}$$
$$\varsigma(r_0) = \frac{1}{\vartheta k}\sum_{j=1}^{3}g_jC_jI_1(m_jr_0), \quad j = 1, 2, 3.$$

[6] The freedom to choose the physical parameters of the LC seems to admit a degeneration of cubic equation (73), when some of the roots m_j^2 can coincide in different ways. This coincidence is not important because it can occur only at specific wave vectors k^* on which the coefficients D_2, D_1, and D_0 depend. On the other hand, this kind of degeneration might be interesting if k^* is accidentally close to the cut-off wave vector k_d at which the breakage of the LC jet develops.

Before proceeding, we discuss the distribution of the roots m_j^2 of cubic equation (73) in the complex plane.

First, m_1^2 is always positive because $D_j > 0$, as mentioned above and as follows from the Descartes rule of sign interchange in the sequence of coefficients for real algebraic equations. The other two roots $m_{2,3}^2$ are either positive or complex conjugate with positive real parts. The last case leads to Bessel functions of complex arguments in (72). This fact can indicate that the separation of the two groups of functions P_1, ζ and Ψ, Θ, F_r by the phase angle $\pi/2$ is more elaborate than assumed in (54). Another consequence of the existence of complex conjugate roots m_j^2, which is more important from the physical standpoint, is the appearance of imaginary contributions to the dispersion equation. This can lead to a complex value of the growth rate $s = \bar{s} + i\omega$ as its solution and to the nonsteady (oscillatory) evolution of the jet, e.g.,

$$\zeta(z, t) \propto \varsigma(r_0)e^{\bar{s}t}e^{i(\omega t + kz)},$$

where ω is the frequency of oscillations.

4.2. Dispersion Equation

In what follows, we derive the dispersion equation $s = s(kr_0)$ that determines the evolution of the Rayleigh instability in LC jets. The revised version of boundary conditions (58) at $r = r_0$, which utilizes the new stream function $\chi(r)$, is given by

$$s\frac{\partial\theta}{\partial r} = k^2\frac{\partial\chi}{\partial r},$$

$$s\gamma_1\mu_2\frac{\partial\theta}{\partial r} = B_3\frac{\partial}{\partial r}\Delta_{2c}\chi + B_5k^2\frac{\partial\chi}{\partial r}, \qquad (76)$$

$$\frac{s}{k}p = 2s\beta_1\frac{\partial^2\chi}{\partial r^2} + \Gamma\frac{\partial\chi}{\partial r},$$

where $B_5 = \beta_2 + \gamma_1\mu_1\mu_2$. Substituting (72) and (75) into (76) and eliminating the coefficients C_1, C_2, and C_3 from the linear equations leads to a (3×3)-determinant equation:

$$\det S_{ij} = 0, \qquad (77)$$

where

$$S_{1j} = k^2 - \frac{s}{\vartheta}g_j,$$

$$S_{2j} = B_3m_j^2 + B_5k^2 - \frac{s}{\vartheta}\gamma_1\mu_2g_j, \qquad (78)$$

$$S_{3j} = \Gamma - s\left[\frac{l_j}{m_j}\frac{I_0(m_jr_0)}{I_1(m_jr_0)} - 2\beta_1m_j\frac{I_1'(m_jr_0)}{I_1(m_jr_0)}\right],$$

and $I_1'(y) = dI_1(y)/dy$. Equation (77) is an implicit form of the exact dispersion relation, which is highly complex and cannot be solved analytically in the general case. Nevertheless, here we can verify that the cut-off wavelength Λ_d coincides with Λ_s obtained from the Plateau theory. Indeed, the cut-off regime corresponds to boundary conditions (76) when $s = 0$ and is satisfied for $\Gamma = 0$, i.e., $\Lambda_d = \Lambda_s$. The implications of Eq. (77) can be extended further, for the study of different modes of the LC flow, including oscillations, and in order to describe the asymptotic behavior of LC jets. This is outside the scope of this paper. In the next section, we consider a case that facilitates decoupling of hydrodynamic and orientational modes, and consequently the solution of the Rayleigh instability problem in a closed form.

5. DECOUPLING OF HYDRODYNAMIC AND ORIENTATIONAL MODES

In this section, we discuss the case in which dispersion equation (77) becomes solvable. Here, we encounter another problem: the elasticity of the LC and anisotropy of its viscous properties have the same origin and therefore cannot be considered separately. Nevertheless, we investigate the case where dispersion equation (77) can be simplified. The large number of physical parameters involved (three viscous moduli, two kinetic coefficients, λ and γ_1, orientational (ϑ) and kinematic (ν_i) viscosities, and the dimensionless parameter κ) call for such a treatment.

We consider an LC with rodlike molecules ($\lambda \approx 1$) and low orientational viscosity ϑ

$$\mu_1 \approx 1, \quad \mu_2 \approx 0, \quad \vartheta \ll \nu_i, \quad k^2 \ll \frac{s}{\vartheta}, \qquad (79)$$

where the first three relations apply to known nematic LC fluids (see Tables 1 and 2). The last inequality in (79) applies to the low-viscosity limit, which was considered for the kinematic viscosity in ordinary liquids by Rayleigh [1].

In this case, characteristic equation (73) reduces to

$$m^6 - \frac{s}{\vartheta}m^4 + \frac{s}{\vartheta}\left(\mathcal{B}k^2 + \frac{s}{\bar{\nu}_2}\right)m^2 - \frac{s}{\vartheta}k^2\left(k^2 + \frac{s}{\bar{\nu}_2}\right) = 0, \qquad (80)$$

$$\bar{\nu}_i = \frac{\beta_i}{\rho}, \quad \mathcal{B} = \frac{\beta_3 + \beta_4}{\beta_2}.$$

The three roots m_j^2 of Eq. (73) become

$$2m_{1,2}^2 = \mathcal{B}k^2 + \frac{s}{\bar{\nu}_2} \pm \sqrt{(\mathcal{B}^2-4)k^4 + 2(\mathcal{B}-2)k^2\frac{s}{\bar{\nu}_2} + \left(\frac{s}{\bar{\nu}_2}\right)^2}, \qquad (81)$$

$$m_3^2 = \frac{s}{\vartheta}.$$

A simple analysis of the last expression shows that the dimensionless parameter $\mathcal{B}$ has a critical value of 2 that separates two different evolution scenarios of the LC jet. If $\mathcal{B} > 2$, both roots, m_1^2 and m_2^2, are positive and the capillary instability always appears via trivial bifurcation (steady-state instability). This scenario applies to MBBA and PAA LCs with $\mathcal{B}_{MBBA} = 5.92$ and $\mathcal{B}_{PAA} = 7.11$ (see Tables 1 and 2). In the opposite case, $\mathcal{B} < 2$, we can find the regime where the above roots are complex conjugate. This leads to the oscillatory evolution of the jet, which appears via the Hopf bifurcation (see Section 4.1).

Significant simplification can be obtained if we assume degeneration of the three viscosities at the critical value $\mathcal{B}_* = 2$. Indeed, if the viscous moduli β_j satisfy the relation

$$\mathcal{B}_*(\beta_j) = 2 \longrightarrow 2\beta_1 + \beta_3 = 3\beta_2, \qquad (82)$$

the three roots m_j^2 of Eq. (73) are

$$m_{1*}^2 = k^2, \quad m_{2*}^2 = k^2 + \frac{s}{\bar{\nu}_2}, \quad m_3^2 = \frac{s}{\vartheta}. \qquad (83)$$

We note that relation (82) cancels the last term in (9). Expressions (83) indicate that the problem is decomposed into two parts, or, in other words, the cross terms in Eqs. (67) and (68) are dropped. Thus, the first part of the problem is associated with the Rayleigh instability, described by

$$(\Delta_{2c} - m_{1*}^2)(\Delta_{2c} - m_{2*}^2)\chi = 0, \qquad (84)$$

with boundary conditions that account for elasticity,

$$\frac{\partial}{\partial r}\Delta_{2c}\chi + k^2\frac{\partial\chi}{\partial r} = 0,$$
$$\frac{s}{k}p = 2s\beta_1\frac{\partial^2\chi}{\partial r^2} + \Gamma\frac{\partial\chi}{\partial r} \text{ at } r = r_0. \qquad (85)$$

The second part is associated with an orientational instability of the director field $\mathbf{n}(\mathbf{r}, t)$,

$$(\Delta_{2c} - m_3^2)\theta = 0,$$

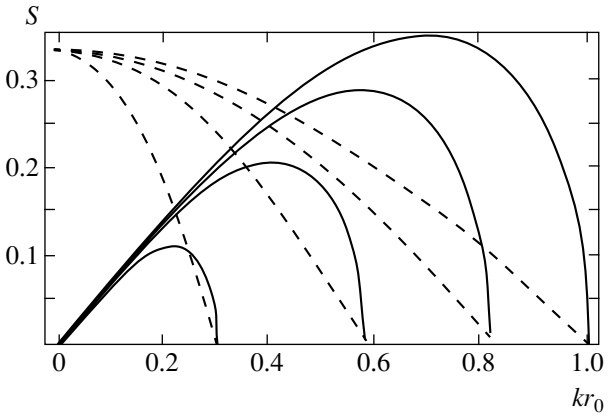


Fig. 2. A plot of the rescaled growth rate S versus kr_0 for low viscosity $\sqrt{\rho r_0^3/\sigma_0}\, s_-(kr_0)$ (solid curves) and high viscosity $2\beta_2 r_0/\sigma_0\, s_+(kr_0)$ (dashed curves) for different values of κ in descending order from top down, $\kappa = 0, 0.25, 1, 5$. If $\vartheta/\nu = 4\kappa$, then the scaling for both viscous regimes is the same.

with the boundary condition

$$s\frac{\partial\theta}{\partial r} = k^2\frac{\partial\chi}{\partial r} \text{ at } r = r_0. \qquad (86)$$

The solutions of Eqs. (84) and (86) are

$$\chi(r) = \frac{c_1}{m_{1*}}I_0(m_{1*}r) + \frac{c_2}{m_{2*}}I_0(m_{2*}r),$$
$$\theta(r) = \frac{c_3}{m_3}I_0(m_3 r). \qquad (87)$$

With these solutions, the hydrodynamic pressure $p(r)$, stream function $\psi(r)$, and surface displacement $\varsigma(r_0)$ are obtained as

$$p(r) = -c_1 s\rho I_0(m_{1*}r),$$
$$\psi(r) = r[c_1 I_1(m_{1*}r) + c_2 I_1(m_{2*}r)],$$
$$\varsigma(r_0) = \frac{c_3}{k}I_1(m_3 r_0),$$

where the only indeterminates are c_1 and c_2, while c_3 can be expressed as their linear combination,

$$c_3\frac{s}{k^2} = c_1\frac{I_1(m_{1*}r_0)}{I_1(m_3 r_0)} + c_2\frac{I_1(m_{2*}r_0)}{I_1(m_3 r_0)}, \qquad (88)$$

provided that $s = s(kr_0)$ satisfies the dispersion relation that follows from (85) and (87),

$$s^2 + \frac{2\bar{\nu}_1 k^2}{I_0(kr_0)}$$

$$\times\left[I_1'(kr_0)-\frac{2km_{2*}}{k^2+m_{2*}^2}\frac{I_1(kr_0)}{I_1(m_{2*}r_0)}I_1'(m_{2*}r_0)\right]s \quad (89)$$

$$=\frac{\sigma_0 k}{\rho r_0^2}[1-k^2r_0^2(1+2\kappa)]\frac{I_1(kr_0)}{I_0(kr_0)}\frac{m_{2*}^2-k^2}{m_{2*}^2+k^2}.$$

If $\kappa = 0$ and $\bar{\nu}_1 = \bar{\nu}_2$, Eq. (89) is known as the Weber equation for a viscous isotropic liquid [6]. For low viscosity,[7] $\beta_1 \sim \beta_2 \ll \sqrt{\rho\sigma_0 r_0}$, a Rayleigh-type expression is obtained (see Fig. 2),

$$s_-^2(kr_0) = \frac{\sigma_0 k}{\rho r_0^2}[1-k^2r_0^2(1+2\kappa)]\frac{I_1(kr_0)}{I_0(kr_0)}, \quad (90)$$

where the subscript "–" denotes low viscosity.

The maximum $s_-^{\max}$ in Eq. (90), which corresponds to the wave number $k_-^{\max}$, leads to evolution of the largest capillary instability. Numerical calculations show that $s_-^{\max}$ and $k_-^{\max}$ are both proportional to $(1+2\kappa)^{-1/2}$:

$$s_-^{\max}\approx\frac{1}{3\sqrt{1+2\kappa}}\sqrt{\frac{\sigma_0}{\rho r_0^3}},$$
$$k_-^{\max}\approx\frac{a}{r_0\sqrt{1+2\kappa}}, \quad a = 0.697. \quad (91)$$

When high viscosity prevails, $\beta_1 \sim \beta_2 \gg \sqrt{\rho\sigma_0 r_0}$, the dispersion equation is given by (see Fig. 2)

$$s_+(kr_0) = \frac{\sigma_0}{2\beta_2 r_0^2 k}$$
$$\times\frac{[1-k^2r_0^2(1+2\kappa)]I_1^2(kr_0)}{I_0(kr_0)I_1(kr_0)+kr_0[I_1'(kr_0)]^2}, \quad (92)$$
$$s_+^{\max}\approx\frac{\sigma_0}{6\beta_2 r_0}, \quad k_+^{\max} = 0,$$

where the subscript "+" denotes high viscosity. In this limit, similar to ordinary liquids [4], there is no finite mode of the maximum instability for any κ. In this case, we have

$$\varsigma(r_0)$$
$$=\frac{k_+^{\max}}{s_+^{\max}}[c_1I_1(k_+^{\max}r_0)+c_2I_1(m_{2*}r_0)] = 0. \quad (93)$$

[7] In the theory of viscoisotropic liquid jets, this case is known [7] to pertain to the range of low Ohnesorge numbers $\mathrm{Oh} = \eta/\sqrt{\rho\sigma_0 r_0}$ that determine the competition between the hydrodynamic and surface tension forces. Expression (92) corresponds to the case of high Ohnesorge numbers.

Nevertheless, there exists a continuous range $[0, (1+2\kappa)^{-1/2}r_0^{-1}]$ of wave numbers k with a finite disturbance growth rate $s_+(kr_0)$, which affect the cylindrical jet.

We note that the dispersion curves shown in Fig. 2 and those in Fig. 5 in [21] appear to be similar, but are characterized by different physical parameters. The reason for this observation is the similarity between Weber equation (89) and dispersion equation (36) in [21], which are obtained from different models. Our approach was to develop a general axisymmetric solution in the framework of the three-dimensional model. This model dates back to the Rayleigh–Weber theory [2, 3] and accounts for the radial inhomogeneity of the disturbed director field. The implicit solutions of Eq. (77) reflect the radial dependence of both the hydrodynamic $\mathbf{V}(r, z, t)$ and orientational $\mathbf{n}(r, z, t)$ modes, and they include all types of LC jet evolution. A specific case where the hydrodynamic and orientational modes are decoupled exhibits this radial dependence and yields dispersion equation (89) in explicit form.

In contrast, the one-dimensional analysis of the LC jet evolution, used in [21], is hardly compatible with the distortion of the director field $\mathbf{n}(r, z, t)$, and therefore must be supported by assuming a fixed axial direction of $\mathbf{n}^0$ (see detailed comments in Section 1). This endows their model with an inherent "decoupling" that results from the a priori elimination of elastic forces. Obviously, the similarity between the above-mentioned dispersion curves disappears if we consider the general solution given by (77).

5.1. Hydrodynamic Influence on the Orientational Instability of LCs

We conclude this section with a brief discussion regarding the hydrodynamic influence on the orientational instability of the director field $\mathbf{n}(\mathbf{r}, t)$. As the effect of hydrodynamics changes the wave number k_s of the Plateau instability to $k_{\max}$, the flow drives the orientational instability (41) of the director field $\mathbf{n}(\mathbf{r}, t)$. Indeed, in accordance with (87),

$$n_r^1(r,z) = c_3I_1(m_3^{\max}r), \quad m_3^{\max} = \sqrt{\frac{s^{\max}}{\vartheta}}. \quad (94)$$

It is convenient to consider the following two marginal viscous regimes.

1. The low-viscosity limit,

$$(m_{3-}^{\max}r_0)^2\approx\frac{1}{3\sqrt{1+2\kappa}}\frac{1}{\sqrt{\kappa\varepsilon}}, \quad \varepsilon = \frac{\rho K}{\gamma_1^2}, \quad (95)$$

where $\varepsilon \sim 10^{-6}$–10^{-4} is a small dimensionless parameter.

2. The high-viscosity limit,

$$(w_{3+}^{\max}r_0)^2\approx\frac{1}{6\kappa}\frac{\gamma_1}{\beta_2}. \quad (96)$$

In both limits, the distribution of the director field $\mathbf{n}(\mathbf{r}, t)$ in the jet is always nontrivial and definitely far from static distribution (41).

6. CONCLUSIONS

1. The capillary instability of an LC jet with a strong tangential anchoring of the director at the surface is considered within the linear hydrodynamics of the uniaxial nematic LC. Its static version, which is called the Plateau instability and corresponds to the variational problem of minimal free energy, predicts a substantial dependence of the disturbance cut-off wavelength on the dimensionless parameter $\kappa = K/\sigma_0 r_0$.

2. The hydrodynamic problem of the capillary instability in LC jets is solved exactly and the dispersion relation is derived. This relation, which is represented as a determinant equation, implicitly expresses the dispersion $s = s(k)$ of the growth rate s as a function of the wave number k of axisymmetric disturbances of the jet.

3. The case where the dispersion equation becomes explicitly solvable is considered in detail. It corresponds to the regime in which the hydrodynamic and orientational modes become decoupled. The hydrodynamics changes the wave number k_s of the Plateau instability into $k_{\max}$, which produces evolution of the largest capillary instability. Similarly, a hydrodynamic flow influences the static orientational instability of the director field $\mathbf{n}(\mathbf{r}, t)$.

4. The present theory can easily be extended to nonuniaxial nematic LCs that possess finite point symmetry groups $G \subset O(3)$ as distinguished from the uniaxial group $D_{\infty h}$ The corresponding expressions for the free energy density $E_d(G)$ and the dissipative function $D(G)$ were derived in [31].

5. In this work, the effect of external fields was not considered. However, the theory developed here facilitates the treatment of the Rayleigh instability in nematic LCs in the presence of static electromagnetic fields.

ACKNOWLEDGMENTS

The research was supported by the Gileadi Fellowship program of the Ministry of Absorption of the State of Israel. The useful comments of E.I. Kats are hereby acknowledged.

REFERENCES

1. Lord Rayleigh, Proc. London Math. Soc. **10**, 4 (1879).
2. Lord Rayleigh, Philos. Mag. **34**, 145 (1892); **34**, 177 (1892).
3. C. Weber, Z. Angew. Math. Mech. **11**, 136 (1931).
4. S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Instability* (Oxford Univ. Press, Oxford, 1961).
5. G. I. Taylor, Proc. R. Soc. London, Ser. A **253**, 289 (1959); **253**, 296 (1959); **253**, 313 (1959); Proc. R. Soc. London, Ser. A **313**, 453 (1969).
6. V. G. Levich, *Physicochemical Hydrodynamics* (Fizmatgiz, Moscow, 1959; Prentice-Hall, Englewood Cliffs, N.J., 1962).
7. A. L. Yarin, *Free Liquid Jets and Films: Hydrodynamics and Rheology* (Wiley, New York, 1993).
8. Y. Tomita, T. Shimbo, and Y. Ishibashi, J. Non-Newtonian Fluid Mech. **5**, 497 (1979).
9. R. Bar-Ziv and E. Moses, Phys. Rev. Lett. **73**, 1392 (1994).
10. K. L. Gurin, V. V. Lebedev, and A. R. Muratov, Zh. Éksp. Teor. Fiz. **110**, 600 (1996) [JETP **83**, 321 (1996)].
11. W. Helfrich, Z. Naturforsch. B **103**, 67 (1975).
12. E. I. Kats and V. V. Lebedev, *Fluctuational Effects in the Dynamics of Liquid Crystals* (Springer, New York, 1994).
13. J. L. Ericksen, Arch. Ration. Mech. Anal. **4**, 231 (1960).
14. J. L. Ericksen, Phys. Fluids **9**, 1205 (1966).
15. F. M. Leslie, Q. J. Mech. Appl. Math. **19**, 357 (1966).
16. F. M. Leslie, Arch. Ration. Mech. Anal. **28**, 265 (1968).
17. O. Parodi, J. Phys. (Paris) **31**, 581 (1970).
18. D. Forster, T. C. Lubensky, P. C. Martin, *et al.*, Phys. Rev. Lett. **26**, 1016 (1971).
19. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, 2nd ed. (Oxford Univ. Press, London, 1993; Mir, Moscow, 1982).
20. L. G. Fel and G. Lasiene, Acta Phys. Pol. A **70**, 165 (1986).
21. A.-G. Cheong, A. D. Rey, and P. T. Mather, Phys. Rev. E **64**, 41701 (2001).
22. G. E. Durand and E. G. Virga, Phys. Rev. E **59**, 4137 (1999).
23. J. Cognard, *Alignment of Nematic Liquid Crystals and Their Mixtures* (Gordon and Breach, London, 1982).
24. A. Fergusson and S. J. Kennedy, Philos. Mag. **26**, 41 (1938).
25. D. Langevin, J. Phys. (Paris) **33**, 249 (1972).
26. S. Krishnaswamy and R. Shashidhar, in *Proceedings of International Liquid Crystal Conference* (Bangalore, 1973), Pramana Suppl., Vol. 1, p. 247.
27. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity*, 4th ed. (Nauka, Moscow, 1987; Pergamon Press, Oxford, 1986).
28. A. Rapini and M. Papoular, J. Phys. Colloq. **30**, C4-54 (1971).
29. J. Plateau, *Statique Expérimentale et Théorique des Liquides Soumis aux Seules Forces Moléculaires* (Gauthier-Villars, Paris, 1873).
30. W. H. de Jeu, *Physical Properties of Liquid Crystalline Materials* (Gordon and Breach, London, 1980).
31. L. G. Fel, Kristallografiya **34**, 1222 (1989) [Sov. Phys. Crystall. **34**, 737 (1989)]; Mol. Cryst. Liq. Cryst. **206**, 1 (1991).

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 974–980.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1115–1121.

FLUIDS

Superfluid ^{3}He Phases in Aerogel

I. A. Fomin

Kapitza Institute for Physical Problems, Russian Academy of Sciences,
ul. Kosygina 2, Moscow, 119334 Russia
e-mail: fomin@kapitza.ras.ru

Received December 1, 2003

Abstract—A criterion for selecting the order parameter in the immediate vicinity of the transition temperature is derived within the framework of a phenomenological description of superfluid ^{3}He in aerogel. The order parameter of the BW phase of pure ^{3}He satisfies this criterion, whereas that of the ABM phase does not. A class of order parameters that could be used to describe the properties of the A-like phase of ^{3}He observed in aerogel was found. The influence of a magnetic field on the order parameters from this class is considered. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Superfluid ^{3}He phases are the most thoroughly studied example of unconventional Cooper pairing. Pairing is considered unconventional if, in addition to gauge symmetry, other symmetries of the normal phase are broken by the pairing-induced phase transition. The introduction of aerogel into liquid ^{3}He makes it possible to use ^{3}He for studying the influence of impurities on unconventional Cooper pairing [1]. These studies may be of interest for understanding the properties of metallic superconductors with unconventional Cooper pairing, such as UPt_3, UBe_{13}, $Sr_2Ru_2O_4$, UGe_2, etc., which unavoidably contain impurities. Aerogel can be thought of as a rigid frame formed by strands approximately 30 Å thick. The estimated mean distance between aerogel strands is 200 Å; this is close to the correlation length ξ_0 in superfluid ^{3}He, which changes in the range 160–500 Å depending on pressure. Like usual impurities, aerogel limits the mean free path l of Fermi quasi-particles in liquid ^{3}He. According to estimates, $l \sim 1500$–1800 Å for 98%-porous aerogel, which fills less than 2% of its total volume. This length is large compared with ξ_0. According to the theory of superconducting alloys [2], impurities decrease the superfluid transition temperature T_c proportionally to ξ_0/l [3]. Two superfluid phases are observed in the ^{3}He + aerogel system below T_c [4]. By analogy with pure (free of aerogel) ^{3}He, these phases are called A- and B-like. This implies correspondence, which is exact for the B-like phase. The observation of a uniformly precessing domain in it [5] shows that the difference between its order parameter and that of the BW phase is small if any. The observed properties of the A-like phase differ more substantially from those of the A phase of pure ^{3}He. To identify this phase, we must answer the question of whether or not aerogel can influence the form of the order parameter, and if it can, then what phases are admissible. The purpose of this work was to answer this question. Next, we will formulate a procedure for determining the order parameters of the superfluid ^{3}He phases in aerogel in the vicinity of the transition point T_c. This procedure will be applied to the A-like phase.

2. AEROGEL INTERACTION WITH THE ORDER PARAMETER

Under the conditions of Cooper pairing with $l \neq 0$, we must expect effects that arise because of fluctuations in the arrangement of aerogel strands in addition to the already mentioned overall decrease in T_c determined by the mean free path. Near T_c, these effects can be described phenomenologically on the assumption that gel creates a random field that acts on the order parameter. The corresponding change in the free energy of superfluid phases is found from symmetry considerations and supposed aerogel properties. Pairing occurs in ^{3}He with orbital momentum $l = 1$ and spin $s = 1$. The corresponding order parameter is a complex 3×3 matrix $A_{\mu j}$, where μ and j are the spin and orbital indices, respectively. The interaction of ^{3}He with aerogel strands arises because of the scattering of quasi-particles on them. Scattering changes quasi-particle momenta, and this is the mechanism of aerogel interactions with the orbital part of $A_{\mu j}$. Aerogel can also interact with the spin part of the order parameter. The aerogel material (SiO_2) is nonmagnetic, but when aerogel is immersed into liquid ^{3}He, aerogel strands become covered with a layer of localized ^{3}He atoms that can participate in spin exchange with scattered quasi-particles. To "switch off" interactions with spin, an admixture of ^{4}He is added into the cell for measurements. The atoms that are first deposited on the strands are then ^{4}He, and, if present in a sufficient concentration, they fully replace

localized ^{3}He atoms. It follows that ^{3}He in aerogel with and without an admixture of ^{4}He can have different properties. Its magnetic properties should differ the most strongly. Everywhere below, we assume that aerogel strands are covered by a layer of ^{4}He. Aerogel then only acts on the orbital part of $A_{\mu j}$, and, in the principal order in $A_{\mu j}$, the corresponding free energy increment can be written in the form [6]

$$F_\eta = N(0)\int \eta_{jl}(\mathbf{r})A_{\mu j}A^*_{\mu l}d^3r, \quad (1)$$

where $N(0)$ is the density of states at the Fermi boundary and $\eta_{jl}(\mathbf{r})$ is the random static tensor field. Because of the $t \longrightarrow -t$ invariance, the $\eta_{jl}(\mathbf{r})$ tensor is real and symmetrical, and its isotropic part $(1/3)\eta_{ll}(\mathbf{r})\delta_{jl}$ describes the local change in $T_c = T_c(\mathbf{r})$ caused by scatterer density fluctuations. The anisotropic part

$$\eta_{jl}(\mathbf{r}) - \frac{1}{3}\eta_{ll}(\mathbf{r})\delta_{jl} \equiv \eta^{(a)}_{jl}$$

describes the local splitting of T_c caused by spherical symmetry violation by gel strands. The isotropic random field part will further be considered included into $T_c = T_c(\mathbf{r})$. The results obtained in [7] allow the random field to be estimated as

$$|\eta_{jl}| \sim x\xi_0/R \sim \xi_0/l,$$

where l is the mean free path, R is the strand radius, and x is the volume fraction occupied by aerogel. This estimate gives the correct order of magnitude. For 98%-porous aerogel, $\xi_0/l \sim 1/10$. The spatial scale of field $\eta_{jl}(\mathbf{r})$ changes is the distance between strands $d \sim R/\sqrt{x}$; this distance is comparable to ξ_0 in 98%-porous aerogel. If the order parameter is deformed on a scale of the order of d, the relative gradient energy loss is of about $(\xi_0/d)^2$ and larger than the energy gain from the interaction with the field $\eta_{jl}(\mathbf{r})$ by a factor proportional to $\xi_0/R \gg 1$. It is unfavorable for the order parameter to follow field changes or form states localized on a scale of the order of d. The weakness of field $\eta_{jl}(\mathbf{r})$ does not rule out the possibility of formation of localized states on a scale of $L \gg d$. This possibility exists because the mean order parameter $\bar{A}_{\mu j}$ is degenerate with respect to orbital rotations. According to Imry and Ma [8], continuous degeneracy of the order parameter can be responsible for disordering under the action of a random field. In particular, Imry and Ma show that an arbitrarily weak field destroys long-range order for the vector order parameter $\mathbf{s}(\mathbf{r})$ that interacts with random field $\mathbf{h}(\mathbf{r})$ as

$$F_{IM} = -\int \mathbf{s}(\mathbf{r})\cdot\mathbf{h}(\mathbf{r})d^3r. \quad (2)$$

Indeed, the mean random field $\mathbf{h}(\mathbf{r})$ value vanishes; that is, $(1/L^3)\int \mathbf{h}(\mathbf{r})\,d^3r \longrightarrow 0$ as $L \longrightarrow \infty$, where L is the linear dimension of the region over which the integration is performed. The mean $\mathbf{h}(\mathbf{r})$ value tends to zero proportionally to $(d/L)^{3/2}$. The energy gain of the orientation of the order parameter along the mean field in the region with linear dimension on the order of L tends to zero by the same law. The energy loss decreases faster, by the law $(\xi_0/L)^2$, and, at large L, partitioning into domains becomes favorable. This partitioning destroys long-range order. When this general argument is applied to superfluid ^{3}He in aerogel, it should be borne in mind that interaction F_η differs from F_{IM}; namely, there exist nonzero $\bar{A}_{\mu j}$ for which F_η vanishes at all admissible $\eta^{(a)}_{jl}$. These $\bar{A}_{\mu j}$ are found from the equation

$$\eta^{(a)}_{jl}\bar{A}_{\mu j}\bar{A}^*_{\mu l} = 0. \quad (3)$$

Its solutions satisfy the equation

$$\bar{A}_{\mu l}\bar{A}^*_{\mu j} + \bar{A}_{\mu j}\bar{A}^*_{\mu l} = \delta_{jl}\cdot \text{const} \quad (4)$$

which is independent of η_{jl}. Equation (4) determines the real part of the $\bar{A}_{\mu j}\bar{A}^*_{\mu l}$ product, its imaginary part may be an arbitrary antisymmetric tensor. When interaction with a random field vanishes, changes in the orientation of the order parameter give no energy gain and long-range order is not destroyed. Already these qualitative arguments are evidence that condition (3) is a necessary criterion of the stability of the corresponding value $\bar{A}_{\mu j}$ with respect to random field $\eta_{jl}(\mathbf{r})$ if order parameter fluctuations can be ignored [9].

In the next section, we formulate a procedure for finding the order parameter in the presence of random field $\eta_{jl}(\mathbf{r})$.

3. THE SELECTION OF SUPERFLUID PHASES

The Ginzburg–Landau functional including interaction (1) is written as

$$F_{GL} = N(0)\int d^3r\Bigg[\tau A_{\mu j}A^*_{\mu j} + \eta_{jl}(\mathbf{r})A_{\mu j}A^*_{\mu l} + \frac{1}{2}\sum_{s=1}^{5}\beta_s I_s + \frac{1}{2}\Bigg(K_1\frac{\partial A_{\mu l}}{\partial x_j}\frac{\partial A^*_{\mu l}}{\partial x_j} + K_2\frac{\partial A_{\mu l}}{\partial x_j}\frac{\partial A^*_{\mu j}}{\partial x_l} + K_3\frac{\partial A_{\mu j}}{\partial x_j}\frac{\partial A^*_{\mu l}}{\partial x_l}\Bigg)\Bigg], \quad (5)$$

where $\tau = (T - T_c)/T_c$ and I_s is the nth fourth-order invariant in the expansion of free energy in $A_{\mu j}$; we will

not need expressions for I_s (see [10]) here. The coefficients $\beta_1, \ldots, \beta_5$ and K_1, K_2, and K_3 are phenomenological constants. In what follows, we assume that $K_1 = K_2 = K_3 \equiv K$, in accordance with the weak-coupling approximation. The gradient terms can also contain random components, for instance, of the form $u_j(\mathbf{r})A_{\mu l}\partial A^*_{\mu l}/\partial x_j$, where $u_j(\mathbf{r})$ is a random vector.[1] The component written above is, by its meaning, the local random velocity to within a factor of $\hbar/m$. All terms of this type are obtained by "elongating" the derivatives,

$$\frac{\partial}{\partial x_j} \longrightarrow \frac{\partial}{\partial x_j} + u_j(\mathbf{r}),$$

in energy equation (5). Further reasoning shows that these terms do not influence the selection of phases; for this reason, they are not taken into account here. The variation of functional (5) with respect to $A^*_{\mu j}$ gives the following equation for the equilibrium order parameter:

$$\tau A_{\mu j} + \frac{1}{2}\sum_{s=1}^{5}\beta_s\frac{\partial I_s}{\partial A^*_{\mu j}} - \frac{1}{2}K\left(\frac{\partial^2 A_{\mu j}}{\partial x_l^2} + 2\frac{\partial^2 A_{\mu l}}{\partial x_l \partial x_j}\right) = -A_{\mu l}\eta_{lj}. \quad (6)$$

The variation in A_{mj} gives the equation complex conjugate to (6). According to the estimate given above, random field $\eta_{jl}(\mathbf{r})$ is low. Larkin and Ovchinnikov [11] estimated the influence of a low random field on a one-component order parameter (conventional pairing) near T_c. We will use similar reasoning. A more complex form of the order parameter and the degeneracy mentioned above, however, require the introduction of nontrivial changes into the procedure used in [11].

A random field causes order parameter fluctuations $a_{\mu j}$ about its mean value,

$$A_{\mu j}(\mathbf{r}) = \bar{A}_{\mu j} + a_{\mu j}(\mathbf{r}).$$

The condition $\bar{A}_{\mu j} \neq 0$ is the criterion of long-range ordering, and it determines the transition temperature T_c. Not too close to T_c, $a_{\mu j}$ can be considered a value of the first order of smallness with respect to η_{jl}. We restrict our consideration to such temperatures. Let us expand (6) in this temperature region in the vicinity of $A_{\mu j} = \bar{A}_{\mu j}$ and retain terms up to second order in $a_{\mu j}$ and η_{jl} in this expansion,

$$\tau\bar{A}_{\mu j} + \tau a_{\mu j} + \frac{1}{2}\sum_{s=1}^{5}\beta_s\left[\frac{\partial I_s}{\partial A^*_{\mu j}} + \frac{\partial^2 I_s}{\partial A^*_{\mu j}\partial A_{\nu n}}a_{\nu n} + \frac{\partial^2 I_s}{\partial A^*_{\mu j}\partial A^*_{\nu n}}a^*_{\nu n} + \frac{1}{2}\left(\frac{\partial^3 I_s}{\partial A^*_{\mu j}\partial A_{\nu n}\partial A_{\beta l}}a_{\nu n}a_{\beta l} + 2\frac{\partial^3 I_s}{\partial A^*_{\mu j}\partial A^*_{\nu n}\partial A_{\beta l}}a^*_{\nu n}a_{\beta l}\right)\right] - \frac{1}{2}K\left(\frac{\partial^2 a_{\mu j}}{\partial x_l^2} + 2\frac{\partial^2 a_{\mu l}}{\partial x_l \partial x_j}\right) = -\bar{A}_{\mu l}\eta_{lj} - a_{\mu l}\eta_{lj}. \quad (7)$$

Averaging (7) over scales much larger than the mean distance between aerogel strands yields

$$\tau\bar{A}_{\mu j} + \frac{1}{2}\sum_{s=1}^{5}\beta_s\left[\frac{\partial I_s}{\partial A^*_{\mu j}} + \frac{1}{2}\left(\frac{\partial^3 I_s}{\partial A^*_{\mu j}\partial A_{\nu n}\partial A_{\beta l}}\langle a_{\nu n}a_{\beta l}\rangle + 2\frac{\partial^3 I_s}{\partial A^*_{\mu j}\partial A^*_{\nu n}\partial A_{\beta l}}\langle a^*_{\nu n}a_{\beta l}\rangle\right)\right] = -\langle a_{\mu l}\eta_{lj}\rangle. \quad (8)$$

Along with $\bar{A}_{\mu j}$, (8) contains the mean products of fluctuation components $\langle a_{\nu n}a_{\beta l}\rangle$, etc. In order to determine these components, we must gather the rapidly varying terms in (7) and in its complex conjugate,

$$\tau a_{\mu j} + \frac{1}{2}\sum_{s=1}^{5}\beta_s\left[\frac{\partial^2 I_s}{\partial A^*_{\mu j}\partial A_{\nu n}}a_{\nu n} + \frac{\partial^2 I_s}{\partial A^*_{\mu j}\partial A^*_{\nu n}}a^*_{\nu n} - \frac{1}{2}K\left(\frac{\partial^2 a_{\mu j}}{\partial x_l^2} + 2\frac{\partial^2 a_{\mu l}}{\partial x_l \partial x_j}\right)\right] = -\bar{A}_{\mu l}\eta_{lj}, \quad (9)$$

$$\tau a^*_{\mu j} + \frac{1}{2}\sum_{s=1}^{5}\beta_s\left[\frac{\partial^2 I_s}{\partial A_{\mu j}\partial A^*_{\nu n}}a^*_{\nu n} + \frac{\partial^2 I_s}{\partial A_{\mu j}\partial A_{\nu n}}a_{\nu n} - \frac{1}{2}K\left(\frac{\partial^2 a^*_{\mu j}}{\partial x_l^2} + 2\frac{\partial^2 a^*_{\mu l}}{\partial x_l \partial x_j}\right)\right] = -\bar{A}^*_{\mu l}\eta_{lj}, \quad (10)$$

This is a linear inhomogeneous system of equations. Because of the degeneracy of $\bar{A}_{\mu j}$ mentioned above, the corresponding homogeneous system has solutions. These are the increments of $\bar{A}_{\mu j}$ and $\bar{A}^*_{\mu j}$ that correspond to small rotation Ω_q,

$$\omega_{\mu j} = \Omega_q e^{jqr}\bar{A}_{\mu r}, \quad \omega^*_{\mu j} = \Omega_q e^{jqr}\bar{A}^*_{\mu r}, \quad (11)$$

where e^{jqr} is the absolutely antisymmetric tensor. To solve (9), (10), we must pass to Fourier transforms of $\eta_{jl}(\mathbf{k})$ and $a_{\mu j}(k)$ in these equations. In what follows, the character of the singularity of $a_{\mu j}(k)$ as $k \longrightarrow 0$ will only

[1] It was V.I. Marchenko who brought the existence of such energy terms to my attention.

be essential; the anisotropy of the gradient terms can therefore be ignored. This allows us to replace the terms with the derivatives

$$\frac{1}{2}K\left(\frac{\partial^2 a_{\mu j}}{\partial x_l^2} + 2\frac{\partial^2 a_{\mu l}}{\partial x_l \partial x_j}\right)$$

and

$$\frac{1}{2}K\left(\frac{\partial^2 a^*_{\mu j}}{\partial x_l^2} + 2\frac{\partial^2 a^*_{\mu l}}{\partial x_l \partial x_j}\right)$$

by

$$\frac{1}{2}\bar{K}\left(\frac{\partial^2 a_{\mu j}}{\partial x_l^2}\right)$$

and

$$\frac{1}{2}\bar{K}\left(\frac{\partial^2 a^*_{\mu j}}{\partial x_l^2}\right),$$

respectively, in (9) and (10). Multiplying both sides of (9) and (10) by $\omega^*_{\mu j}$ and $\omega_{\mu j}$, respectively, and summing the resulting equations yields

$$a^{\omega}(k) = -\frac{2}{K}\frac{(\omega^*_{\mu j}\bar{A}_{\mu l} + \omega_{\mu j}\bar{A}^*_{\mu l})\eta^{(a)}_{lj}}{k^2} \qquad (12)$$

for the projection $a_{\mu j}(k)\omega^*_{\mu j} + a^*_{\mu j}(k)\omega_{\mu j} \equiv a^{\omega}(k)$. In calculating the means $\langle a_{\nu n}a_{\beta l}\rangle$, the $a_{\mu j}$ components parallel to $\omega_{\mu j}$ give the contribution proportional to

$$[(\omega^*_{\mu j}\bar{A}_{\mu l} + \omega_{\mu j}\bar{A}^*_{\mu l})\eta^{(a)}_{lj}]^2\int\frac{d^3k}{k^4}.$$

This integral diverges at the lower limit. The diverging terms cannot be excluded by renormalizing the constants in the equation for the energy. For a solution to (8) to exist, it is necessary that the coefficient of the integral vanish; that is,

$$\Omega_n e^{jnr}Q_{rl}\eta^{(a)}_{lj} = 0, \qquad (13)$$

where $Q_{rl} = \bar{A}_{\mu r}\bar{A}^*_{\mu l} + \bar{A}_{\mu l}\bar{A}^*_{\mu r}$. Because Ω_n is selected arbitrarily, it follows from (13) that

$$Q_{rl}\eta^{(a)}_{lj} = Q_{jl}\eta^{(a)}_{lr}. \qquad (14)$$

The Q_{rl} matrix is Hermitian and can therefore be transformed to the diagonal form with real diagonal matrix elements q_r. In the corresponding basis, (14) can then be rewritten as

$$(q_r - q_j)\eta^{(a)}_{rj} = 0.$$

This equality should be satisfied at all admissible $\eta^{(a)}_{rn}$ values. It follows that all q_r are equal; that is,

$$\bar{A}_{\nu r}\bar{A}^*_{\nu j} + \bar{A}_{\nu j}\bar{A}^*_{\nu r} = q\delta_{rj},$$

which coincides with condition (4).

The order parameters that satisfy (4) can naturally be called "quasi-isotropic," because the energy of their interaction with aerogel does not change under arbitrary orbital rotations; that is, continuous degeneracy is also retained when random tensor field $\eta_{jl}(\mathbf{r})$ is taken into account. Also note that the order parameter $A_{\mu j}$ enters into the tensor of superfluid densities in the combination $A_{\mu l}\bar{A}^*_{\mu j} + A_{\mu j}\bar{A}^*_{\mu l}$; that is, condition (4) is a requirement that this tensor be isotropic.

It follows that random field $\eta_{jl}(\mathbf{r})$ does not destroy long-range order only of those ^{3}He phases that have a quasi-isotropic order parameter. The procedure for finding order parameters that correspond to the observed or possible superfluid phases should therefore begin with selecting a family of $\bar{A}_{\mu j}$ matrices that satisfy condition (4). These matrices are "the correct zeroth approximation" to the sought order parameter. Next, we must use (9) and (10) to express $a_{\mu j}$ and $a^*_{\mu j}$ via $\bar{A}_{\mu j}$ and $\eta_{jl}(\mathbf{r})$. In practice, the Fourier components of $a_{\mu j}(\mathbf{k})$ can be found more conveniently. After calculating the means $\langle a_{\nu n}a_{\beta l}\rangle$, etc., and substituting them into (8), (8) becomes a closed equation for determining $\bar{A}_{\mu j}$. The coefficients β_1, …, β_5 and K and the correlation functions $\langle\eta_{\nu n}(\mathbf{k})\eta_{\beta l}(-\mathbf{k})\rangle$ that (8) contains should be considered given. At $\eta_{jl}(\mathbf{r}) = 0$, we return to the usual equation for determining free energy extrema for pure ^{3}He.

The order parameter of the BW phase,

$$A^{BW}_{\mu j} = \Delta e^{i\varphi}R_{\mu j}, \qquad (15)$$

where $R_{\mu j}$ is a real orthogonal matrix, satisfies condition (4). Ignoring dipole interactions allows $R_{\mu j}$ to be transformed into an identity matrix by rotating the spin axes with respect to the orbital axes. Aerogel can naturally be considered uniform and isotropic. The tensor structure of the $\langle\eta_{\nu n}(\mathbf{k})\eta_{\beta l}(-\mathbf{k})\rangle$ correlation functions is then determined by symmetry [6]. It is also clear from symmetry considerations that the order parameter, which is proportional to the unit matrix, satisfies (8). It follows that the BW phase remains stable in the presence of aerogel. Compared with pure ^{3}He, the phenomenological coefficients β_1, …, β_5 change in the BW phase. This influences the region of BW phase stability and the thermodynamic properties of the phase, which depend on these coefficients. We will, however, omit explicit calculations of such fluctuation corrections to the β_1, …, β_5 coefficients.

The order parameter of the ABM phase,

$$A_{\mu j} = \Delta\frac{1}{\sqrt{2}}\hat{d}_\mu(\hat{m}_j + i\hat{n}_j), \quad (16)$$

does not satisfy criterion (4). This raises the question of the search for the order parameter that would be capable of describing the observed properties of the A-like phase.

4. THE ZEROTH APPROXIMATION FOR THE ESP PHASES

The measured magnetic susceptibility of the A phase equals that of the normal phase [4]. It follows that this phase does not contain Cooper pairs with a zero spin projection onto the magnetic field direction; that is, it is an equal spin pairing (ESP) phase. The order parameter of an arbitrary ESP phase can be written as

$$A_{\mu j} = \Delta\frac{1}{\sqrt{3}}[\hat{d}_\mu(m_j + in_j) + \hat{e}_\mu(l_j + ip_j)], \quad (17)$$

where $\hat{d}_\mu$ and $\hat{e}_\mu$ are the mutually orthogonal unit vectors and the m_j, n_j, l_j, and p_j vectors are arbitrary at this stage. Substituting order parameter (17) into condition (4) shows that, for this condition to be met, the m_j, n_j, l_j, and p_j vectors should satisfy the equation

$$m_j m_l + n_j n_l + l_j l_l + p_j p_l = \delta_{jl}. \quad (18)$$

It is assumed that the order parameter is normalized by the condition $A_{\mu j}A^*_{\mu j} = \Delta^2$. One of the solutions to (18) ($\boldsymbol{p} = 0$ and **m**, **n**, and **l** are the set of three orthonormalized vectors) was discussed in detail in [9]. All solutions can conveniently be found using the following procedure. Consider four four-dimensional vectors M_s, N_s, L_s, and P_s (s = 1, 2, 3, 4) that satisfy the equation

$$M_r M_s + N_r N_s + L_r L_s + P_r P_s = \delta_{rs}. \quad (19)$$

To within collective rotations and reflections, the only solution to (19) is a set of four vectors $\hat{q}^{(a)}$ such that $\hat{q}^{(a)} \cdot \hat{q}^{(b)} = \delta^{ab}$. Let us select an arbitrary four-dimensional unit vector $\hat{\nu} = (\nu_1, \nu_2, \nu_3, \nu_4)$ and project the $\hat{q}^{(a)}$ vectors onto the three-dimensional hyperplane orthogonal to $\hat{\nu}$. This yields the four three-dimensional vectors

$$\mathbf{m} = \hat{q}^{(1)} - \nu_1\hat{\nu}, \quad \mathbf{n} = \hat{q}^{(2)} - \nu_2\hat{\nu},$$
$$\mathbf{l} = \hat{q}^{(3)} - \nu_3\hat{\nu}, \quad \mathbf{p} = \hat{q}^{(4)} - \nu_4\hat{\nu}. \quad (20)$$

Multiplying the combination $m_j m_l + n_j n_l + l_j l_l + p_j p_l$ by an arbitrary vector a_l normal to $\hat{\nu}$ and using (20) for **m**, **n**, **l**, and **p** shows that these vectors satisfy (18). Equations (20) allow us to find some other properties of the vectors **m**, **n**, **l**, and **p**, namely,

$$m^2 + n^2 + l^2 + p^2 = 3, \quad (21)$$

$$\mathbf{m}\cdot\mathbf{n} = -\nu_1\nu_2, \quad \mathbf{m}\cdot\mathbf{l} = -\nu_1\nu_3,$$
$$\mathbf{n}\cdot\mathbf{l} = -\nu_2\nu_3, \ldots, \quad (22)$$

$$m^2 = 1 - \nu_1^2, \quad n^2 = 1 - \nu_2^2, \ldots. \quad (23)$$

Property (22) can be used to show that $[\mathbf{m}\times\mathbf{n}]\cdot[\mathbf{l}\times\mathbf{p}] = 0$; that is, the normals to the planes spanned by the **m**, **n** and **l**, **p** pairs of vectors, respectively, are mutually perpendicular. This property is retained for all pairs selected from the set of four vectors **m**, **n**, **l**, and **p**. It follows that (17) with the vectors **m**, **n**, **l**, and **p** given by (20) determines a three-parameter family of quasi-isotropic order parameters of the ESP type. Substituting these order parameters into (5) yields their energies in the zeroth approximation with respect to $\eta_{jl}(\mathbf{r})$,

$$\frac{F^{(0)}_{GL}}{N(0)} = \tau\Delta^2 + \frac{\Delta^4}{18}[\beta_1 + 9\beta_2 + \beta_3 + 5(\beta_4 + \beta_5)$$
$$- 4(\beta_1 + \beta_5)(\nu_1\nu_4 - \nu_2\nu_3)^2]. \quad (24)$$

The ν_1, ν_2, ν_3, and ν_4 parameters appear in (24) only in the combination $\Lambda \equiv \nu_1\nu_4 - \nu_2\nu_3$. If $\beta_1 + \beta_5 \equiv \beta_{15} < 0$, the free energy minimum is reached at $\Lambda = 0$, that is, at

$$\nu_1\nu_4 = \nu_2\nu_3. \quad (25)$$

In the weak coupling approximation, both coefficients β_1 and β_5 are negative, and the inequality $\beta_1 + \beta_5 < 0$ is satisfied with a large safety margin. Condition (25) has simple physical meaning. The order parameters determined by (17) are not unitary. The phases that correspond to them may have spin densities proportional to $e_{\mu\nu\lambda}A_{\mu j}A^*_{\nu j}$, that is, to $(2\Delta^2/3)[\hat{d}\times\hat{e}\,][\mathbf{n}\cdot\mathbf{l} - \mathbf{m}\cdot\mathbf{p}]$ for the order parameter given by (17). Property (22) allows us to easily ascertain that spontaneous spin density vanishes if condition (25) is satisfied. This condition determines a two-parameter family of nonferromagnetic quasi-isotropic phases that may include the A-like phase. This family can be parameterized as follows: $\nu_1 = \sin\alpha\sin\beta$, $\nu_2 = \sin\alpha\cos\beta$, $\nu_3 = \cos\alpha\sin\beta$, and $\nu_4 = \cos\alpha\cos\beta$. The parameters $\alpha = \pi/4$ and $\beta = \pi/4$ correspond to the most symmetrical nonferromagnetic configuration. We then have $\nu_1 = \nu_2 = \nu_3 = \nu_4 = 1/2$, the lengths of the **m**, **n**, **l**, and **p** vectors are all equal to $\sqrt{3}/2$, and the angles between two arbitrary vectors of this set are also equal. The vectors connecting the center of a regular tetrahedron and its vertices constitute such a set of four vectors.

If $\beta_{15} > 0$, the Λ^2 value should be as large as possible in the equilibrium state. The maximum Λ^2 value is 1/4 and is attained at $\nu_1 = \nu_4$, $\nu_2 = -\nu_3$ or $\nu_1 = -\nu_4$, $\nu_2 = \nu_3$.

The solutions make up a one-parameter family. In the first case, it can be parameterized as $\nu_1 = \nu_4 = \frac{1}{\sqrt{2}}\sin\gamma$ and $\nu_2 = -\nu_3 = \frac{1}{\sqrt{2}}\cos\gamma$. For instance, the set $\nu_1 = -1/2$, $\nu_2 = \nu_3 = \nu_4 = 1/2$ corresponds to the most symmetrical ferromagnetic solution; that is, this solution is obtained from the most symmetrical nonferromagnetic solution m_j, n_j, l_j, p_j by reversing one of the vectors.

5. MAGNETIC FIELD EFFECTS

In a magnetic field, two terms should be added to the free energy. One of these is quadratic in the field,

$$f_H^{(2)} = -\frac{1}{2}\chi_{\mu\nu}H_\mu H_\nu. \quad (26)$$

By the definition of ESP phases, one of the principal values of the magnetic susceptibility tensor $\chi_{\mu\nu}$ coincides with the susceptibility of the normal phase χ_n. Near T_c, the form of the $\chi_{\mu\nu}$ tensor is determined from symmetry considerations, $\chi_{\mu\nu} = \chi_n\delta_{\mu\nu} - \kappa(A_{\mu j}A_{\nu j}^* + A_{\nu j}A_{\mu j}^*)$. The second term on the right-hand side describes a decrease in the transverse susceptibility compared with χ_n. This is a two-dimensional tensor with the principal values $2\Delta^2\lambda_{1,2}/3$, where $\lambda_{1,2}$ are the roots of the equation

$$\lambda^2 - 3\lambda + 2 + \Lambda^2 = 0.$$

In the nonferromagnetic phase, $\Lambda = 0$ and, in addition, $\lambda_1 = 2$ and $\lambda_2 = 1$; that is, the transverse susceptibility is anisotropic. In the ferromagnetic phase, $\Lambda^2 = 1/4$, one obtains $\lambda_1 = \lambda_2 = 3/2$, and the transverse susceptibility is isotropic. In the equilibrium state, the order parameter is oriented in such a way that the largest principal value of $\chi_{\mu\nu}$ corresponds to the magnetic field direction. Additional energy (26) then has equal values for all A-like phases.

Apart from the term quadratic in field, the free energy contains the linear term

$$f_H^{(1)} = i\zeta e_{\mu\nu\lambda}A_{\mu j}A_{\nu j}^*H_\lambda. \quad (27)$$

In pure ^{3}He, this term splits the transition to the A phase into two transitions closely spaced along the temperature axis. First, a ferromagnetic A_1 phase is formed. This phase only contains Cooper pairs with a single spin projection. At a lower temperature, the transition to an A_2 phase occurs. In this phase, both spin projections are present. The ζ coefficient is proportional to the derivative of the density of states with respect to the energy, and the temperature interval in which the A_1 phase exists is narrow in measure of the smallness of $\mu H/\varepsilon_F$, where μ is the magnetic moment of the ^{3}He nucleus and ε_F is the Fermi energy.

In aerogel, the term linear in field also influences the sequence of phase transitions. Let us include it into energy (24),

$$\frac{F_{GL}^{(0)}}{N(0)} = \left(\tau - \frac{\zeta H\Lambda}{3}\right)\Delta^2 - \frac{2\Delta^4}{9}\beta_{15}\Lambda^2 + \frac{\Delta^4}{18}[\beta_1 + 9\beta_2 + \beta_3 + 5(\beta_4 + \beta_5)]. \quad (28)$$

This equation should be minimized with respect to Λ and Δ^2. The result depends on the sign of the sum of β_{15}. If $\beta_{15} > 0$, the energy minimum is reached at $|\Lambda| = 1/2$ at all Δ^2; that is, the ferromagnetic phase is stable. The transition to the superfluid state occurs at $\tau = \zeta H/6$. At $\tau < \zeta H/6$, we have $\Delta^2 = -9\tau_H/B$, where $\tau_H = \tau - \zeta_H/6$ and $B = 9\beta_2 + \beta_3 + 5\beta_4 + 4\beta_5$. The magnetic moment includes a small component $M = N(0)\zeta\Delta^2/6$, which is independent of the field but proportional to Δ^2.

However, if $\beta_{15} < 0$, the ferromagnetic phase corresponds to the minimum of energy (28) only in the temperature range $(B/\beta_{15})(\zeta_H/6) < \tau < \zeta H/6$. At $\tau_2 = \zeta HB/6\beta_{15}$, the transition to another phase with $\Lambda = -3\zeta H/4\beta_{15}\Delta^2$ occurs. As the temperature departs from τ_2, Λ tends to zero; that is, the additional magnetic moment vanishes. The transition at $\tau = \tau_2$ is similar to the $A_1 \longrightarrow A_2$ transition in pure ^{3}He. It follows that the ferromagnetic phases considered above are similar to the A_1 phase of pure ^{3}He. The pairing amplitudes for both spin projections $s = 1$ and $s = -1$ are, however, nonzero.

6. DISCUSSION

To summarize, the superfluid ^{3}He phases whose order parameter satisfies condition (4) can only form in aerogel. It was shown for the example of the A-like and A_1-like phases that this condition does not determine the $\bar{A}_{\mu j}$ matrix unambiguously but specifies quite a family of such matrices. The order parameter most favorable energetically can be selected by using an approximation of a higher order in $\eta_{jl}(\mathbf{r})$. The procedure for finding the solution becomes substantially more complex, and the solution then explicitly depends on the unknown correlation functions of the random field $\eta_{jl}(\mathbf{r})$. An attempt can be made to narrow the class of admissible solutions based on the physical properties of the observed superfluid phases. For instance, transition splitting in a magnetic field would be evidence that, of the two possibilities considered in the preceding section, we have $\beta_{15} < 0$ and the nonferromagnetic phase is stable far from T_c. Conversely, if the $A_1 \longrightarrow A_2$ transition is absent, the stable phase is ferromagnetic ($\beta_{15} > 0$). No definite data on this point are available. In the main approximation with respect to the random field, all quasi-isotropic phases have an isotropic tensor

of superfluid densities. In particular, this feature should distinguish the A-like phase from the ABM phase of pure ^{3}He.

ACKNOWLEDGMENTS

The author thanks V.V. Dmitriev and J. Parpia for discussions and useful comments and E.I. Kats for the invitation to Institut Laue–Langevin in Grenoble, where part of this work was accomplished, and for stimulating discussions.

This work was supported, in part, by the Russian Foundation for Basic Research (project no. 01-02-16714) and by the Ministry of Industry, Science, and Technology of the Russian Federation.

REFERENCES

1. J. V. Porto and J. M. Parpia, Phys. Rev. Lett. **74**, 4667 (1995).
2. A. A. Abrikosov and L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **39**, 1781 (1961) [Sov. Phys. JETP **12**, 1243 (1961)].
3. A. I. Larkin, Pis'ma Zh. Éksp. Teor. Fiz. **2**, 205 (1965) [JETP Lett. **2**, 130 (1965)].
4. B. I. Barker, Y. Lee, L. Polukhina, *et al.*, Phys. Rev. Lett. **85**, 2148 (2000).
5. V. V. Dmitriev, V. V. Zav'yalov, D. E. Zmeev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 371 (2002) [JETP Lett. **76**, 312 (2002)].
6. I. A. Fomin, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 220 (2002) [JETP Lett. **75**, 187 (2002)].
7. D. Rainer and M. Vuorio, J. Phys. C: Solid State Phys. **10**, 3093 (1977).
8. Y. Imry and S. Ma, Phys. Rev. Lett. **35**, 1399 (1975).
9. I. A. Fomin, Pis'ma Zh. Éksp. Teor. Fiz. **77**, 285 (2003) [JETP Lett. **77**, 240 (2003)].
10. D. Vollhardt and P. Wölfle, *The Superfluid Phases of Helium 3* (Taylor and Francis, London, New York, Philadelphia, 1990).
11. A. I. Larkin and Yu. N. Ovchinnikov, Zh. Éksp. Teor. Fiz. **61**, 1221 (1971) [Sov. Phys. JETP **34**, 651 (1971)].

Translated by A. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 981–985.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1122–1127.

SOLIDS
Structure

High-Pressure Behavior of the Bond-Bending Mode of AlN¶

E. V. Yakovenko[a], M. Gauthier[b], and A. Polian[b]

[a]*Institute for High-Pressure Physics, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia*
e-mail: katia@ns.hppi.troitsk.ru
[b]*Physique des Milieux Condensés, Université P&M Curie F 75252, Paris Cedex 05, France*
e-mail: Michel.Gauthier@pmc.jussieu.fr; Alain.Polian@pmc.jussieu.fr

Received July 8, 2003

Abstract—Lattice vibrations of the wurtzite-type AlN have been studied by Raman spectroscopy under high-pressure up to the phase transition to the rock salt structure at 20 GPa. Five fundamental bands E_2^2, $A_1(TO)$, $E_1(TO)$, $A_1(LO)$, and $E_1(LO)$ have a strong, positive pressure shift, whereas the shift of the low-frequency E_2^1 band is weakly positive. We have found that the bond-bending E_2^1 mode has a positive mode Grüneisen parameter $\gamma_i = 0.04$, which is qualitatively consistent with the recently reported value $\gamma_i = 0.12$ [21]. Thus, we confirm that AlN remains stable with respect to the bond-bending mode, while in most tetrahedral semiconductors, bond-bending modes soften on compression. Experimental results are compared with the first-principle calculations. © *2004 MAIK "Nauka/Interperiodica".*

The pressure-induced phonon softening is a significant characteristic property of tetrahedral semiconductors and has been reported in many experimental [1–10] and theoretical [11–18] works. Negative frequency shift of the low-energy modes of tetrahedral semiconductors on compression manifests itself in their well-known negative thermal expansions at low temperatures. These "soft" modes are shearing modes, involving bond bending in the first order of the strain [15, 18]. Phonon frequency drop, more pronounced for high-*Z* materials, reaches about 30% at the threshold of the pressure-induced phase transitions, when covalent tetrahedral structures lose their stability and transform into more densely packed arrangements. Experimentally, it has been found that the stability of tetrahedral structures with respect to the bond-bending modes correlates with their absolute stability under pressure to such an extent that the drop in frequency is faster for less stable compounds. Weinstein [4, 6] has found that for six diamond and zinc-blende structure ZnTe, Ge, Si, ZnSe, ZnS, and GaP semiconductors, there is a remarkable linearity between the mode Grüneisen parameter

$$\gamma_i = -\frac{d\ln\nu_i}{d\ln V}$$

(where ν_i is the frequency of the mode i and V is the volume) for the purely bond-bending $TA(X)$ mode and the transition pressure P_{tr} for these materials.

Previously, the only known experimental examples of bond-bending modes with a positive pressure shift were the bond-bending $TA(X)$ mode of diamond [19][1] and the E_2^1 mode of wurtzite-type BeO [20]. This behavior might be regarded as characteristic of the low-*Z* second-row semiconductors; however, recent Raman measurements found a similar behavior for the E_2^1 mode of the wurtzite-type AlN (w-AlN) at pressures up to 6 GPa [21]. Previous high-pressure Raman studies of w-AlN [22, 23] failed to measure the pressure shift of the E_2^1 mode, most probably due to the lack of high-quality crystals. We believe that this problem deserves special attention, because the bond-bending elasticity is one of the most prominent manifestations of directional covalent bonding and its pressure behavior should be studied in depth. From the fundamental standpoint, AlN represents an interesting and complicated case of covalent versus ionic bonding [24]: although its valence charge distribution is highly ionic [25], AlN adopts the tetrahedrally coordinated wurtzite structure and therefore belongs to covalent materials [26]. To ensure that the pressure coefficient of the E_2^1 mode of AlN is indeed positive, we have taken a complementary high-pressure Raman study of AlN up to its stability limit at

¶This article was submitted by the authors in English.

[1] We note that the experimental error bar for $\gamma_{TA}(X)$ for diamond obtained in this work is twice the value of $\gamma_{TA}(X)$ itself. However, the positive sign of $\gamma_{TA}(X)$ is indirectly corroborated by the positive thermal expansion coefficient of diamond at low temperatures.

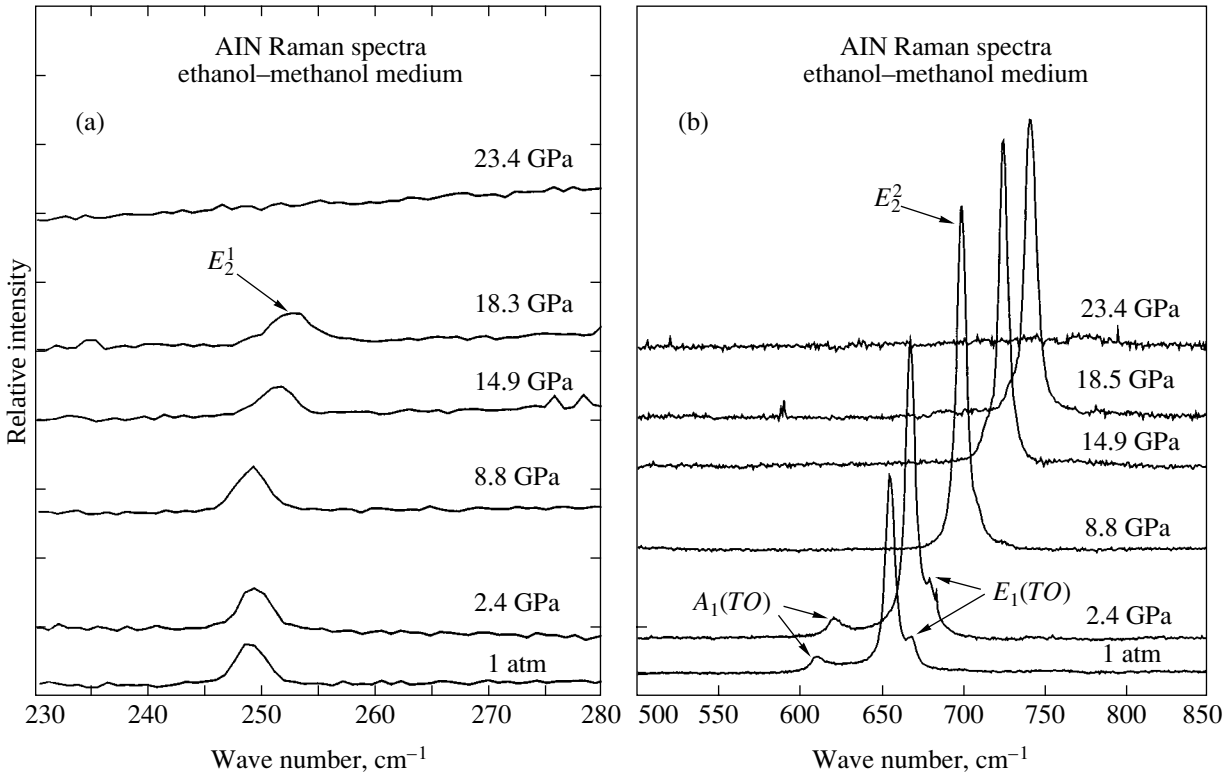


Fig. 1. Raman spectra of AlN as a function of pressure in the low-energy region (a) and in the high-energy region (b). The spectral resolution is 0.5 cm^{-1}. Ethanol–methanol mixture was used as a pressure transmitting medium.

about 20 GPa. The pressure dependence of the low-frequency bond-bending E_2^1 mode was traced up to the threshold of the pressure-induced phase transition for the first time.

The AlN samples were 20-μm-thick crystals grown on a sapphire substrate by vapor phase epitaxy. Pressure was produced using a diamond-anvil pressure cell. Compressed helium and a methanol–ethanol mixture were used as a pressure-transmitting medium in the first and in the second experimental run, respectively. Pressure was measured *in situ* by the ruby luminescence technique. The Raman spectra were measured using a THR-1000 triple spectrometer equipped with an OSMA detector (the first run), and a Dilor XY double spectrometer equipped with a CCD detector (the second run). An Ar^+ laser (λ = 514.5 nm) was used as the source of excitation. All spectra were recorded in the backscattering geometry at ambient temperature.

For the hexagonal wurtzite structure with the space group $P6_3mc$ (Z = 2), factor-group analysis predicts six sets of optical modes at k = 0 [27],

$$\Gamma_{op} = A_1 + 2B_1 + E_1 + 2E_2,$$

where A_1, E_1, and E_2 are Raman active modes and B_1 modes are silent. A_1 and E_1 are also infrared active and split into longitudinal and transverse components (*LO*

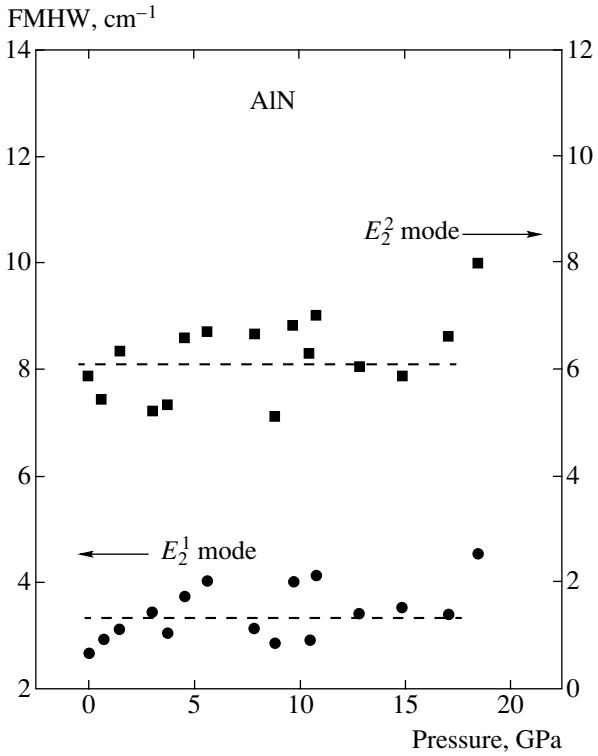


Fig. 2. Raman band width (FMHW) for the E_2^1 and E_2^2 modes of AlN as a function of pressure. An ethanol–methanol mixture was used as a pressure-transmitting medium. The dashed horizontal lines are drawn to guide the eye.

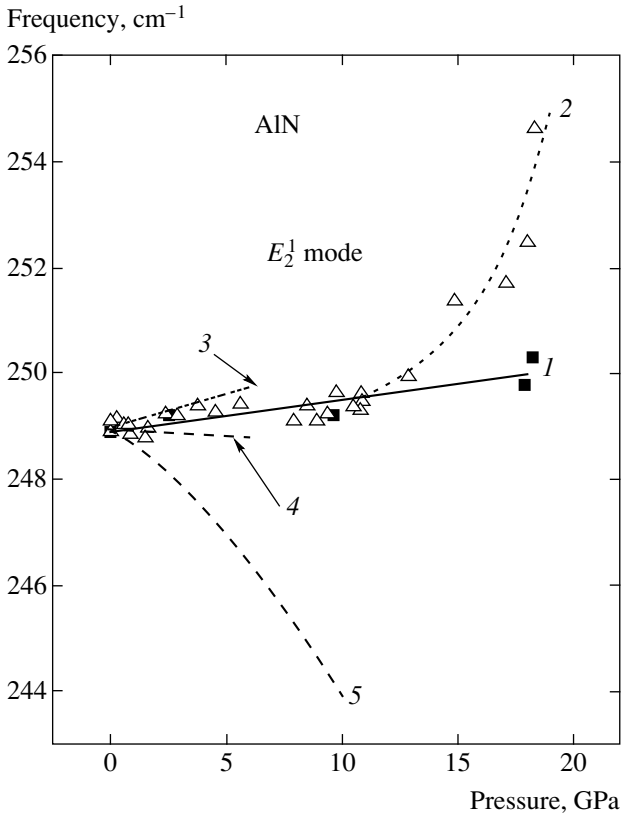


Fig. 3. Comparison of the measured and calculated pressure dependence of the Raman frequency for the E_2^1 mode of AlN. The squares are the first run data, obtained with compressed helium as a pressure-transmitting medium. The triangles are the second run data, obtained with the ethanol–methanol mixture as a pressure transmitting medium. Solid line *1* is a linear fit of the first run data. Dotted line *2* is a guide for the eye. Line *3* is the experimental dependence obtained in [21]. Lines *4* and *5* are the calculated dependences obtained in [21] and [33], respectively. All data are shifted along the vertical axis in order that the ambient pressure frequencies coincide with the value of 249 cm^{-1} obtained in our experiment.

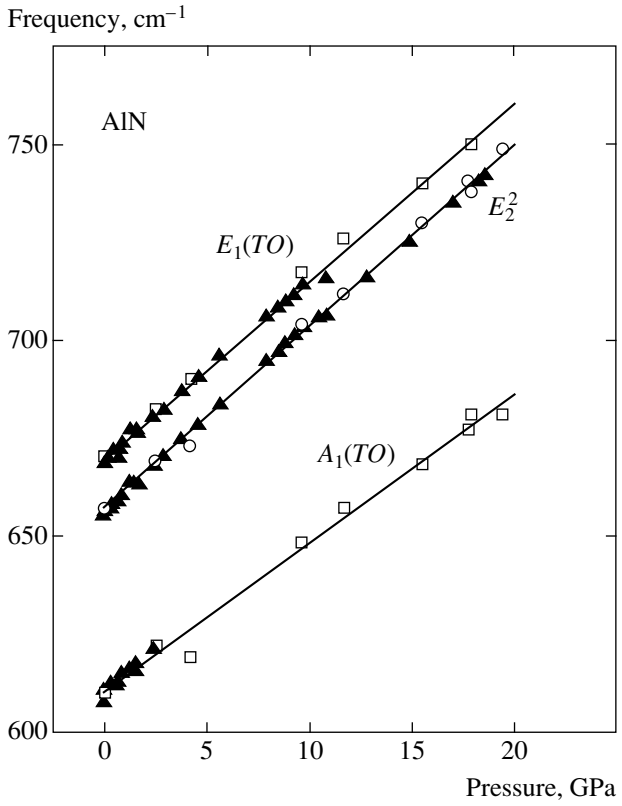


Fig. 4. The measured pressure dependence of the Raman frequency for the E_2^2 and *TO* modes of AlN. The squares and circles are the first run data, obtained with compressed helium as a pressure-transmitting medium. The triangles are the second run data, obtained with the ethanol–methanol mixture as a pressure-transmitting medium. The solid lines are linear fits of the first run data.

and *TO*). The lowest frequency mode E_2^1 is a bond-bending mode.

The Raman spectrum of w-AlN has been measured previously under ambient conditions and has been analyzed in some detail, including the effects of polarization and anisotropy [28–30]. Our ambient pressure Raman frequencies are 249, 610, 657, 669, 890, and 910 cm^{-1} for the respective modes E_2^1, $A_1(TO)$, E_2^2, $E_1(TO)$, $A_1(LO)$, and $E_1(LO)$. These values agree with very reliable data [28–30] to within 1%. On an increase in pressure, all Raman bands shift continuously to higher phonon energy, with neither broadening nor intensity loss, to about 18 GPa. Above 18 GPa, the bands weaken and disappear at about 21 GPa in both experimental runs due to the phase transition to the rock salt structure [31, 32]. Representative Raman spectra of AlN in the low- and high-energy region as a function of pressure are shown in Fig. 1. Pressure dependences of the E_2^1 and E_2^2 bandwidths are shown in Fig. 2.

Figure 3 compares measured and calculated pressure dependences of the Raman frequency for the E_2^1 mode of AlN. In our experiment, the pressure dependence of the E_2^1 frequency is weak but apparently positive and linear up to 13 GPa in both runs. Above 13 GPa, the first-run data fall well on the low-pressure dependence, while the second-run data indicate a sudden rise of the E_2^1 frequency. This is possibly associated with solidification of the ethanol–methanol medium, resulting in nonhydrostatic sample stress. Therefore, above 13 GPa, the first run data obtained in hydrostatic conditions with compressed helium as a pressure-transmitting medium are the most reliable. We note that solidification of the ethanol–methanol medium did not result in any detectable anomaly in the pressure dependence of the high-frequency bond-stretching modes (see Fig. 4). We believe that the non-

Table 1. Measured and calculated linear pressure coefficients ν_i' [cm^{-1} GPa^{-1}] obtained for the AlN Raman frequencies

	E_2^1	$A_1(TO)$	E_2^2	$E_1(TO)$	$A_1(LO)$	$E_1(LO)$
Experiment						
Present 1	0.05(1)	3.8(2)	4.9(2)	4.5(1)	–	–
Present 2	0.05(1)	4.3(2)	4.65(3)	4.55(6)	4.0(1)	3.6(7)
[21]	0.12(5)	4.4(1)	4.99(3)	4.55(3)	–	4.6(1)
Calculations						
[33]	−0.29	4.29	4.79	4.36	–	–
[21]	−0.03	3.0	4.2	3.8	3.5	4.0

hydrostatic stresses result in a much weaker response of these modes in comparison to their strong dependence on the high hydrostatic pressure.

The mode pressure coefficients ν_i' calculated using the linear least-square fit

$$\nu_i = \nu_{i0} + \nu_i' P,$$

where ν_i is the frequency of the mode i at the pressure P, are listed in Table 1. The E_2^1 frequency in the second run was fitted only to 12 GPa. Our mode pressure coefficients ν_i' are consistent with the results reported in [21], although with somewhat lower pressure slope for the E_2^1 frequency (see Table 1).

Ab initio calculations [21, 33] give a weak negative pressure shift for the E_2^1 mode of AlN. Nevertheless, the pressure coefficient of the E_2^1 mode is nearly zero, and the differences between experimental and calculated pressure shifts are quite small on the absolute scale. The agreement between measured and calculated coefficients of the E_2^2, *TO*, and *LO* modes is quite satisfactory (see Table 1).

As mentioned earlier, the bond-bending phonon modes of most tetrahedral semiconductors soften under compression and thus have negative Grüneisen parameters γ_i. Table 2 compiles experimental mode Grüneisen parameters obtained on the basis of the first- and the second-order Raman measurements and inelastic neutron scattering for the bond-bending modes in a series of tetrahedral compounds. The calculated Grüneisen parameters for diamond and BP are also displayed. A negative value of γ_i is observed in most cases except for diamond, BeO, and AlN. For BP, calculations [36] predict an exotic combination of a negative γ_i for the bond-bending $TA(X)$ mode and a positive γ_i for the bond-bending $TA(L)$ mode. The E_2 mode of SiC-6H has zero pressure slope and hence zero γ_i, but the quadratic pressure coefficient of the mode frequency is

Table 2. Mode-Grüneisen parameters γ_i for the bond-bending modes in a series of $A^N B^{8-N}$ compounds (W—wurtzite structure, ZB—zinc blende structure, D—diamond structure, L—hexagonal diamond structure)

Material	Structure	Mode	γ_i	Reference
Experiment				
CdS	W	E_2^1	–2.7	[6]
InP	ZB	$TA(L)$	–2.0	[6]
ZnO	W	E_2^1	–1.8	[6]
GaAs	ZB	$TA(L)$	–1.7	[6]
ZnSe	ZB	$TA(L)$	–1.5	[6]
ZnS	ZB	$TA(L)$	–1.5	[6]
Ge	D	$TA(L)$	–1.52	[10]
Si	D	$TA(L)$	–1.3	[3]
ZnTe	ZB	$TA(L)$	–1.0	[6]
GaP	ZB	$TA(L)$	–0.81	[2]
GaN	W	E_2^1	–0.426	[8]
SiC–6H	Hex.	E_2	0.0	[7]
BeO	W	E_2^1	0.04	[20]
AlN	W	E_2^1	0.04	This study
			0.10	[21]
C	D	$TA(X)$	0.4	[19]
Calculations				
C	D	$TA(X)$	0.3	[34]
C	D	$TA(L)$	0.17	[34]
C	L	E_{2u}	0.16	[35]
BP	ZB	$TA(X)$	−0.64	[36]
			−0.27	[37]
BP	ZB	$TA(L)$	0.121	[36]

negative [7]. We see that AlN is one of the most stable materials with respect to the bond-bending mode on compression.

Despite this, w-AlN undergoes a first-order phase transition to the rock-salt structure at a rather low pressure of 20 GPa. At the same time, SiC-6H and w-BeO, which have nearly the same values of γ_i for the bond-bending E_2 modes as AlN does, preserve the tetrahedral structures up to the pressures as high as 100 GPa [38] and 140 GPa [39], respectively. This obviously indicates that the applicability of Weinstein's empirical correlation rule [4, 6] is limited.

The pressure behavior of the bond-bending modes of tetrahedral semiconductors can be elucidated in terms of the pressure-sensitive balance between stabilizing and destabilizing contributions to the restoring force constants [15]. This balance, in turn, can be traced back to the atomic configuration of the constituent atoms, as has been done in the analysis of the thermodynamical stability of the diamond phase of carbon [40]. However, this issue is beyond the scope of the present report and will be discussed in a subsequent paper.

The authors wish to thank A. Dobrynin for growing the AI1N crystals. E. V. Yakovenko is grateful to A.F. Goncharov for his assistance in Raman measurements.

REFERENCES

1. R. T. Payne, Phys. Rev. Lett. **13**, 53 (1964).
2. B. A. Weinstein and G. J. Piermarini, Phys. Lett. A **48**, 14 (1974).
3. B. A. Weinstein and G. J. Piermarini, Phys. Rev. B **12**, 1172 (1975).
4. B. A. Weinstein, Solid State Commun. **24**, 595 (1977).
5. D. Olego and M. Cardona, Phys. Rev. B **25**, 1151 (1982).
6. B. A. Weinstein and R. Zallen, in *Light Scattering in Solids*, Ed. by M. Cardona and G. Guntherodt (Springer, Heidelberg, 1984), Issue 4, p. 463.
7. E. V. Yakovenko, A. F. Goncharov, and S. M. Stishov, High Press. Res. **7**, 433 (1991).
8. P. Perlin, C. Jauberthie-Carillon, J. P. Itié, *et al.*, Phys. Rev. B **45**, 83 (1992).
9. H. Olijnyk, High Press. Res. **10**, 461 (1992).
10. S. Klotz, J. M. Besson, M. Braden, *et al.*, Phys. Rev. Lett. **79**, 1313 (1997).
11. G. Dolling and R. A. Cowley, Proc. Phys. Soc. London **88**, 463 (1966).
12. R. M. Martin, Phys. Rev. **186**, 871 (1969).
13. H. Jex, Phys. Status Solidi B **45**, 343 (1971).
14. H. Wendel and R. M. Martin, Phys. Rev. B **19**, 5251 (1979).
15. M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 3259 (1982); **26**, 5668 (1982).
16. K. J. Chang and M. L. Cohen, Phys. Rev. B **31**, 7819 (1985); **34**, 8581 (1986).
17. O. H. Nielsen, Phys. Rev. B **34**, 5808 (1986).
18. D. J. Chadi and R. M. Martin, Solid State Commun. **19**, 643 (1976).
19. B. J. Parsons, Proc. R. Soc. London, Ser. A **352**, 397 (1976).
20. A. P. Jephcoat, R. J. Hemley, H. K. Mao, *et al.*, Phys. Rev. B **37**, 4727 (1988).
21. A. R. Goñi, H. Siegle, K. Syassen, *et al.*, Phys. Rev. B **64**, 035205 (2001).
22. J. A. Sanjurjo, E. Lopez-Cruz, P. Vogl, *et al.*, Phys. Rev. B **28**, 4579 (1983).
23. P. Perlin, A. Polian, and T. Suski, Phys. Rev. B **47**, 2874 (1993).
24. K. Karch and F. Bechstedt, Phys. Rev. B **56**, 7404 (1997).
25. E. Gabe, Y. LePage, and S. L. Mair, Phys. Rev. B **24**, 5634 (1981).
26. W. A. Harrison, *Electronic Structure and the Properties of Solids: The Physics of the Chemical Bond* (Freeman, San Francisco, 1980; Mir, Moscow, 1983).
27. W. G. Fateley, F. R. Dollish, N. T. McDevitt, *et al.*, in *Infrared and Raman Selection Rules for Molecular and Lattice Vibrations: The Correlation Method* (Wiley–Interscience, New York, 1972).
28. L. E. McNeil, M. Grimsditch, and R. H. French, J. Am. Ceram. Soc. **76**, 1132 (1993).
29. L. Filippidis, H. Siegle, A. Hoffman, *et al.*, Phys. Status Solidi B **198**, 621 (1996).
30. V. Yu. Davydov, Yu. E. Kitaev, I. N. Goncharuk, *et al.*, Phys. Rev. B **58**, 12899 (1998).
31. M. Ueno, A. Onodera, O. Shimomura, *et al.*, Phys. Rev. B **45**, 10123 (1992).
32. Q. Xia, H. Xia, and A. L. Ruoff, J. Appl. Phys. **73**, 8198 (1993).
33. I. Gorczyca, N. E. Christensen, E. L. Peltzer, *et al.*, Phys. Rev. B **51**, 11936 (1995).
34. J. Xie, S. P. Chen, J. S. Tse, *et al.*, Phys. Rev. B **60**, 9444 (1999).
35. B. R. Wu and Ji-an Xu, Phys. Rev. B **60**, 2964 (1999).
36. D. N. Talvar, G. Thaler, S. Zaranek, *et al.*, Phys. Rev. B **55**, 11293 (1997).
37. H. W. Leite Alves and K. Kunc, J. Phys.: Condens. Matter **4**, 6603 (1992).
38. M. Yoshida, A. Onodera, M. Ueno, *et al.*, Phys. Rev. B **48**, 10587 (1993).
39. Y. Mori, T. Ikai, and K. Takarabe, in *Abstracts of High-Pressure Conference* (Japan, 2002); Rev. High-Pressure Sci. Technol. **12**, 2D05 (2002).
40. M. T. Yin and M. L. Cohen, Phys. Rev. Lett. **50**, 2006 (1983).

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 986–996.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1128–1138.

SOLIDS
Structure

Weakly Nonlinear Analysis of the Morphological Stability of a Two-Dimensional Cylindrical Crystal

L. M. Martyushev*, E. M. Sal'nikova, and E. A. Chervontseva

Institute of Industrial Ecology, Ural Division, Russian Academy of Sciences, Yekaterinburg, 620219 Russia

**e-mail: mlm@ecko.uran.ru*

Received July 22, 2003

Abstract—We present the first weakly nonlinear analysis of the morphological stability of a two-dimensional cylindrical crystal growing from solution in an arbitrary regime (with the growth rate proportional to supersaturation). A quadratic (with respect to the perturbation amplitude) correction to the critical radius of a stable crystal determined in the linear theory is obtained in an analytical form and studied as a function of the perturbation frequency and the growth regime. It is established that an increase in the perturbation amplitude virtually always leads to a decrease in the critical radius. Factors accounting for this nontrivial effect are considered. © *2004 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Problems pertaining to the morphological stability of growing crystals are of importance both in practical applications (e.g., for predicting the microstructure of solidifying ingots) and in basic science (e.g., for describing dissipative structures formed under nonequilibrium conditions). After the classical study of Mullins and Sekerka [1], presenting an original linear analysis of the morphological stability for a growing spherical crystal, this approach was extended to other geometries and more general conditions (including attachment kinetics, surface self-diffusion, etc.) [2]. The results of this linear analysis of stability were qualitatively and quantitatively confirmed by the results of numerous experiments [2–5].

Hardy and Coriell [3] studied the growth of perturbations in a cylindrical crystal of ice growing from distilled water under conditions of low supercooling. Initially, the crystal grew in the form of a smooth cylinder. Then, when the crystal reached certain critical sizes, its shape had visible, rapidly growing distortions. The experimentally measured critical radius of a stable cylindrical crystal was in good agreement with the value predicted by the theory, which allowed a rather accurate method to be developed for the indirect determination of the interfacial free energy of the water–ice system.

However, experiments revealed an interesting peculiarity that could not be described within the framework of the linear analysis of stability: the possible coexistence of particles in various forms (morphological phases) in some regions of the control parameters (e.g., supersaturation) [6–8]. Analysis of this and some other experimental features [9–11] led to the idea that the morphological transition from one crystal shape to another can be considered by analogy with the usual phase transitions of the first order. In the case of nonequilibrium crystallization, the role of thermodynamic potential is played by the production of entropy.

The results of our recent calculations [11–13] performed within a linear analysis of the morphological stability, with determination of the production of entropy, allowed the boundaries of metastable regions (the regions of coexistence of various morphological phases) to be determined for the first time. In this context, there is the need for independent theoretical verification and justification of the main hypotheses underlying the calculations [11–13]. One of these hypotheses is that an increase in the amplitude of perturbations (experimentally implemented by applying thermal or acoustic action, introducing special impurities, etc.) on the initially smooth crystal surface, leads to a decrease in the critical radius of a stable crystal.

The most natural way of considering the critical radius as dependent on the perturbation amplitude is to perform an analysis of stability at higher (second and above) orders of perturbation theory. Only a few such attempts (using shape perturbations with harmonics of small but finite amplitude) have been made because of complex and tedious calculations [14–16]. The weakly nonlinear analysis of diffusion-limited growth was performed in these investigations up to the third order with respect to a small parameter. It was confirmed that an increase in the perturbation amplitude almost always leads to a decrease in the critical radius of a stable crystal [14–16]. However, the assumption of a diffusion-limited crystal growth regime used in these investigations posed a significant limitation upon their generality: the question naturally arises as to whether the above

important result has a universal character, rather than being related to the specific growth regime.

The study was aimed at performing a weakly nonlinear analysis of the morphological stability of a two-dimensional cylindrical crystal growing from solution in an arbitrary regime (with the local growth rate proportional to supersaturation). This would significantly generalize the results obtained in [14], showing how universal is the dependence of the critical radius of a stable crystal on the perturbation amplitude, which is very important for justification of the approach developed in [11].

2. FORMULATION OF THE PROBLEM

We consider the growth of an initially smooth, round single two-dimensional crystal from a supersaturated solution. The main assumptions are as follows.

(i) Crystallization proceeds under isothermal–isobaric conditions and both the free surface energy and the kinetic coefficient are isotropic.

(ii) The field of concentrations C is described by the Laplace equation

$$\nabla^2 C = 0, \tag{1}$$

where the symbol ∇ denotes the nabla operator.

(iii) It is assumed that an arbitrarily small distortion of the circle can be represented as a superposition of harmonic functions of the type $\cos k\varphi$, where φ is the polar angle and k is a positive integer.

(iv) The solution concentration obeys the following boundary conditions:

$$C(R_\lambda) = C_\infty, \tag{2}$$

$$D\frac{\partial C}{\partial \tilde{\mathbf{n}}}\bigg|_{\tilde r = R + a\cos k\varphi} = \beta(C|_{\tilde r = R + a\cos k\varphi} - C_S), \tag{3}$$

$$G_S = C_0 + C_0\Gamma\tilde K,$$

$$\tilde K = \frac{\tilde r^2 + 2(\partial\tilde r/\partial\varphi)^2 - \tilde r\partial^2\tilde r/\partial\varphi^2}{(\tilde r^2 + (\partial\tilde r/\partial\varphi)^2)^{3/2}}. \tag{4}$$

Here, $\tilde r = R + a\cos k\varphi$ is the shape of a distorted circular boundary, R is the radius of the unperturbed circle, $a(t)$ is the perturbation amplitude ($a \ll R$), t is the current time, D is the diffusion coefficient, β is the kinetic coefficient of crystallization, C_∞ is the solution concentration at a large distance R_λ ($R_\lambda \gg R$) from the crystal surface, C_S is the equilibrium solution concentration near an arbitrary boundary, C_0 is the equilibrium solution concentration near a plane boundary, Γ is the coefficient of surface tension, and $\tilde K$ is the curvature.

There are two important remarks concerning the boundary conditions.

(i) Condition (2) can be given two interpretations. First, we can assume that concentration C_∞ is maintained at a certain distance R_λ and seek a stationary concentration distribution satisfying Eqs. (1)–(3). However, another possible interpretation is more useful from the standpoint of analysis of the loss of stability. It was demonstrated [17, 18] that, under conditions of low supersaturation, the solution of the nonstationary problem with the boundary conditions $C(\infty) = C_\infty$ and (3) coincides with the solution of problem (1)–(3), provided that R_λ corresponds to a certain boundary determined by the relations

$$R_\lambda = R/\nu\lambda, \quad \lambda^2\ln(\nu^2\lambda^2) + S = 0.$$

Here, $S = (C_\infty - C_S)/(C_{\rm sol} - C_S)$, $C_{\rm sol}$ is the crystal density, λ is a parameter, and $\ln\nu^2 = 0.5772$ is the Euler constant.

(ii) Boundary condition (3) describes the balance of a substance under the assumption that the solution concentration is negligibly small as compared to the crystal density. This assumption, significantly simplifying the problem, is well satisfied in many real systems featuring crystallization from solution.

For the convenience of calculations, we will pass to dimensionless variables in Eqs. (1)–(4), by scaling distances to the radius of nucleation in a saturated solution, $R^* = C_0\Gamma/(C_\infty - C_0)$ [17] and representing the concentration field as $u = (C - C_0)/C_0$. In these variables, the Laplace equation is written as

$$\nabla^2 u = 0, \tag{5}$$

and the boundary conditions appear as

$$u(\rho_\lambda) = \Delta, \tag{6}$$

$$\alpha\frac{\partial u}{\partial \mathbf{n}}\bigg|_{r = \rho + \delta\cos k\varphi} = u|_{r = \rho + \delta\cos k\varphi} - u_s,$$
$$u_s = K\Delta. \tag{7}$$

Here, $\mathbf{n} = \tilde{\mathbf{n}}/R^*$, $r = \tilde r/R^*$, $\rho = R/R^*$, $\delta(t) = a(t)/R^*$, $\Delta = (C_\infty - C_0)/C_0$ is the degree of supersaturation, $K = \tilde K R^*$, $\rho_\lambda = R_\lambda/R^*$, and $\alpha = D/\beta R^*$.

In boundary condition (7), we can pass from the operator $\partial/\partial\mathbf{n}$ to the vector components using the relation

$$\frac{\partial u}{\partial \mathbf{n}} = \nabla u\cdot\mathbf{e}_n,$$

where

$$\nabla u = \frac{\partial u}{\partial \mathbf{r}}\mathbf{i}_r + \frac{1}{r}\frac{\partial u}{\partial\varphi}\mathbf{i}_\varphi,$$

$$\mathbf{e}_n = \frac{\nabla\Phi}{|\nabla\Phi|}, \quad \Phi = r - \rho - \delta\cos k\varphi,$$

and $\mathbf{e}_n$ is the unit vector normal to the surface $\Phi = 0$. As

a result, condition (7) transforms to

$$\frac{\alpha}{\sqrt{1+\frac{\delta^2 k^2}{r^2}\sin^2 k\varphi}}\left(\frac{\partial u}{\partial r}+\frac{\delta k}{r^2}\sin k\varphi\frac{\partial u}{\partial \varphi}\right)\Bigg|_r = u|_r - u_s. \quad (8)$$

3. WEAKLY NONLINEAR ANALYSIS OF THE MORPHOLOGICAL STABILITY

3.1. Calculation of the Concentration Field

Let us expand the concentration field into series in powers of δ,

$$u(r,\varphi) = u_0(r) + u_1(r,\varphi)\delta + u_2(r,\varphi)\delta^2 + u_3(r,\varphi)\delta^3, \quad (9)$$

where u_0, …, u_3 are the expansion coefficients. Substituting expression (9) into the initial equation (5) and the boundary conditions (6) and (8), we expand each term into Taylor's series in δ up to the third power in the vicinity of ρ. The curvature is expressed as

$$K = K_0 + K_1\delta + K_2\delta^2 + K_3\delta^3, \quad (10)$$

where K_0, K_1, K_2, and K_3 are the coefficients calculated in the Appendix. Equating the coefficients at like powers of δ, we obtain four sets of equations for determining the functions $u_0(r)$, $u_1(r, \varphi)$, $u_2(r, \varphi)$, and $u_3(r, \varphi)$:

1.
$$\nabla^2 u_0 = 0,$$
$$u_0(\rho_\lambda) = \Delta,$$
$$\alpha\frac{\partial u_0}{\partial r}\Big|_\rho - u_0|_\rho = K_0\Delta; \quad (11)$$

2.
$$\nabla^2 u_1 = 0,$$
$$u_1(\rho_\lambda) = 0,$$
$$\alpha\frac{\partial u_1}{\partial r}\Big|_\rho - u_1|_\rho = \left(\frac{\partial u_0}{\partial r}\Big|_\rho - \alpha\frac{\partial^2 u_0}{\partial r^2}\Big|_\rho\right)\cos k\varphi - K_1\Delta; \quad (12)$$

3.
$$\nabla^2 u_2 = 0,$$
$$u_2(\rho_\lambda) = 0,$$
$$\alpha\frac{\partial u_2}{\partial r}\Big|_\rho - u_2|_\rho = \frac{1}{2}\left(\frac{\partial^2 u_0}{\partial r^2}\Big|_\rho - \alpha\frac{\partial^3 u_0}{\partial r^3}\Big|_\rho\right)\cos^2 k\varphi + \frac{\alpha k^2\sin^2 k\varphi}{2\rho^2}\frac{\partial u_0}{\partial r}\Big|_\rho + \left(\frac{\partial u_1}{\partial r}\Big|_\rho - \alpha\frac{\partial^2 u_1}{\partial r^2}\Big|_\rho\right)\cos k\varphi + \frac{\alpha k\sin k\varphi}{\rho^2}\frac{\partial u_1}{\partial r}\Big|_\rho - K_2\Delta; \quad (13)$$

4.
$$\nabla^2 u_3 = 0,$$
$$u_3(\rho_\lambda) = 0,$$
$$\alpha\frac{\partial u_3}{\partial r}\Big|_\rho - u_3|_\rho = \frac{1}{6}\left(\frac{\partial^3 u_0}{\partial r^3}\Big|_\rho - \alpha\frac{\partial^4 u_0}{\partial r^4}\Big|_\rho\right)\cos^3 k\varphi + \frac{1}{2}\left(\frac{\partial^2 u_1}{\partial r^2}\Big|_\rho - \alpha\frac{\partial^3 u_1}{\partial r^3}\Big|_\rho\right)\cos^2 k\varphi + \left(\frac{\partial u_2}{\partial r}\Big|_\rho - \alpha\frac{\partial^2 u_2}{\partial r^2}\Big|_\rho\right)\cos k\varphi - \alpha\frac{k\sin k\varphi\cos k\varphi}{\rho^2}\left(\left(\frac{\partial^2 u_1}{\partial r\partial\varphi}\right)\Big|_\rho + \frac{2}{\rho}\left(\frac{\partial u_1}{\partial\varphi}\right)\Big|_\rho\right) - \alpha\frac{k\sin k\varphi}{\rho^2}\left(\frac{\partial u_2}{\partial\varphi}\right)\Big|_\rho + \frac{\alpha k^2\sin^2 k\varphi\cos k\varphi}{2\rho^2}\left(\frac{\partial^2 u_0}{\partial r^2}\Big|_\rho - \frac{2}{\rho}\frac{\partial u_0}{\partial r}\Big|_\rho\right) + \frac{\alpha k^2\sin^2 k\varphi}{2\rho}\frac{\partial u_1}{\partial r}\Big|_\rho - K_3\Delta. \quad (14)$$

The solution of the Laplace equation in the ring ($\rho < r < \rho_\lambda$) for each *i*th system can be written in the following form [19]:

$$u_i = A_{i0} + B_{i0}\ln r + \sum_{n=1}^{\infty} r^{-n}(A_{in}\cos n\varphi + B_{in}\sin n\varphi) + \sum_{n=1}^{\infty} r^{n}(E_{in}\cos n\varphi + F_{in}\sin n\varphi), \quad (15)$$

where $i = 0, 1, 2,$ or 3.

3.1.1. Unperturbed solution (zero order). Substituting expression (15) for $i = 0$ into Eqs. (11) and equating

the coefficient at the like trigonometric functions, we obtain

$$A_{0n} = B_{0n} = E_{0n} = F_{0n} = 0, \quad (16)$$

$$A_{00} = \Delta - B_{00}\ln\rho_\lambda, \quad (17)$$

$$B_{00} = \frac{\Delta(\rho-1)}{\alpha+\rho A_\lambda}, \quad (18)$$

where $A_\lambda = \ln(\rho_\lambda/\rho)$. A solution to the problem in the zero approximation is

$$u_0 = \Delta + \frac{\Delta(\rho-1)}{\alpha+\rho A_\lambda}\ln\frac{r}{\rho_\lambda}. \quad (19)$$

3.1.2. First-order perturbation solution. Substituting expression (15) for $i = 1$ into the boundary conditions of set (12) and equating the coefficients at like trigonometric functions for the same harmonic numbers, we obtain two nonzero constants of the first-order approximation for $n = k$:

$$A_{1k} = A_1 z\rho_\lambda^k, \quad (20)$$

$$E_{1k} = -A_1 z\rho_\lambda^{-k}, \quad (21)$$

where

$$A_1 = \frac{\Delta(k^2-1) - B_{00}(\rho+\alpha)}{\rho(z^2(\alpha k-\rho)+(\alpha k+\rho))}, \quad z = \frac{\rho^k}{\rho_\lambda^k}.$$

The solution to Eqs. (12) is

$$u_1(r,\varphi) = A_1 z\cos k\varphi\left(\frac{\rho_\lambda^k}{r^k} - \frac{r^k}{\rho_\lambda^k}\right). \quad (22)$$

3.1.3. Second-order perturbation solution. Substituting expression (15) for $i = 2$ into the boundary conditions of set (13), expressing the powers of trigonometric functions via trigonometric functions of the corresponding multiple arguments, and equating the coefficients at the same harmonic numbers, we obtain four nonzero constants of the second-order approximation for $n = 0$ and $n = 2k$:

$$A_{20} = -B_{20}\ln\rho_\lambda, \quad (23)$$

$$\begin{aligned} B_{20} = -\{&2A_1\rho k(\alpha+\rho)(z^2+1) \\ &+ B_{00}((\alpha+\rho)-\alpha(k^2-1)) \\ &+ \Delta(2-3k^2)\}/\{4\rho^2(\alpha+A_\lambda\rho)\}, \end{aligned} \quad (24)$$

$$A_{2,2k} = A_2 z^2\rho_\lambda^{2k}, \quad (25)$$

$$E_{2,2k} = -A_2 z^2\rho_\lambda^{-2k}, \quad (26)$$

where the expression for A_2 is given in the Appendix. Substituting formulas (23)–(26) into Eqs. (15), for $i = 2$ we obtain a solution to Eqs. (13):

$$u_2(r,\varphi) = B_{20}\ln\frac{r}{\rho_\lambda} + A_2 z^2\cos 2k\varphi\left(\frac{\rho_\lambda^{2k}}{r^{2k}} - \frac{r^{2k}}{\rho_\lambda^{2k}}\right). \quad (27)$$

3.1.4. Third-order perturbation solution. Substituting expression (15) for $i = 3$ into the boundary conditions of set (14), expressing the powers of trigonometric functions via trigonometric functions of the corresponding multiple arguments, and equating the coefficients at the same harmonic numbers, we obtain four nonzero constants of the third-order approximation for $n = k$ and $n = 3k$:

$$E_{3,k} = -A_{3,k}\rho_\lambda^{-2k}, \quad (28)$$

$$E_{3,3k} = -A_{3,3k}\rho_\lambda^{-6k}, \quad (29)$$

where the expressions for $A_{3,k}$ and $A_{3,3k}$ via the constants of previous orders are presented in Appendix. Substituting these quantities into Eqs. (15) for $i = 3$, we eventually obtain the solution to Eqs. (14):

$$\begin{aligned} u_3(r,\varphi) = &\frac{A_{3,k}}{\rho_\lambda^k}\cos k\varphi\left(\frac{\rho_\lambda^k}{r^k} - \frac{r^k}{\rho_\lambda^k}\right) \\ &+ \frac{A_{3,3k}}{\rho_\lambda^{3k}}\cos 3k\varphi\left(\frac{\rho_\lambda^{3k}}{r^{3k}} - \frac{r^{3k}}{\rho_\lambda^{3k}}\right). \end{aligned} \quad (30)$$

Finally, substituting solutions (19), (22), (27), and (30) into expansion (9), we obtain an expression for the concentration field $u(r, \varphi)$ in the form of a series in powers of δ.

3.2. Calculation of the Radius of Stability of a Circular Crystal

Once the concentration field is known, we can determine the radius of a stable crystal. To within a positive constant factor, the local crystal growth rate V can be written as

$$V \sim \left.\frac{\partial u}{\partial \mathbf{n}}\right|_{r=\rho+\delta\cos k\varphi}. \quad (31)$$

Substituting the above expression for $u(r, \varphi)$ into Eq. (31) and expanding the resulting expression into series in δ in the vicinity of ρ, we can eventually represent the growth rate as

$$\begin{aligned} V = &V_0 + \frac{1}{\rho^4} \\ &\times (V_1\cos k\varphi + V_2\cos 2k\varphi + V_3\cos 3k\varphi), \end{aligned} \quad (32)$$

where

$$V_0 = \frac{B_{00}}{\rho} + \frac{1}{4\rho^3} \times (2A_1\rho k(z^2+1) + 4B_{02}\rho^2 - B_{00}(k^2-2))\delta^2, \quad (33)$$

$$V_1 = \left(\frac{5}{8}A_1\rho k^2(z^2-1) - k^2\rho^2 A_2(z^4-1) + A_2\rho^2 k(z^4+1) - \frac{3}{4}A_1\rho k(z^2+1) - A_{3,k}k\rho^{(3-k)}(z^2+1) - B_{20}\rho^2 + \frac{3}{8}(k^2-2)B_{00}\right)\delta^3 - (A_1\rho kz^2 + B_{00} + A_1\rho k)\delta\rho^2, \quad (34)$$

$$V_2 = \rho\left(-A_1\rho k^2(z^2-1) - 2A_2\rho^2 k(z^4+1) + \frac{1}{4}(k^2+2)B_{00} + \frac{1}{2}A_1\rho k(z^2+1)\right)\delta^2, \quad (35)$$

$$V_3 = \left(-\frac{1}{8}(3k^2+2)B_{00} - 3A_{3,3k}k\rho^{(3-3k)}(z^6+1) - \frac{1}{4}A_1\rho k(z^2+1) - \frac{1}{2}k^3A_1\rho(z^2+1) - 3k^2\rho^2A_2(z^4-1) + A_2\rho^2k(z^4+1) + \frac{7}{8}A_1\rho k^2(z^2-1)\right)\delta^3. \quad (36)$$

Let us consider the initial stage of the loss of morphological stability of the growing particle upon the onset of perturbation on the boundary and determine the critical particle radius for which the perturbation amplitude will grow. To this end, following [14–16], we have to solve for ρ the equation

$$V_1 = 0. \quad (37)$$

This will determine the critical radius above which the rate of variation of the amplitude of the basic (initially applied) harmonic $\cos k\varphi$ changes sign from minus (corresponding to decay) to plus (corresponding to increase). Obviously, an increase in the amplitude of the basic harmonic will lead to the appearance of secondary harmonics of the other frequencies (see, e.g., Eq. (32)). Thus, from the standpoint of the initial stage of the loss of stability, behavior of the basic harmonic is the determining factor.

Let us seek for a solution of Eq. (37) in the following form:

$$\rho = g + g_1\delta^2, \quad (38)$$

where g and g_1 are the expansion coefficients. In this expression, the term proportional to δ is omitted because its presence in Eq. (37) has no physical meaning for symmetry considerations. Indeed, in the presence of a term proportional to δ, a half-period translation of the $\cos k\varphi$ perturbation (equivalent to alternation in the amplitude sign) would lead to a change in the critical radius for stable growth in the isotropic medium.

Substituting ansatz (38) into the explicit expression (37) and expanding it into series in powers of δ up to the third-order term, we arrive at an expression containing δ to the first and third powers. Grouping and equating the terms at like powers, we obtain two equations for determining g and g_1. The first quantity cannot be expressed explicitly and is determined from the equation

$$(-g^2k - gkL + k^3\alpha + gk^3L + gk - k\alpha)(h^2+1) + g(g-1)(1-h^2) = 0. \quad (39)$$

The second coefficient can be explicitly expressed as

$$g_1 = [(M_1 - M_2)h^8 + (M_3 + M_4)h^6 + M_5h^4 + (M_3 - M_4)h^2 + M_1 + M_2 + 6\alpha^4k^3(5k^2 - k^4 - 2)][8g[(2\alpha k - g)(M_6 - M_7)h^8 + ((2\alpha k - g)h^6 + (2\alpha k + g)h^2)M_8 + ((2\alpha k + g)(M_6 - M_7) + (2\alpha k - g) \times (M_6 + M_7))h^4 + (2\alpha k + g)(M_6 + M_7)]]^{-1}, \quad (40)$$

where $h = g^k/\rho_\lambda^k$, $L = \ln(\rho_\lambda/g)$, and the expressions for coefficients M_1, M_2, …, M_8 are presented in the Appendix.

It should be noted that Eq. (39) for g in the diffusion limit ($\alpha \longrightarrow 0$) coincides with an analogous equation derived in [14].

4. DISCUSSION OF RESULTS

Figures 1 and 2 show the plots of the linear radius of stability g and the correction of the second order of smallness g_1 versus the parameters α and k (the value of ρ_λ was selected following [14]). The g and g_1 values numerically calculated for the diffusion-controlled growth ($\alpha = 0$) (see Figs. 1 and 2) coincide with the analogous values reported in [14], which is additional evidence in favor of correctness of our calculation.

According to the results of calculations (Figs. 1a and 2a), the linear radius of stability (g) increases with the harmonic frequency (number) and the parameter α. This behavior can be explained as follows. According to the classical results [1], the main factor responsible for the loss of stability is the inhomogeneity of the con-

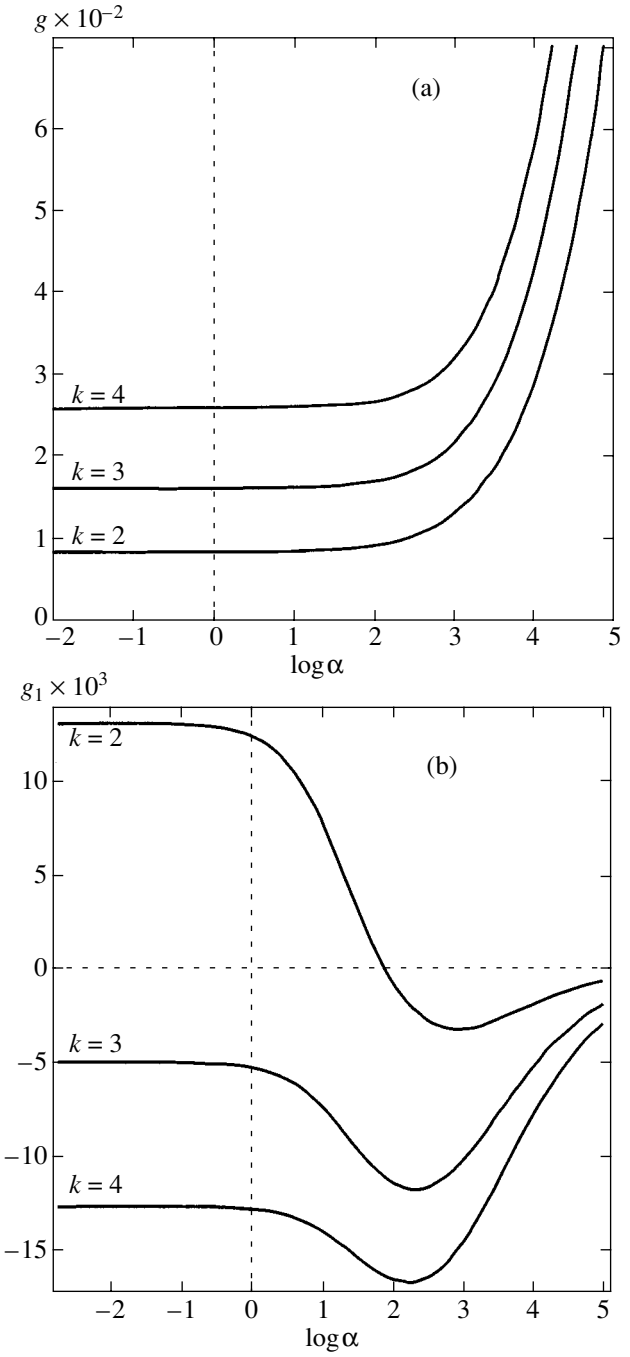


Fig. 1. Plots of (a) the linear critical radius of stability g and (b) the quadratic correction g_1 versus parameter α for various perturbation frequencies (harmonic numbers) k for $\rho_\lambda = 10^8$.

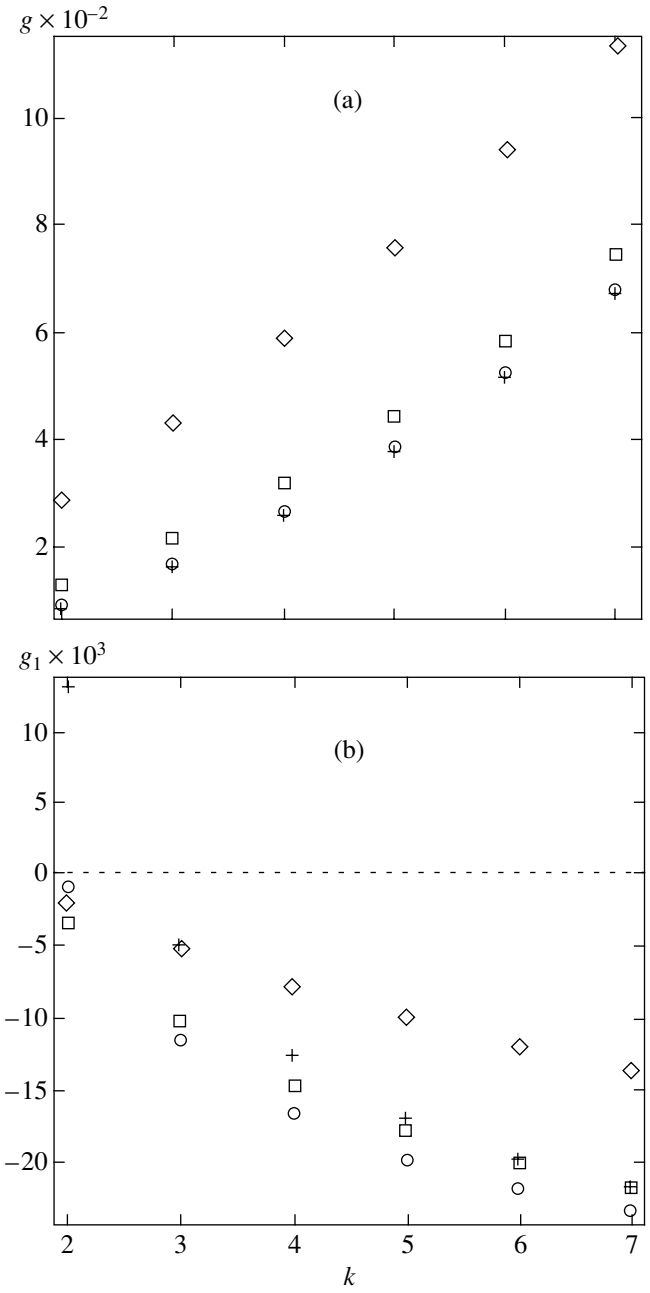


Fig. 2. Plots of (a) the linear critical radius of stability g and (b) the quadratic correction g_1 versus perturbation frequency k for $\rho_\lambda = 10^8$ and various parameters α: (+) 0; (o) 100; (□) 1000; (◇) 10000.

centration field at the crystal boundary: the supersaturation increases with distance from the crystal. For this reason, a perturbation arising on the crystal surface occurs under more favorable conditions and can develop more rapidly. The main stabilizing factor is the curvature (i.e., the surface energy): the greater the curvature at the appearing protrusion, the more readily it can be dissipated. Thus, for two nuclei with perturbations of the same amplitude but different frequencies, the curvature is greater for the nucleus with a higher frequency and, accordingly, the stabilizing factor is stronger here. As a result, the nucleus loses stability under this perturbation at a greater critical radius. As the α value increases, the field at the crystal becomes more homogeneous and, accordingly, the destabilizing factor decreases and the critical radius g increases.

In contrast to the behavior of g, the character of variation and the sign of g_1 are not as obvious. As can be

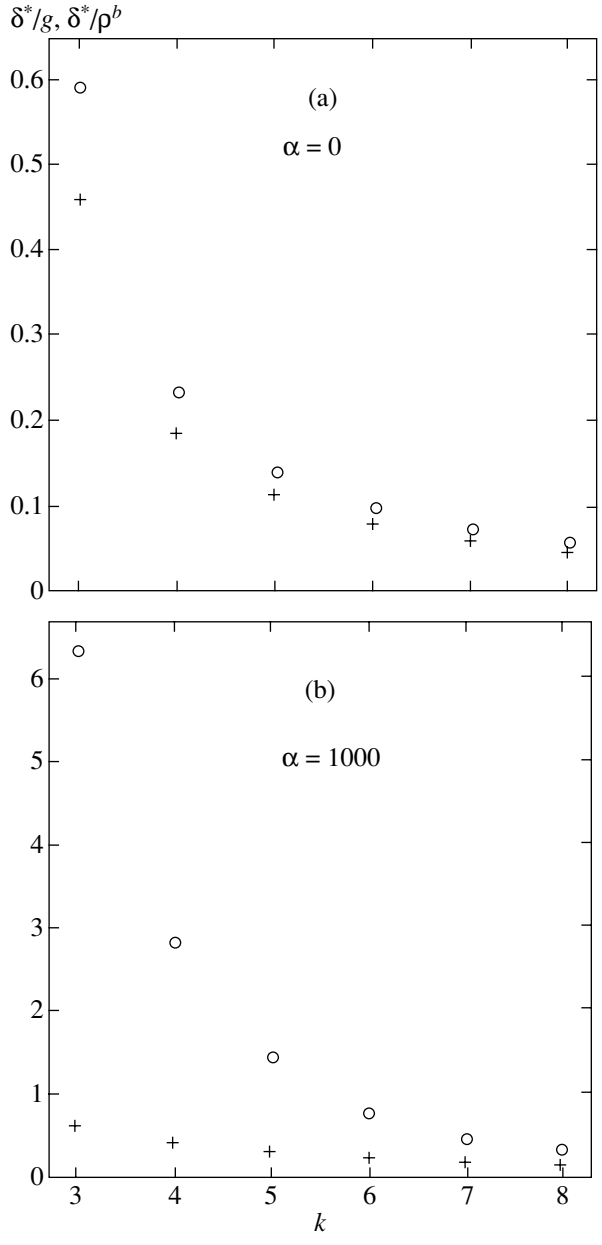


Fig. 3. Plots of (+) δ^*/g and (o) δ^*/ρ^b versus perturbation frequency k for $\rho_\lambda = 10^8$.

seen from Figs. 1b and 2b, an increase in the perturbation amplitude leads in most cases to a decrease in the critical radius as compared to that in the linear case (an increase is observed only in the diffusion-controlled growth regime for the harmonic with $k = 2$). However, with an increase in the perturbation amplitude (for fixed parameters of α and k), the perturbation falls within a region of higher supersaturation (i.e., a stronger destabilizing factor) and greater curvature (i.e., a stronger stabilizing factor). The fact that, in the competition of two factors, the former almost always predominates (even in the kinetically controlled growth regime with $\alpha \gg 1$) is a very interesting result.

A decrease in the critical radius with increasing perturbation amplitude allows us to explain the phenomenon of coexistence of two different morphological phases, which is frequently observed in experiment. Indeed, consider two round crystal nuclei growing at a sufficiently large distance from each other in a supersaturated solution. Since the appearance of fluctuations in solutions is, in the general case, a stochastic process, let a fluctuation of small but finite amplitude δ to arise at one of the two particles. According to the above calculations, this nucleus will lose stability when its radius will increase to the critical value of ρ, while the other nucleus may still retain its round shape. Thus, in the growth interval from ρ to g, we can observe both the crystals of round shape and those with developing perturbations corresponding to the kth harmonic. When the radius of a round nucleus will increase to g according to the linear analysis, it will lose stability to even a negligibly small fluctuation.

This provides ground for one of the main assumptions used in our calculations of the boundaries of metastable regions [11–13]. It would be also interesting to use the above results for qualitatively evaluating the perturbation amplitude for which the crystal radius reaches a value ρ^b called binodal. This value was previously determined via a calculation of the production of entropy based on the linear analysis [13] as

$$\rho^b = \frac{1}{2}\left\{1 - \frac{\alpha k}{2k-1} + \frac{2A_\lambda k(k^2-1)}{2k-1} + \sqrt{\left[1 - \frac{\alpha k}{2k-1} + \frac{2A_\lambda k(k^2-1)}{2k-1}\right]^2 + 4\alpha\frac{k(2k^2-1)}{2k-1}}\right\}. \quad (41)$$

According to formula (38), the amplitude of this perturbation is

$$\delta^* = \sqrt{(\rho^b - g)/g_1}. \quad (42)$$

Figure 3 shows plots of the ratios of the perturbation amplitude δ^* to the above characteristic dimensions g and ρ^b versus k for the diffusion and kinetic growth regimes. These data reveal an interesting fact: the size ρ^b found in [13] corresponds to the perturbation amplitudes comparable to the radius of stability of the crystal. This result indicates that, from the standpoint of the above weakly nonlinear analysis, ρ^b represents the critical radius with respect to the maximum perturbations δ^* of the crystal radius (whereby the perturbation amplitude is comparable to the crystal size), which justifies use of the name "binodal radius."

Let us specially consider the cases of $k = 1$ and $k = 2$. According to the results of numerical analysis, representing the solution in the form (38) for $k = 1$ is not justified because a correction to the solution in the linear

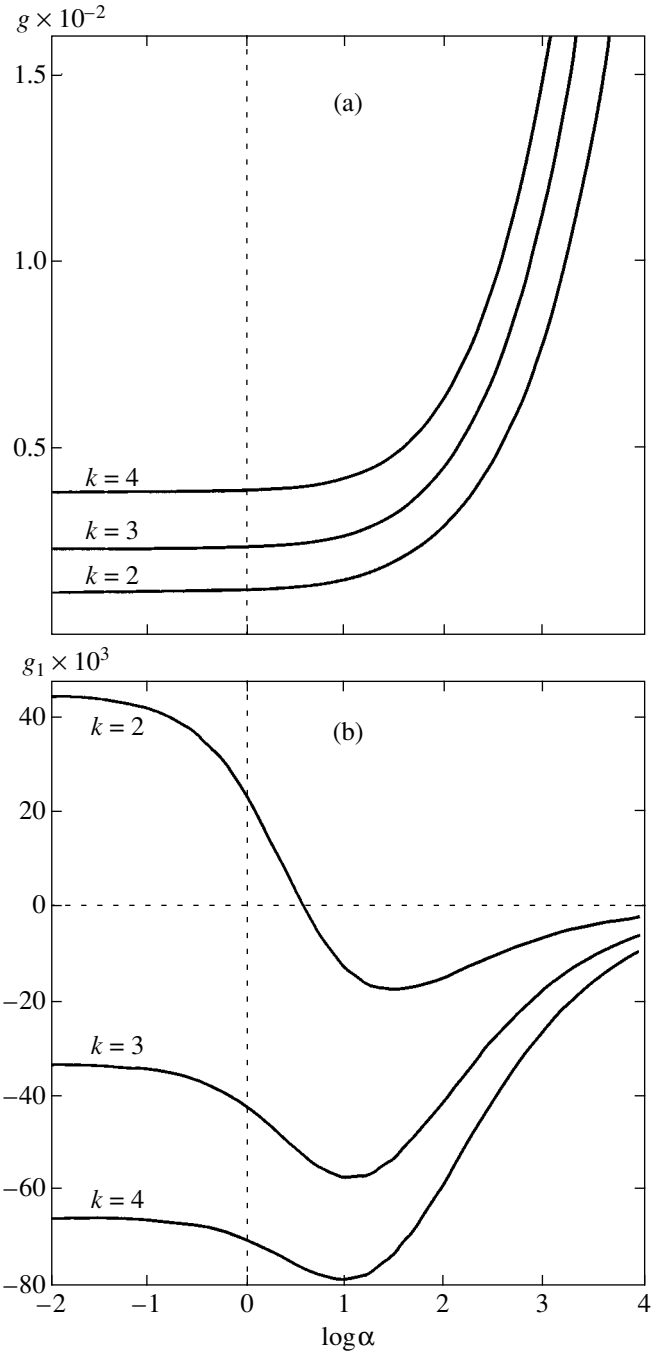


Fig. 4. Plots of (a) the linear critical radius of stability g and (b) the quadratic correction g_1 versus parameter α for various perturbation frequencies (harmonic numbers) k for $S = 0.05$.

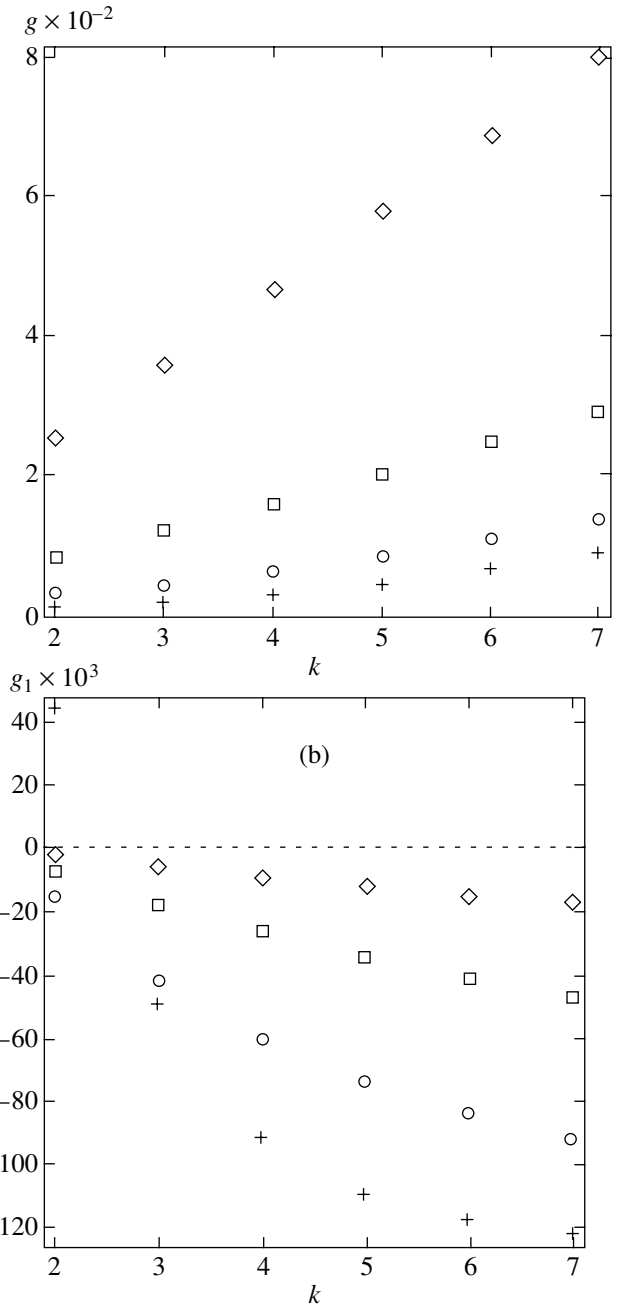


Fig. 5. Plots of (a) the linear critical radius of stability g and (b) the quadratic correction g_1 versus perturbation frequency k for $S = 0.05$ and various parameters α: (+) 0; (o) 100; (□) 1000; (◇) 10000.

approximation is much greater: $g_1\delta^2/g \sim 10^{17}\delta^2$ for $\alpha = 0$–100). This violates the main assumption of perturbation theory (introducing a so-called inhomogeneity of the expansion [20]). The perturbation with $k = 2$, for which the g_1 correction is positive at $\alpha < 73$, is the next neighboring case to that of $k = 1$. Does this fact account for the special behavior (see Figs. 1b and 2b) because the point occurs in the region of inhomogeneous expansion [20], or is everything correct from the mathematical standpoint and the result has a physical meaning? These questions require additional investigation.

The data considered above (Figs. 1 and 2) are related to a constant ρ_λ, which is not convenient from the standpoint of experiment. As was pointed out in Section 2, the quantity ρ_λ appears as a fitting parameter on the passage from nonstationary to stationary diffusion

problem. This parameter is directly related to the quantity S, which has a clear physical meaning and can be readily determined for any supersaturation solution. In connection with this, Figs. 4 and 5 present analogs of Figs. 1 and 2, respectively, plotted as a fixed S. A comparison between these representations shows that the character of dependences is qualitatively the same.

5. CONCLUSIONS

We have carried out a weakly nonlinear analysis of morphological stability for a round crystal growing in an arbitrary regime from a supersaturated solution. It was found that, similar to the diffusion-controlled growth considered in [14–16], an increase in the amplitude of perturbations leads to a decrease in the critical radius of stable crystal. This behavior is an additional argument in favor of the hypothesis that such a dependence of the critical radius on the perturbation amplitude is rather universal and may be considered as evidence that this nonequilibrium transition belongs to first-order phase transitions.

We would like to make a final remark concerning the possibility of comparing the values of critical radii calculated in this study to the experimental data. At the present time, it is hardly possible to speak of such verification due to both the technical complexity of the experiment and the still relatively large number of simplifying assumptions made in the theory. The main task at this stage is to provide qualitative explanation and to propose a method for analytically calculating the experimentally observed phenomenon of coexisting morphologies during crystallization under identical conditions. We believe that the results reported above are an important step in this direction.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to E.F. Lelikova for her attention and numerous consultations on mathematical aspects of this study. Also gratefully acknowledged are the permanent interest, support, and creative criticism of V.D. Seleznev.

APPENDIX

$$K_0 = \frac{1}{\rho},$$

$$K_1 = \frac{(k^2-1)\cos k\varphi}{\rho^2},$$

$$K_2 = \frac{1}{\rho^3}\left[\frac{k^2}{2} + \left(1 - \frac{5}{2}k^2\right)\cos^2 k\varphi\right],$$

$$K_3 = \frac{1}{\rho^4}\left[\left(\frac{3}{2}k^4 + \frac{9}{2}k^2 - 1\right)\cos^3 k\varphi - \frac{3}{2}k^2(1+k^2)\cos k\varphi\right],$$

$$A_2 = \frac{2A_1\rho k(\alpha+\rho-2\alpha k)z^2 + B_{00}(2\alpha+\rho+\alpha k^2) + 2A_1\rho k(\alpha+\rho+2\alpha k)}{4\rho^2(z^4(2\alpha k-\rho)+(2\alpha k+\rho))},$$

$$\begin{aligned}A_{3,k} = \{&8A_2\rho^2 k(\alpha+\rho-\alpha k)z^4 \\ &+ A_1\rho k(3\rho k - 6\alpha - 3\rho + 5\alpha k)z^2 \\ &+ 8A_2\rho^2 k(\alpha+\rho+\alpha k) \\ &- A_1\rho k(3\rho k + 6\alpha + 3\rho + 5\alpha k) + 8B_{20}\rho^2(\alpha+\rho) \\ &+ B_{00}(3\alpha k^2 - 6\alpha - 2\rho) \\ &+ 3\Delta(5k^2 - k^4 - 2)\}\rho^{k-3}/\{8(z^2(\alpha k-\rho)+(\alpha k+\rho))\},\end{aligned}$$

$$\begin{aligned}A_{3,3k} = \{&24A_2\rho^2 k(\alpha+\rho-3\alpha k)z^4 \\ &+ 3A_1\rho k(\rho k - 2\alpha - \rho + 7\alpha k - 4\alpha k^2)z^2 \\ &+ 24A_2\rho^2 k(\alpha+\rho+3\alpha k) \\ &- 3A_1\rho k(\rho k + 2\alpha + \rho + 7\alpha k + 4\alpha k^2) \\ &- B_{00}(9\alpha k^2 + 6\alpha + 2\rho), \\ &+ 3\Delta(9k + 3k^4 - 2)\}\rho^{3k-3}/\{24(z^6(3\alpha k-\rho) \\ &+ (3\alpha k+\rho))\},\end{aligned}$$

$$\begin{aligned}M_1 = &-3\alpha^4 k^7 - 6\alpha^4 k^3 + 15\alpha^4 k^5 \\ &+ 18g^4 kL - 10g^5 kL + 12\alpha g^3 k - 10\alpha^2 g^3 k^3 \\ &- 3\alpha g^3 k^3 - 13g^4 Lk^3 - 6g^4 L^2 k^5 + 34g^4 L^2 k^3 \\ &- 16g^4 L^2 k - 7g^4 k^3\alpha - 4g^4 k\alpha + 3g^5 Lk^3\end{aligned}$$

$$-6g^4k - 12L\alpha^2k^5g^2 - 48\alpha g^3kL + 78\alpha g^3Lk^3$$
$$-4\alpha g^3L^2k^5 + 4\alpha g^3L^2k^3 - 4\alpha g^3Lk^5 - 4g^4k^3\alpha L$$
$$+8g^4k\alpha L + 8L\alpha^2k^3g^2 - 24k\alpha^2g^2 + 2\alpha^2k^5g^2$$
$$-4\alpha^3k^5g^2 + 50\alpha^2k^3g^2 + 6g^5k - 12L\alpha^3k^7g$$
$$-24L\alpha^3k^3g + 60L\alpha^3k^5g - 12L^2\alpha^2k^3g^2$$
$$-6L^2\alpha^2k^7g^2 + 30L^2\alpha^2k^5g^2 + 4\alpha^3k^3g - 4\alpha^3k^5g,$$

$$M_2 = 6g^5L - 4g^5k^2 + 4g^4k^2 - 6\alpha g^3 + 2g^4\alpha + 3\alpha^4k^7$$
$$+6\alpha^4k^3 - 15\alpha^4k^5 - 2g^5 + 2g^4 + 7g^4k^2\alpha$$
$$-14g^4k^2\alpha L - 10g^4L - 15\alpha^3k^6g + 2\alpha^3k^4g^2$$
$$+12k^2\alpha^2g^2 - 24\alpha^2k^4g^2 + 6\alpha^2g^3k^2 - 5\alpha g^3k^2$$
$$+4\alpha g^3k^4 - 2g^4L^2k^4 + 2g^4L^2k^2 + 3g^4Lk^2$$
$$+4g^4Lk^4 - 30L\alpha^2k^6g^2 + 144k^4\alpha^2Lg^2 - 60L\alpha^2k^2g^2$$
$$-2\alpha^2g^3k^4L + 69\alpha g^3L^2k^4 - 30\alpha g^3L^2k^2$$
$$+34\alpha g^3Lk^2 - 32\alpha g^3Lk^4 - 15\alpha g^3L^2k^6 + 6g^4k^4\alpha L$$
$$-g^5Lk^2 + 71\alpha^3k^4g - 30\alpha^3k^2g,$$

$$M_3 = 4k(-6k^2\alpha^4 + 15k^4\alpha^4 + 6g^2\alpha^2 + 2g^4L^2 + g^5L$$
$$-3\alpha g^3 + 3g^4\alpha + 12\alpha g^3L - 4g^4L\alpha + 2k^2\alpha^3g^2$$
$$-3k^6\alpha^4 - 6g^4k^2\alpha + 5g^4k^2\alpha L + 2k^2\alpha^2g^3L - 6k^6L\alpha^3g$$
$$-12k^2L\alpha^3g + 30k^4L\alpha^3g - 6k^2L^2\alpha^2g^2 - 3k^6L^2\alpha^2g^2$$
$$+15k^4L^2\alpha^2g^2 - g^4L - \alpha^3k^4g^2 - 14k^2\alpha^2g^2 + 6\alpha^2k^4g^2$$
$$-3\alpha^2g^3k^2 + 6\alpha g^3k^2 + 2g^4L^2k^4 - 7g^4L^2k^2 + 2g^4Lk^2$$
$$-17k^4\alpha^2Lg^2 + 6L\alpha^2k^2g^2 + \alpha^2g^3k^4L - 7\alpha g^3L^2k^4$$
$$+4\alpha g^3L^2k^2 - 29\alpha g^3Lk^2 + 8\alpha g^3Lk^4 - 2g^5Lk^2$$
$$-8\alpha^3k^4g + 2\alpha^3k^2g),$$

$$M_4 = 2g(4g^4k^2 + 2\alpha g^3 - 4g^3k^2 - 10g^3L + 6\alpha^3k^2$$
$$-15\alpha^3k^4 - 6\alpha g^2 - 5\alpha^3k^6 - 2g^4 + 2g^3 + 15k^2\alpha g^2$$
$$+23g^3k^2L - 12g^3L^2k^2 + 18g^3L^2k^4 - 4\alpha k^4g^2 - 4g^3Lk^4$$
$$-10L\alpha^2k^6g - 22k^4\alpha^2Lg + 12L\alpha^2k^2g - 7L^2\alpha k^4g^2$$
$$+6L^2k^2\alpha g^2 - 14k^2\alpha Lg^2 + 32L\alpha k^4g^2 - 5g^2\alpha L^2k^6$$
$$-2k^2\alpha^2g + 12\alpha^2k^4g + 6g^4L - 10k^2\alpha^2g^2 + 8\alpha^2k^4g^2$$
$$-5\alpha g^3k^2 - 13g^4Lk^2 + 8k^4\alpha^2Lg^2$$
$$-10\alpha g^3Lk^2 + 6\alpha g^3Lk^4 + 8\alpha^3k^4g),$$

$$M_5 = 2k(-12k^2\alpha^4 + 30k^4\alpha^4 + 8g^4L^2 + 6g^5L - 8g^4\alpha$$
$$-6g^5 + 6g^4 + 8g^4L\alpha - 6k^6\alpha^4 - 17g^4k^2\alpha - 16g^4k^2\alpha L$$
$$-12k^6L\alpha^3g - 24k^2L\alpha^3g + 60k^4L\alpha^3g - 12k^2L^2\alpha^2g^2$$
$$-6k^6L^2\alpha^2g^2 + 30k^4L^2\alpha^2g^2 - 14g^4L + 6k^2\alpha^2g^2$$
$$+22\alpha^2k^4g^2 - 26\alpha^2g^3k^2 + 27\alpha g^3k^2 + 14g^4L^2k^4$$
$$-22g^4L^2k^2 + 21g^4Lk^2 - 40k^4\alpha^2Lg^2 + 8L\alpha^2k^2g^2$$
$$+4\alpha^2g^3k^4L - 16\alpha g^3L^2k^4 + 4\alpha g^3L^2k^2 - 26\alpha g^3Lk^2$$
$$+36\alpha g^3Lk^4 - 11g^5Lk^2 - 20\alpha^3k^4g + 4\alpha^3k^2g),$$

$$M_6 = \alpha g^3 - g^4 + g^3 - 2\alpha g^2 + \alpha^2k^2g + \alpha g^3k^2 - k^2\alpha g^2$$
$$-2\alpha^3k^2 + 2\alpha^3k^4 - 3g^3L - \alpha g^3Lk^2 + 4L\alpha^2k^4g$$
$$-4L\alpha^2k^2g + 2Lk^2\alpha g^2 + 2L^2\alpha k^4g^2$$
$$-2L^2k^2\alpha g^2 + 2g^4L,$$

$$M_7 = gk(-g^2 + g^3 + 3k^2\alpha^2 - 4\alpha^2$$
$$-2\alpha g^2 - 2g^3L + 3\alpha g + 3k^2g^2L^2 - 3g^2L^2 + 3g^2L$$
$$+g^2\alpha L + 6k^2gL\alpha - 8g\alpha L),$$

$$M_8 = 4g^4Lk^2 + 2\alpha g^3Lk^2 + 2\alpha^2k^2g$$
$$-8L\alpha^2k^2g - 4g^3Lk^2 - 2g^3 + 8Lk^2\alpha g^2$$
$$-4g^4L - 8L\alpha k^4g^2 - 4\alpha^2k^4g + 6g^3L + 4L^2\alpha k^4g^2$$
$$-4g^3L^2k^4 + 2g^4 - 4L^2k^2\alpha g^2 - 6k^2\alpha g^2 + 4\alpha^3k^4$$
$$-4\alpha^3k^2 + 6\alpha g^3k^2 + 4\alpha^2k^2g^2 + 4\alpha g^2$$
$$+8L\alpha^2k^4g + 4g^3L^2k^2 - 2\alpha g^3.$$

REFERENCES

1. W. W. Mullins and R. F. Sekerka, J. Appl. Phys. **34**, 323 (1963).
2. S. R. Coriell and G. B. McFadden, in *Handbook of Crystal Growth*, Ed. by D. T. J. Hurle (North-Holland, Amsterdam, 1993), Vol. 1, Part B, p. 785.
3. S. C. Hardy and S. R. Coriell, J. Cryst. Growth **5**, 329 (1969).
4. P. Oswald, J. Malthete, and P. Pelce, J. Phys. (Paris) **50**, 221 (1989).

5. D. E. Ovsienko, G. A. Alfintsev, G. P. Chemerinskiĭ, *et al.*, Kristallografiya **21**, 801 (1976) [Sov. Phys. Crystallogr. **21**, 454 (1976)].
6. A. A. Shibkov, Yu. I. Golovin, M. A. Zheltov, *et al.*, Kristallografiya **46**, 549 (2001) [Crystallogr. Rep. **46**, 496 (2001)].
7. O. Shochet and E. Ben-Jacob, Phys. Rev. E **48**, 4168 (1993).
8. Y. Sawada, B. Perrin, P. Tabeling, *et al.*, Phys. Rev. A **43**, 5537 (1991).
9. E. Ben-Jacob and P. Garik, Nature **343**, 523 (1990).
10. Y. Sawada, J. Stat. Phys. **34**, 1039 (1984).
11. L. M. Martyushev, V. D. Seleznev, and I. E. Kuznetsova, Zh. Éksp. Teor. Fiz. **118**, 149 (2000) [JETP **91**, 132 (2000)].
12. L. M. Martyushev, I. E. Kuznetsova, and V. D. Seleznev, Zh. Éksp. Teor. Fiz. **121**, 363 (2002) [JETP **94**, 307 (2002)].
13. L. M. Martyushev and E. M. Sal'nikova, Pis'ma Zh. Tekh. Fiz. **28** (6), 57 (2002) [Tech. Phys. Lett. **28**, 242 (2002)].
14. L. N. Brush, R. F. Sekerka, and G. B. McFadden, J. Cryst. Growth **100**, 89 (1990).
15. P. P. Debroy and R. F. Sekerka, Phys. Rev. E **51**, 4608 (1995).
16. P. P. Debroy and R. F. Sekerka, Phys. Rev. E **53**, 6244 (1996).
17. S. R. Coriell and R. L. Parker, J. Appl. Phys. **36**, 632 (1965).
18. S. R. Coriell and S. C. Hardy, J. Res. Natl. Bur. Stand., Sect. A **73**, 65 (1969).
19. A. D. Polyanin, *Handbook on Linear Equations of Mathematical Physics* (Fizmatlit, Moscow, 2001).
20. A. H. Nayfeh, *Introduction to Perturbation Techniques* (Wiley, New York, 1981; Mir, Moscow, 1984).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 98, No. 5, 2004, pp. 997–1001.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 125, No. 5, 2004, pp. 1139–1143.

SOLIDS
Structure

Incommensurate Phases Induced in a Crystal by External Effects

D. G. Sannikov

Institute of Crystallography, Russian Academy of Sciences, Moscow, 117333 Russia
e-mail: sannikov@ns.crys.ras.ru

Received September 5, 2003; in final form, December 12, 2003

Abstract—Lifshitz invariants emerging in the presence of external effects in a crystal are considered for 2D irreducible representations of the D_{4h} class. It is shown that, for a 2D irreducible vector representation, the electric field exceeding a critical value leads to the formation of an incommensurate phase on the phase diagram.

1. INTRODUCTION

In studying incommensurate (*I*) phases and phase transitions (see reviews [1–3]), two types of such transitions are distinguished. The first type is associated with the presence of the Lifshitz (*L*) invariant in the thermodynamic potential, while the second type is characterized by the presence of a Lifshitz-type (*LT*) invariant. Here, we consider the first type of transitions. In the case when a multidimensional representation of the symmetry group of the initial phase in the crystal rules out the existence of an *L* invariant, the incommensurate invariant (abbreviated below as *IL*) can be induced by appropriate effects exerted on a crystal (electric or magnetic fields, mechanical stresses, etc.). The coefficient of the *IL* invariant is the external factor itself. This problem was considered for the first time in [4], where the phase diagram with an *I* phase was constructed from general considerations. However, apparently not a single specific case has been investigates as yet.

The goal of this study is analysis of a 2D vector representation of the crystal symmetry class D_{4h}, the determination of possible *IL* invariants for this representation, and the calculation of the corresponding phase diagrams. To run a few steps ahead, we can state that the induction of *I* phases on these diagram is not as obvious and simple as appears at first glance and the results of our analysis differ from those described in [4].

2. TWO-DIMENSIONAL REPRESENTATIONS OF SYMMETRY CLASS D_{4h}

The choice of the crystal symmetry class D_{4h} is dictated by the existence of two irreducible representations, each of which permits two independent invariants of the second and fourth order in the bases of the representations. The latter circumstance determines the simplest algebra of these representations among all nonone-dimensional representations of crystal symmetry groups. The table shows the representations of class D_{4h}, according to which the components of vector $x_i = \{x, y, z\}$ and rank two tensor $\{u_\alpha\}$ are transformed (we use the matrix form of notation: $\alpha = 1, \ldots, 6$). The polarization vector components P_i, which will be treated as the order parameter for representation of E_u, is transformed like the components of vector x_i; u_α stands for components of the strain tensor u_{ij} (in particular, $u_4 = 2u_{yz}$ and $u_5 = 2u_{zx}$). The existence of two representations instead of one as in classes D_4, C_{4v}, and D_{2d} (with the same independent invariants) rules out the existence of the piezoelectric effect in the initial phase of the crystal, which simplifies the problem.

None of 2D representations of E_u and E_g permits an *L* invariant. We will consider external forces (electric field E_i and mechanical stresses σ_α) that can induce *IL* invariants (we confine our analysis only to invariants linear in forces). Using the table, we can obtain two *IL* invariants for the E_u and E_g representations, respectively:

$$E_x\{P_x, P_y\}_y - E_y\{P_x, P_y\}_x,$$

$$E_x\{u_4, u_5\}_y - E_y\{u_4, u_5\}_x, \qquad (1)$$

$$\{\eta, \xi\}_z \equiv \eta\partial_z\xi - \xi\partial_z\eta, \quad \partial_z \equiv \partial/\partial z.$$

Since invariants (1) contain the electric field components E_x and E_y, we will find the invariants which are linear in E_x, E_y and also linear in P_x, P_y and u_4, u_5. There exists only one such invariant (the energy of interaction of polarization with the electric field):

$$-E_xP_x - E_yP_y. \qquad (2)$$

There are no invariants linear in E_x, E_y and quadratic in P_x, P_y or quadratic in u_4, u_5. We will confine our subse-

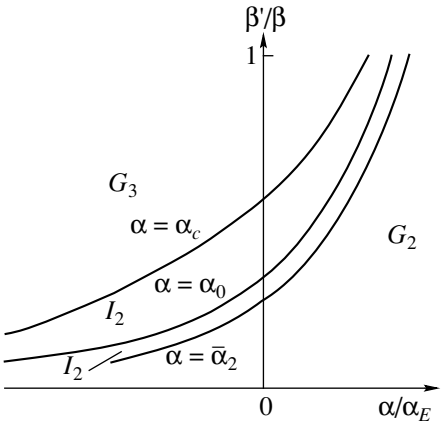


Fig. 1. Phase diagram on the αβ' plane for $F > F_0$ ($\Delta > 0$). The curves demarcate the regions of existence of phases G_2, G_3, and I_2 (see relations (6), (13), and (17)); $\alpha_E \equiv 2(2\beta E^2)^{1/3}$.

quent analysis to the vector representation of E_u only.

3. HOMOGENEOUS THERMODYNAMIC POTENTIAL

It is convenient to represent components P_x and P_y of the order parameter in polar coordinates:

$$P_x = \rho\cos\varphi, \quad P_y = \rho\sin\varphi. \tag{3}$$

In the simplest case and in the absence of fields ($E_x = 0$ and $E_y = 0$), the homogeneous thermodynamic potential (containing no gradient invariants) has the form

$$\Phi = \alpha\rho^2 + \beta\rho^4 + \beta'\rho^4\cos 4\varphi. \tag{4}$$

We assume that we are dealing with second-order transitions; i.e., coefficients $\beta > 0$ and $\beta - |\beta'| > 0$.

From the conditions of minimum for potential (4) in variables ρ and φ, we obtain the three solutions

$$G_0\colon \rho = 0,$$
$$G_1\colon \cos 4\varphi = -1, \quad \rho^2 = -\alpha/(\beta - \beta'), \tag{5}$$
$$G_2\colon \cos 4\varphi = 1, \quad \rho^2 = -\alpha/(\beta + \beta'),$$

corresponding to three phases with symmetry groups

Irreducible representations of crystal symmetry class D_{4h}

A_{1g}	$u_1 + u_2$	A_{1u}	
	u_3		
A_{2g}		A_{2u}	z
E_u	x	E_g	u_4
	y		$-u_5$
B_{1g}	$u_1 - u_2$	B_{1u}	
B_{2g}	u_6	B_{2u}	

$G_0 = D_{4h} = 4/mmm$ (initial phase), $G_1 = C_{2v} = 2_{xy}\ \ m_z$, and $G_2 = C_{2v} = 2_x m_y m_z$ (see, for example, [5]). Phase G_0 is stable for $\alpha > 0$, and G_1 is stable for $\alpha < 0$ and $\beta' > 0$, while G_2 exhibits stability for $\alpha < 0$ and $\beta' < 0$ (the same notation of phases and their groups should not lead to misunderstanding). Group $G_3 = C_s = m_z$ is one more subgroup of group G_0 (as well as of groups G_1 and G_2) in the E_u representation. The solution for phase G_3 can be obtained if we take into account the second-order invariant in $\rho^4\cos 4\varphi$ in expression (4). It should be noted that phase G_1, as well as phase G_2, has four different domains with different orientations of components P_x and P_y (different values of angle φ; see relations (5)).

We now supplement potential (4) with invariant (2). In the case of a one-component order parameter (e.g., P_z) transformed in accordance with a one-dimensional representation (A_{2u}; see table), field E_z conjugate to it leads to disappearance of the second-order phase transition. In other words, the symmetry groups of the initial and polar phases become identical. Characteristic anomalies in physical quantities are smoothed. In the case of a 2D representation, however, we can always choose the field orientation such that one of the second-order phase transitions is preserved. Fields with orientation $E_x = 0$ or $E_y = 0$ lower the symmetry of the G_0 and G_1 phases to the symmetry of the G_2 and G_3 phases, respectively. Consequently, transition $G_{0(2)} \longrightarrow G_{1(3)}$ is preserved (the index of the group to which the group of the corresponding phase is lowered in the presence of the field is given in parentheses). Fields with orientation $E_x = \pm E_y$ lower the symmetry of the G_0 and G_2 phases to the symmetry of the G_1 and G_3 phases, respectively. Consequently, transition $G_{0(1)} \longrightarrow G_{2(3)}$ is preserved. It is sufficient to consider one of these cases. We choose $E_y = 0$ and $E_x = E$.

The phase diagram corresponding to potential (4), (2) is depicted in Fig. 1 (so far we ignore the presence of the I phase, which will follow from the subsequent analysis). The curve $\alpha = \alpha_0$ describing the second-order phase transition between phases G_2 and G_3 (here, we simplify the notation of phases: $G_2 \equiv G_{0(2)}$ and $G_3 \equiv G_{1(3)}$) is defined by the expression

$$\alpha = \frac{1}{2}(3\beta' - \beta)F^2 \equiv \alpha_0, \quad \rho = \frac{1}{2}F \equiv \rho_0, \tag{6}$$
$$E = 2\beta' F^3,$$

where the expression for ρ on this curve is also given and a convenient new notation F for field E is introduced (we note that the value of F depends on β' and that $\beta' > 0$; see Fig. 1).

The solution for phases G_2 and G_3 in field E can be obtained in a small neighborhood of transition curve (6).

In the G_2 phase, one of four domains is preserved (see above), in which vector P_i is directed along field E_i:

$$\sin\varphi = 0, \quad \cos\varphi = 1, \quad \rho = \rho_0(1-\Delta_2),$$
$$\alpha - \alpha_0 = (\beta + 3\beta')F^2\Delta_2, \tag{7}$$
$$\Phi_2 = -\frac{1}{16}(\beta + 9\beta')F^4 + \frac{1}{4}(\alpha-\alpha_0)F^2 - \frac{(\alpha-\alpha_0)^2}{4(\beta+3\beta')}.$$

Here, we also have the expression for thermodynamic potential Φ_2 in the form of a series expansion in $\alpha - \alpha_0$ (to be more precise, in Δ_2). The region of validity of solutions (7) is defined by the inequality $\Delta_2 \ll 1$. In the G_3 phase, two of four domains are preserved:

$$\sin^2\varphi = \frac{6}{5}\Delta_3, \quad \rho = \rho_0(1+\Delta_3),$$
$$\alpha_0 - \alpha = \frac{1}{5}(5\beta - 3\beta')F^2\Delta_3, \tag{8}$$
$$\Phi_3 = -\frac{1}{16}(\beta + 9\beta')F^4 - \frac{1}{4}(\alpha_0-\alpha)F^2 - \frac{5(\alpha_0-\alpha)^2}{4(5\beta-3\beta')}.$$

Solutions (8) are valid for $\Delta_3 \ll 1$.

It can be seen from relations (7) and (8) that the second-order phase transition $G_2 \longrightarrow G_3$ consists in the rotation of vector P_i (which is directed along the x axis in the G_2 phase) through small angles $\pm\varphi$ about the x axis; in other words, the component P_y having opposite directions in the two domains of phase G_3 appears in this phase.

4. PHASE TRANSITION $G_2 \longrightarrow I_2$

We now supplement potential (4), (2) with *IL* invariant (1) with coefficient σ,

$$\sigma E\{P_x, P_y\}_y, \tag{9}$$

as well as with the invariants quadratic in the derivatives and in components P_x and P_y:

$$\delta_1[(\partial_x P_x)^2 + (\partial_y P_y)^2] + \delta_2[(\partial_y P_x)^2 + (\partial_x P_y)^2] + 2\delta_3(\partial_x P_x)(\partial_y P_y) + 2\delta_4(\partial_y P_x)(\partial_x P_y) \tag{10}$$

(see, for example, [6], where both terms, the *L* invariants and *LT* invariant, were introduced simultaneously for the first time).

We analyze the loss of stability in the G_2 and G_3 phases relative to harmonic displacements, which determine the possibility of transitions from these phases to *I* phases. We write P_x and P_y in the form

$$P_x = \rho\cos\varphi + \rho_1\cos(q_x x + q_y y),$$
$$P_y = \rho\sin\varphi + \rho_2\cos(q_x x + q_y y + \psi). \tag{11}$$

Quantities ρ and φ are defined in accordance with relations (7) and (8). It should be noted that, in the presence of field E, quantities P_x and P_y are transformed according to different 1D representations of the G_2 and G_3 groups; for this reason, the values of ρ_1 and ρ_2, as well as the constant phase shifts ψ in relations (11), are different.

We substitute expressions (11) into thermodynamic potential (2), (4), (9), and (10), integrate the result with respect to coordinates x and y, and minimize the part of the integral obtained, which is quadratic in ρ_1 and ρ_2, with respect to ρ_1, ρ_2, q_x, q_y, and ψ. Calculations lead to the following results for phases G_2 and G_3.

Let us first consider the G_2 phase. This phase loses stability relative to harmonic displacements (11) (and experiences a second-order phase transition to phase I_2) for values of F exceeding the critical value F_0:

$$F_0^4 = (\beta + 3\beta')\delta_1/(2\beta'\sigma)^2. \tag{12}$$

Wave vector q_i of the superstructure formed in this way and the stability loss ($\alpha = \bar{\alpha}$) of the G_2 phase are determined by the expressions

$$q^2 = (\beta + 3\beta')F_0^2\Delta/\delta_2 = q_y^2, \quad q_x = 0, \quad \psi = 0,$$
$$\Delta \equiv (F^2 - F_0^2)/F_0^2, \quad \alpha = \bar{\alpha}_2 \equiv \alpha_0 + \bar{\Delta}_2, \tag{13}$$
$$\bar{\Delta}_2 = (\beta + 3\beta')^2\delta_1 F_0^2\Delta^2/6\beta'\delta_2,$$

which were derived for $\Delta \ll 1$.

The phase transition from the G_2 phase to incommensurate phase I_2 involves the emergence of harmonic components P_x and P_y with amplitudes ρ_1 and ρ_2, respectively. It should be emphasized that the homogeneous component P_x, which is present in the G_2 phase, is preserved in phase I_2; in other words, the I_2 phase has a peculiar structure (see relations (11)). (On possible incommensurate polar phases, see [7].)

In the I_2 phase, potential Φ_2 (7) associated with the presence of a homogeneous part of component P_x is supplemented with potential $\bar{\Phi}_2$ associated with inhomogeneous (harmonic) parts of components P_x and P_y with amplitudes ρ_1 and ρ_2:

$$\bar{\Phi}_2 = -\frac{6\beta'^2(\bar{\alpha}_2 - \alpha)^2}{(\beta+\beta')(\beta+3\beta')^2}. \tag{14}$$

5. PHASE TRANSITIONS $G_3 \longrightarrow I_3$ AND $I_2 \longrightarrow G_3$

Let us now consider the G_3 phase. Performing calculations analogous to those for phase G_2, we obtain the following results. Phase G_3 loses stability to harmonic displacements (11) (and experiences a second-order phase transition to phase I_3) for values of field F exceeding the critical value F_0 [see (12)]. The value of wave vector q of the superstructure formed in this way

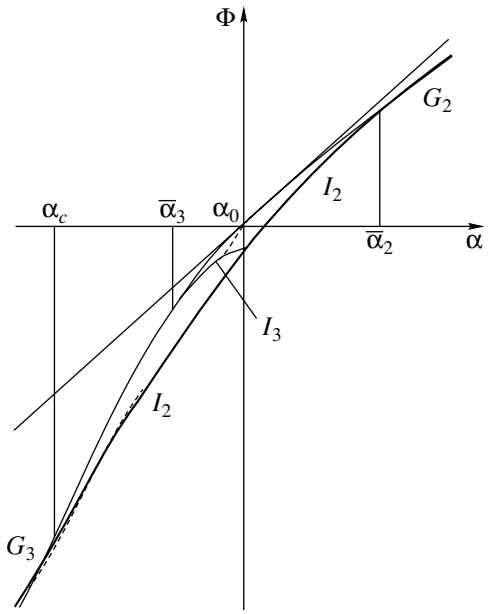


Fig. 2. Dependence of the thermodynamic potentials of phases G_2, G_3, I_2, and I_3 on α for $F > F_0$ ($\Delta > 0$) (see relations (7), (8), (14), and (16)).

coincides with expression (13). Vector q_i is found to be perpendicular to vector P_i in each of the two domains of phase I_3 (note that wave vector q_i is perpendicular to vector P_i in phase G_2 as well):

$$q_x = -q\sin\varphi, \quad q_y = q\cos\varphi,$$

$$\alpha = \bar{\alpha}_3 \equiv \alpha_0 - \bar{\Delta}_3,$$

$$\bar{\Delta}_3 = \frac{(\beta + 3\beta')^2\delta_1 F_0^2\Delta^2}{12\beta'\delta_2} = \frac{1}{2}\bar{\Delta}_2, \tag{15}$$

$$\sin^2\varphi = \frac{6\bar{\Delta}_3}{(5\beta - 3\beta')F_0^2}, \quad \sin^2\psi = \frac{(\beta - 3\beta')^2\Delta}{2\beta'(5\beta - 3\beta')}.$$

These expressions, as well as relations (13), were derived for $\Delta \ll 1$; in formulas (15), Δ coincides with that in formulas (13). Angle φ, as well as phase shift ψ, assumes two values corresponding to two domains of phase I_3.

The phase transition from phase G_3 to incommensurate phase I_3 involves the emergence of harmonic parts in components P_x and P_y with amplitudes ρ_1 and ρ_2. In domains of phase I_3, homogeneous component P_x and P_y, which are present in domains of phase G_3, are preserved. In contrast to the I_2 phase, phase I_3 experiences a constant phase shift ψ between the harmonic components P_x and P_y. The structure of phase I_3 turns out to be even more peculiar and intricate than the structure of phase I_2.

In phase I_3, potential Φ_3 (8) associated with the presence of homogeneous parts of components P_x and P_y is supplemented with potential $\bar{\Phi}_3$ associated with inhomogeneous (harmonic) parts of components P_x and P_y with amplitudes ρ_1 and ρ_2:

$$\bar{\Phi}_3 = -\frac{24\beta'^2(\alpha - \bar{\alpha}_3)^2}{(\beta + \beta')(\beta + 3\beta')^2}. \tag{16}$$

Although the formation of phases I_2 and I_3 is almost symmetric (cf. relations (13) and (15)), phase I_3 turns out to be metastable in the entire range where it exists. Figure 2 shows the dependence of the thermodynamic potentials of phases G_2, G_3, I_2, and I_3 on α. The potential of the I_2 phase passes below the potential of the I_3 phase. The intersection of the potential of the I_2 phase with the potential of the G_3 phase, which is determined by the equality $\Phi_2 + \bar{\Phi}_2 = \Phi_3$ (see relations (7), (8), and (14)) takes place at $\alpha = \alpha_c$, where

$$\alpha_c = \alpha_0 - \bar{\Delta}_2\left[\sqrt{\frac{3(\beta + \beta')(\beta + 3\beta')}{4\beta'(5\beta - 3\beta')}} - 1\right]^{-1}. \tag{17}$$

It should be noted that we are using here the one-harmonic approximation for phases *I*. Analysis shows that the inclusion of higher harmonics affects only slightly the dependence of the potential of phase I_2 on α, while the corresponding dependence for the potential of the I_3 phase changes significantly. The dashed curve in Fig. 2 shows the predicted dependence of the potentials of phases I_2 and I_3 on α.

6. PHASE DIAGRAMS

Thus, there exists a critical field $E_0 = 2\beta' F_0^3$ (see relation (6)) above which ($E > E_0$) the incommensurate phase I_2 appears on the phase diagram. This is due to the existence of *IL* invariant (1), while the existence of the critical field is associated with the presence of invariant (2), i.e., with the linear interaction of the field with the order parameter.

It is convenient to plot the phase diagram with the I_2 phase on the αβ' plane for $F > F_0$ (see Fig. 1). Phase I_2 lies on both sides of the line $\alpha = \alpha_0$ (see relation (6)). The phase diagram plotted on the αF^2 plane is even more illustrative (Fig. 3). On the diagrams depicted in Figs. 1 and 3, the domains of phase I_2 expand with increasing F^2 in proportion to $(F^2 - F_0^2)^2$ (see relations (13) and (17)).

It should also be verified that possible but disregarded invariants cannot appreciably affect the results

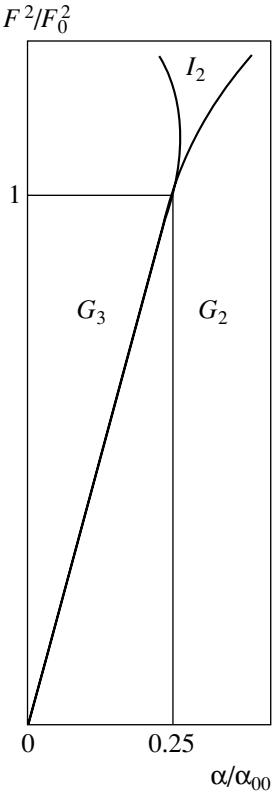


Fig. 3. Phase diagram on the αF^2 plane ($3\beta' > \beta$). The curves demarcate the regions of existence of phases G_2, G_3, and I_2 (see relations (6), (13), and (17)); $\alpha_{00} \equiv 2\beta F_0^2$.

obtained above. The potential can be supplemented with homogeneous invariants of the form

$$(u_1 \pm u_2)(P_x^2 \pm P_y^2), \quad u_3(P_x^2 + P_y^2), \quad u_6(P_xP_y + P_yP_x). \tag{18}$$

However, the elimination of these invariants by varying the potential in variables $u_1 \pm u_2$, u_3, and u_6 would result to renormalization of coefficients of β and β', which can be assumed to have been performed in relation (4).

The potential can be supplemented with invariants of the form

$$(E_x^2 \pm E_y^2)(P_x^2 \pm P_y^2), \quad (E_xE_y + E_yE_x)(P_xP_y + P_yP_x), \tag{19}$$

which are quadratic in the field and quadratic in P_x and P_y, as well as with invariants of the type

$$(E_xP_x \pm E_yP_y)(P_x^2 \pm P_y^2), \quad (E_xP_y + E_yP_x)(P_xP_y + P_yP_x), \quad (E_x^2 \pm E_y^2)(E_xP_x \pm E_yP_y), \quad (E_xE_y + E_yE_x)(E_xP_y + E_yP_x), \tag{20}$$

which are linear in the field and cubic in P_x and P_y, along with those cubic in field and linear in P_x and P_y.

All invariants (19) and (20) are characterized by higher orders in the field or in the components of the order parameter as compared to the invariants included in the potential. In these invariants, series expansions can be carried out. Calculations show that this does not significantly change the results obtained above.

ACKNOWLEDGMENTS

The author is grateful to V.A. Golovko for fruitful discussions.

This study was supported by the Russian Foundation for Basic Research (project no. 03-02-16104).

REFERENCES

1. *Incommensurate Phases in Dielectrics*, Vol. 1: *Fundamentals*, Ed. by R. Blinc and A. P. Levanyuk (North-Holland, Amsterdam, 1986).
2. *Incommensurate Phases in Dielectrics*, Vol. 2: *Materials*, Ed. by R. Blinc and A. P. Levanyuk (North-Holland, Amsterdam, 1985).
3. H. Z. Cummins, Phys. Rep. **185**, 211 (1990).
4. I. M. Vitebskiĭ, Zh. Éksp. Teor. Fiz. **82**, 357 (1982) [Sov. Phys. JETP **55**, 390 (1982)].
5. V. Janovec, V. Dvorak, and J. Petzelt, Czech. J. Phys., Sect. B **25**, 1362 (1975).
6. V. Kopsky and D. G. Sannikov, J. Phys. C **10**, 4347 (1977).
7. V. A. Golovko, Fiz. Tverd. Tela (Leningrad) **23**, 1643 (1981) [Sov. Phys. Solid State **23**, 958 (1981)].

Translated by N. Wadhwa