*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 197–210.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 227–241.*
*Original English Text Copyright © 2001 by Nikishov.*

# Vector Boson in the Constant Electromagnetic Field¶

**A. I. Nikishov**
*Tamm Department of Theoretical Physics, Lebedev Physical Institute,*
*Russian Academy of Sciences, Moscow, 117924 Russia*
*e-mail: nikishov@lpi.ru*

Received March 2, 2001

**Abstract**—The propagator and the complete sets of in- and out-solutions of the wave equation, together with the Bogoliubov coefficients relating these solutions are obtained for the vector $W$-boson (with the gyromagnetic ratio $g = 2$) in a constant electromagnetic field. When only the electric field is present, the Bogoliubov coefficients are independent of the boson polarization and are the same as for the scalar boson. For the collinear electric and magnetic fields, the Bogoliubov coefficients for states with the boson spin perpendicular to the field are again the same as in the scalar case. For the $W^-$ spin parallel (antiparallel) to the magnetic field, the Bogoliubov coefficients and the one-loop contributions to the imaginary part of the Lagrange function are obtained from the corresponding expressions for the scalar case by the substitution $m^2 \longrightarrow m^2 + 2eH$ ($m^2 \longrightarrow m^2 - 2eH$). For the gyromagnetic ratio $g = 2$, the vector boson interaction with the constant electromagnetic field is described by the functions that can be expected by comparing the scalar and Dirac particle wave functions in the constant electromagnetic field. *© 2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

Vector bosons occupy an intermediate place between low-spin particles (with the spins 0 and 1/2) and higher-spin particles. They can therefore share some of the problems encountered in considering higher-spin particle interactions with a strong electromagnetic field. The most conspicuous feature of the vector boson interaction in the case of $g = 2$ is the appearance of tachyonic modes in the overcritical magnetic field. The ways to deal with this problem in the framework of non-abelian theories are analyzed in [1]. But are there any others? According to [2], problems in treating the pair production of the electric field by diagonalizing the Hamiltonian precisely occur for $g = 2$. This is surprising in view of a successful calculation of the Lagrange function of the constant field in the one-loop approximation [3]. We calculate the pair production by the constant field using the Bogoliubov coefficients (which contain all the information about this process); as expected, the results obtained are in agreement with those in [3] and [4].

## 2. VECTOR BOSON
## IN THE CONSTANT ELECTRIC FIELD

We assume $\eta_{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1)$ and set $e = |e|$. The wave function of the $W^-$ boson (with $g = 2$) in a source-free space (where $\partial_\mu F^{\mu\nu} = 0$) satisfies the equation [1, 5]

$$(-D_\sigma D^\sigma + m^2)\psi_\mu - 2ieF_{\mu\nu}\psi^\nu = 0 \tag{1}$$

and the constraint

$$D_\mu \psi^\mu = 0, \quad D_\mu = \partial_\mu + ieA_\mu. \tag{2}$$

With the vector potential chosen such that $A_3 = -Et$ and $A_1 = A_2 = A_0 = 0$, it follows from (1) that $\psi^1$ and $\psi^2$ satisfy the same equation as in scalar case,

$$(-D^2 + m^2)\psi^{1,2} = 0. \tag{3}$$

For $\psi^3$ and $\psi^0$, it follows from (1) that

$$\begin{aligned}(-D^2 + m^2)\psi^3 - 2ieE\psi^0 &= 0,\\ (-D^2 + m^2)\psi^0 - 2ieE\psi^3 &= 0.\end{aligned} \tag{4}$$

Introducing $\psi^\pm = \psi^0 \pm \psi^3$, we rewrite Eqs. (4) as

$$(-D^2 + m^2 \mp 2ieE)\psi^\pm = 0, \tag{5}$$

which can be obtained from (3) by the substitution $m^2 \longrightarrow m^2 \mp 2ieE$. We see that the vector boson wave function can be obtained from the corresponding scalar boson wave function by simple rules.

We now do this. We let ${}_+\psi_{\mathbf{p}}$ denote the positive-frequency in-solution for the (negatively charged) scalar boson. The subscript $\mathbf{p} = (p_1, p_2, p_3)$ is dropped in what follows. Then [6]

$${}_+\psi \propto D_\nu(\tau)\exp(i\mathbf{p}\cdot\mathbf{x}), \tag{6}$$

where $D_\nu(\tau)$ is the parabolic cylinder function [7] and

$$\nu = \frac{i\lambda}{2} - \frac{1}{2},$$

---

¶This article was submitted by the author in English.

$$\tau = -\sqrt{2eE}\exp\left(-i\frac{\pi}{4}\right)\left(t - \frac{p_3}{eE}\right), \qquad (7)$$

$$\lambda = \frac{m^2 + p_1^2 + p_2^2}{eE}.$$

For the vector boson, we obtain

$$ {}_{+}\psi = \begin{bmatrix} \psi^0 \\ \psi^1 \\ \psi^2 \\ \psi^3 \end{bmatrix} = \begin{bmatrix} c_+ D_{\nu+1}(\tau) + c_- D_{\nu-1}(\tau) \\ c_1 D_\nu(\tau) \\ c_2 D_\nu(\tau) \\ c_+ D_{\nu+1}(\tau) - c_- D_{\nu-1}(\tau) \end{bmatrix} \exp(i\mathbf{p}\cdot\mathbf{x}), \qquad (8)$$

where

$$\psi^1 = {}_{+}\psi^1, \quad \psi^2 = {}_{+}\psi^2,$$

and

$$\psi^0 \equiv {}_{+}\psi^0 = \frac{1}{2}({}_{+}\psi^{+} + {}_{+}\psi^{-}),$$

$$\psi^3 \equiv {}_{+}\psi^3 = \frac{1}{2}({}_{+}\psi^{+} - {}_{+}\psi^{-}), \qquad (9)$$

$${}_{+}\psi^{\pm} = 2c_\pm D_{\nu\pm1}\exp(i\mathbf{p}\cdot\mathbf{x}).$$

The function $D_{\nu\pm1}(\tau)$ is obtained from $D_\nu(\tau)$ in Eqs. (6) and (7) by the substitution

$$m^2 \longrightarrow m^2 \mp 2ieE.$$

Arbitrary coefficients $c_1$, $c_2$, and $c_\pm \equiv {}_{+}c_\pm$ determine the polarization of the vector boson. They are not independent because of constraint (2),

$$c_1 p_1 + c_2 p_2 + \sqrt{2eE}e^{i\pi/4}[(1+\nu){}_{+}c_{+} - {}_{+}c_{-}] = 0. \qquad (10)$$

For the negative-frequency in-solution (for the scalar boson), we have

$${}_{-}\psi \propto [D_\nu(\tau)]^* \exp(i\mathbf{p}\cdot\mathbf{x}) \qquad (11)$$

instead of (6). The star denotes the complex conjugation. Similarly to (8), the parabolic cylinder functions entering ${}_{-}\psi^{\pm}$ are obtained from $[D_\nu(\tau)]^*$ in (11) by the substitution

$$m^2 \longrightarrow m^2 \mp 2ieE,$$

and therefore,

$${}_{-}\psi = \begin{bmatrix} c_+ D_{\nu^*-1}(\tau^*) + c_- D_{\nu^*+1}(\tau^*) \\ c_1 D_{\nu^*}(\tau^*) \\ c_2 D_{\nu^*}(\tau^*) \\ c_+ D_{\nu^*-1}(\tau^*) - c_- D_{\nu^*+1}(\tau^*) \end{bmatrix} \exp(i\mathbf{p}\cdot\mathbf{x}). \qquad (12)$$

(We have $c_\pm = {}_{-}c_\pm$ in (12) and similarly in other cases.) The constraint takes the form

$$c_1 p_1 + c_2 p_2 + \sqrt{2eE}e^{-i\pi/4}[{}_{-}c_{+} + \nu\,{}_{-}c_{-}] = 0. \qquad (13)$$

Nothing prevents us from assuming that $c_1$ and $c_2$ in (12) are the same as in (8).

The negative-frequency out-solution is obtained from the positive-frequency in-solution by changing the sign of τ in the parabolic cylinder functions in (8),

$${}^{-}\psi = \begin{bmatrix} c_+ D_{\nu+1}(-\tau) + c_- D_{\nu-1}(-\tau) \\ c_1 D_\nu(-\tau) \\ c_2 D_\nu(-\tau) \\ c_+ D_{\nu+1}(-\tau) - c_- D_{\nu-1}(-\tau) \end{bmatrix} \exp(i\mathbf{p}\cdot\mathbf{x}), \qquad (14)$$

$${}^{-}c_\pm = -{}_{+}c_\pm,$$

see (112a). The constraint is given by

$$c_1 p_1 + c_2 p_2 + \sqrt{2eE}e^{i\pi/4}[{}^{-}c_{-} - (1+\nu){}^{-}c_{+}] = 0. \qquad (15)$$

Similarly, the positive-frequency out-solution can be found from ${}_{-}\psi$ in (12) by changing the sign of $\tau^*$,

$${}^{+}\psi = \begin{bmatrix} c_+ D_{\nu^*-1}(-\tau^*) + c_- D_{\nu^*+1}(-\tau^*) \\ c_1 D_{\nu^*}(-\tau^*) \\ c_2 D_{\nu^*}(-\tau^*) \\ c_+ D_{\nu^*-1}(-\tau^*) - c_- D_{\nu^*+1}(-\tau^*) \end{bmatrix} \exp(i\mathbf{p}\cdot\mathbf{x}). \qquad (16)$$

The corresponding constraint is

$$c_1 p_1 + c_2 p_2 - \sqrt{2eE}e^{-i\pi/4}[\nu\,{}^{+}c_{-} + {}^{+}c_{+}] = 0. \qquad (17)$$

For the scalar boson, the in- and out-solutions are related by [6]

$$\begin{aligned} {}_{+}\psi_n &= c_{1n}\,{}^{+}\psi_n + c_{2n}\,{}^{-}\psi_n, \\ {}_{-}\psi_n &= c_{2n}^{*}\,{}^{+}\psi_n + c_{1n}^{*}\,{}^{-}\psi_n, \end{aligned} \qquad (18)$$

$$c_{1n} = \frac{\sqrt{2\pi}}{\Gamma((1-i\lambda)/2)}\exp\left[-\frac{\pi}{4}(\lambda-i)\right],$$

$$c_{2n} = \exp\left[-\frac{\pi}{2}(\lambda+i)\right],$$

$$|c_{1n}|^2 - |c_{2n}|^2 = 1.$$

The subscript $n$ indicates a set of quantum numbers; here, $n = \mathbf{p}$. By a straightforward calculation similar to the one in the scalar case, we find that Eqs. (18) also hold for the vector boson and that

$$\begin{gathered} {}^{+}c_{-} = \frac{i}{\nu}{}_{+}c_{-} = -{}_{-}c_{-} = -\frac{i}{\nu}{}^{-}c_{-}, \\ {}^{+}c_{+} = -i(1+\nu){}_{+}c_{+} = -{}_{-}c_{+} = i(1+\nu){}^{-}c_{+}. \end{gathered} \tag{19}$$

These relations guarantee that the wave functions ${}_{\pm}\psi$ and ${}^{\pm}\psi$ are normalized in the same manner and that any constraint can be obtained from any other use of (19).

As seen from (18), the Bogoliubov coefficients $c_{1n}$ and $c_{2n}$ do not depend on the boson polarization in the constant electric field. The imaginary part of the Lagrange function is therefore given by $3\mathrm{Im}\mathrm{L}_{\mathrm{spin}0}$ in agreement with [3, 4].

## 3. VECTOR BOSON IN THE CONSTANT ELECTROMAGNETIC FIELD

We now add a collinear constant magnetic field to the constant electric field. For $A_2 = Hx_1$, we obtain from Eq. (1) that

$$\begin{aligned} (-D^2+m^2)\psi_1 - 2ieH\psi_2 &= 0, \\ (-D^2+m^2)\psi_2 + 2ieH\psi_1 &= 0. \end{aligned} \tag{20}$$

Introducing

$$\begin{gathered} \tilde{\psi}_1 = \psi_1 - i\psi_2, \quad \tilde{\psi}_2 = \psi_1 + i\psi_2, \\ \psi_1 = \frac{1}{2}(\tilde{\psi}_1 + \tilde{\psi}_2), \quad \psi_2 = \frac{i}{2}(\tilde{\psi}_1 - \tilde{\psi}_2), \end{gathered} \tag{21}$$

we rewrite Eqs. (20) as

$$\begin{aligned} (-D^2+m^2+2eH)\tilde{\psi}_1 &= 0, \\ (-D^2+m^2-2eH)\tilde{\psi}_2 &= 0, \end{aligned} \tag{22}$$

and therefore, $\tilde{\psi}_{1,2}$ can be obtained from the scalar boson wave function by the substitutions

$$m^2 \longrightarrow m^2 \pm 2eH.$$

We can therefore write

$$\begin{gathered} \tilde{\psi}_1 \propto 2c_1D_{n-1}(\zeta), \quad \tilde{\psi}_2 \propto 2c_2D_{n+1}(\zeta), \\ \zeta = \sqrt{2eH}\left(x_1 + \frac{p_2}{eH}\right). \end{gathered} \tag{23}$$

Instead of (8), we thus have

$${}_{+}\psi_{p_2p_3n} = \begin{bmatrix} [c_+D_{\nu+1}(\tau) + c_-D_{\nu-1}(\tau)]D_n(\zeta) \\ [c_1D_{n-1}(\zeta) + c_2D_{n+1}(\zeta)]D_\nu(\tau) \\ i[c_1D_{n-1}(\zeta) - c_2D_{n+1}(\zeta)]D_\nu(\tau) \\ [c_+D_{\nu+1}(\tau) - c_-D_{\nu-1}(\tau)]D_n(\zeta) \end{bmatrix} \times \exp(i(p_2x_2 + p_3x_3)), \tag{24}$$

and similarly for the other $\psi$ functions. Here,

$$\nu = \frac{i\lambda}{2} - \frac{1}{2}, \quad \lambda = \frac{m^2 + eH(2n+1)}{eE}. \tag{25}$$

The constraints can be obtained from the previous ones by the substitution

$$c_1p_1 + c_2p_2 \longrightarrow -i\sqrt{2eH}[(1+n)c_2 - c_1]. \tag{26}$$

We note that $D_\mu\psi^\mu$ is proportional to the scalar wave function

$$D_n(\zeta)D_\nu(\tau)\exp[i(p_2x_2 + p_3x_3)]$$

(which is dropped in the expressions similar to (10) with modification (26), or in (116)). Equations (67) and (98) were used to obtain the constraints. It follows from the derivation that the presence of $c_1$ in the right-hand side of (26) is due to the assumption that $D_{n-1}(\zeta)$ is not zero in (24), i.e., $n \geq 1$.

Using (24) and (26), we can build three polarization states $\psi(i, x)$, $i = 1, 2, 3$, see Section 7. For these states, the respective minimum values of $n$ in (25) are –1, 0, 1. Thus the Bogoliubov coefficients depend on all the four quantum numbers ($n = p_2, p_3, n, i$) through the minimum value of $n$.

Because

$$2\mathrm{Im}L = \sum_n \ln(1 + |c_{2n}|^2),$$

it is easy to show that in agreement with [4],

$$\begin{aligned} \mathrm{Im}2L_{\mathrm{spin}1} &= 2 \times 3\mathrm{Im}L_{\mathrm{spin}0} \\ &+ \left\{\ln\left[1 + \exp\left(-\pi\frac{m^2 - eH}{eE}\right)\right]\right. \\ &\left.- \ln\left[1 + \exp\left(-\pi\frac{m^2 + eH}{eE}\right)\right]\right\}\frac{\alpha}{\pi}EHVT. \end{aligned} \tag{27}$$

The factors outside the braces give the statistical weight of the “correcting” states, see Eqs. (3.6) and (3.7) in [6].

The Bogoliubov coefficients allow finding the transition probability from any initial to any final state (with arbitrary occupation numbers) [6]. For example, if the initial state is a vacuum, we have

$$|c_{1n}|^{-2}\{1 + w_n + w_n^2 + w_n^3 + \dots\} = 1,$$
$$w_n = \frac{|c_{2n}|^2}{|c_{1n}|^2}, \tag{28}$$

for the cell with the set of quantum numbers $n = p_2, p_3, n, i$. The term $|c_{1n}|^{-2} w_n^k$ gives the probability for the production of $k$ pairs, $k = 0, 1, 2, \dots$. The events occurring in cells with different quantum numbers are independent.

## 4. THE FREE-VECTOR BOSON PROPAGATOR

The wave functions of a free-vector boson with the momentum $p^\mu = (p^0, 0, 0, p^3)$ can be written as

$$\psi^\mu(i, x) = \frac{u^\mu(i)}{\sqrt{2|p^0|}}\exp(ip\cdot x),$$
$$\eta^{\mu\nu} = \mathrm{diag}(-1, 1, 1, 1), \quad \mu = 0, 1, 2, 3, \tag{29}$$
$$u(1) = \begin{bmatrix}0\\1\\0\\0\end{bmatrix}, \quad u(2) = \begin{bmatrix}0\\0\\1\\0\end{bmatrix}, \quad u(3) = \frac{1}{m}\begin{bmatrix}p_3\\0\\0\\p^0\end{bmatrix}.$$

These solutions satisfy wave equation (1) and constraint (2) with the external field switched off. Summing $\psi^\mu(i, x)\psi^{\nu*}(i, x')$ over polarizations, we find

$$\sum_{i=1}^{3}\psi^\mu(i, x)\psi^{\nu*}(i, x') = \frac{1}{2|p^0|}$$
$$\times\begin{bmatrix}\frac{p_3^2}{m^2} & 0 & 0 & \frac{p_3p^0}{m^2}\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \frac{p_3p^0}{m^2} & 0 & 0 & \frac{(p^0)^2}{m^2}\end{bmatrix}\exp[ip(x-x')]. \tag{30}$$

If we use helicity states instead of linear polarization states (29) (cf. §16 in [8]), we obtain the same result (30). In general, we must replace the matrix in the right-hand side of (30) by $\eta^{\mu\nu} + p^\mu p^\nu/m^2$. This case differs from the scalar particle case only by the presence of this matrix in the expression for the propagator (which is similar to (51)). The vector boson propagator can therefore be obtained from the scalar one

$$G_{\mathrm{spin}0}(x, x') = \frac{1}{(2\pi)^4}\int d^4p\frac{\exp[ip(x-x')]}{p^2 + m^2 - i\varepsilon}$$
$$= \frac{1}{(4\pi)^2}\int_0^\infty\frac{ds}{s^2}\exp\left[-ism^2 + \frac{i(x-x')^2}{4s}\right], \tag{31}$$

considered as a unit matrix over discrete indices, by acting on it with the differential operator

$$G^{\mu\nu}(x, x')$$
$$= \left(\eta^{\mu\nu} - \frac{1}{m^2}\frac{\partial^2}{\partial x_\mu\partial x_\nu}\right)G_{\mathrm{spin}0}(x, x'). \tag{32}$$

Because the scalar boson propagator satisfies the equation

$$(-\partial_\mu\partial^\mu + m^2)G_{\mathrm{spin}0}(x, x') = \delta^4(x-x'), \tag{33}$$

we have

$$(-\partial_\sigma\partial^\sigma + m^2)G^{\mu\nu}(x, x')$$
$$= \left(\eta^{\mu\nu} - \frac{1}{m^2}\frac{\partial^2}{\partial x_\mu\partial x_\nu}\right)\delta^4(x-x') \tag{34}$$

for the vector boson. We note that the right-hand side is not simply given by $\delta^4(x-x')$. The complication is due to the existence of constraints. This prevents us from using the well-known methods of constructing propagators of scalar and spinor particles in an external field [9, 10]. An elegant way to circumvent this difficulty was given by Vanyashin and Terentyev [3].

## 5. THE VECTOR BOSON PROPAGATOR IN THE CONSTANT MAGNETIC FIELD

To write the propagator, we need the complete set of orthonormalized solutions. The orthonormalization is performed by expressing the vector current as [5]

$$J_\mu = -i\{\psi^{\nu*}(D_\mu\psi_\nu - D_\nu\psi_\mu)$$
$$-(D_\mu^*\psi_\nu^* - D_\nu^*\psi_\mu^*)\psi^\nu\}, \tag{35}$$
$$D_\mu = \partial_\mu + ieA_\mu.$$

We note that our expression for $D_\mu$ in (35) coincides with that in [5]; although our $\eta_{\mu\nu}$ has a different sign, we also replace $e \longrightarrow -e$, using the fact that $e = |e|$ and assuming that $W^-$ is a particle by analogy with the electron.

In the space without a field, the expression for $J_\mu$ can be written similarly to the scalar case up to divergence terms (see §15 in [8]). It is remarkable that with constraint (2), the same is true in the presence of a field. Indeed,

$$-\psi^{\nu*}D_\nu\psi_\mu = -\partial_\nu(\psi^{\nu*}\psi_\mu) + \psi_\mu D_\nu^*\psi^{\nu*}. \qquad (36)$$

The last term in the right-hand side vanishes because of Eq. (2) for the boson with $g = 2$. Similarly,

$$\begin{aligned}(D_\nu^*\psi_\mu^*)\psi^\nu &= \partial_\nu(\psi^\nu\psi_\mu^*) - \psi_\mu^* D_\nu\psi^\nu \\ &= \partial_\nu(\psi^\nu\psi_\mu^*),\end{aligned} \qquad (37)$$

and therefore,

$$\begin{aligned}J_\mu &= -i\{\psi^{\nu*}D_\mu\psi_\nu \\ &- (D_\nu^*\psi_\mu^*)\psi^\nu - \partial_\nu[\psi^{\nu*}\psi_\mu - \psi^\nu\psi_\mu^*]\}.\end{aligned} \qquad (38)$$

To normalize the wave functions, we need only $J_0$. Straightforward calculations show that the divergence terms do not contribute to $J_0$ for the fields considered here. We then have

$$J^0 = -J_0 = i\{\psi^{\nu*}D_0\psi_\nu - (D_0^*\psi_\nu^*)\psi^\nu\}. \qquad (39)$$

For orthonormalization, we use the expression

$$J^0(\psi', \psi) = i\{\psi'^{\nu*}D_0\psi_\nu - (D_0^*\psi_\nu'^*)\psi^\nu\}. \qquad (40)$$

Our vector potentials are such that $A_0(x) = 0$. It then follows that $D_0 = \partial/\partial t$ and

$$J^0(\psi', \psi) = i\{\psi_k'^*\overleftrightarrow{\partial}_t\psi_k - \psi'^{0*}\overleftrightarrow{\partial}_t\psi^0\}, \qquad (41)$$

where the sum over $k$ runs from 1 to 3.

The positive-frequency solution of wave equation (1) with $A_\mu(x) = \delta_{\mu 2}Hx_1$ is given by

$$\psi^\mu_{p_2,\,p_3,\,n} = \begin{bmatrix} c^0 D_n(\zeta) \\ c_1 D_{n-1}(\zeta) + c_2 D_{n+1}(\zeta) \\ i[c_1 D_{n-1}(\zeta) - c_2 D_{n+1}(\zeta)] \\ c_3 D_n(\zeta) \end{bmatrix} \qquad (42)$$

$$\times \exp[i(p_2x_2 + p_3x_3 - p^0t)].$$

The elements of this column correspond to $\mu = 0, 1, 2, 3$,

$$\zeta = \sqrt{2eH}\left(x_1 + \frac{p_2}{eH}\right),$$

$$p^0 = \sqrt{m^2 + p_3^2 + eH(2n+1)}.$$

The coefficients $c$ determining the boson polarization satisfy the constraint

$$-ip^0c^0 + ip_3c_3 + \sqrt{2eH}[(n+1)c_2 - c_1] = 0. \qquad (43)$$

For states with the polarizations $c'$ and $c$, Eqs. (41) and (42) imply that

$$\begin{aligned}J^0(\psi', \psi) &= 2p^0\{2c_1'^*c_1D_{n-1}^2(\zeta) \\ &+ 2c_2'^*c_2D_{n+1}^2(\zeta) + (c_3'^*c_3 - c'^{0*}c^0)D_n^2(\zeta)\}.\end{aligned} \qquad (44)$$

Integrating over $x_1$, we find

$$\int_{-\infty}^{\infty} dx_1 J^0(\psi', \psi) = 2p^0 n!\sqrt{\frac{\pi}{eH}}$$

$$\times\left\{\frac{2}{n}c_1'^*c_1 + 2(n+1)c_2'^*c_2 + c_3'^*c_3 - c'^{0*}c^0\right\}, \qquad (45)$$

$$\int_{-\infty}^{\infty} dx_1 D_n^2(\zeta) = n!\sqrt{\frac{\pi}{eH}}.$$

Using the orthonormalization conditions

$$\begin{gathered}\int dx_1 J^0({}_\pm\psi(i, x), {}_\pm\psi(j, x)) = \pm\delta_{ij}, \\ i, j = 1, 2, 3,\end{gathered} \qquad (46)$$

and constraint (43), we find the positive-frequency polarization states

$$\psi^\mu(1, x) = N(1)\begin{bmatrix} (n+1)\sqrt{2eH}p^0D_n(\zeta) \\ im_\perp^2 D_{n+1}(\zeta) \\ m_\perp^2 D_{n+1}(\zeta) \\ (n+1)\sqrt{2eH}p_3D_n(\zeta) \end{bmatrix} \qquad (47)$$

$$\times \exp[i(p_2x_2 + p_3x_3 - p^0t)],$$

where

$$\mu = 0, 1, 2, 3, \quad m_\perp^2 = m^2 + eH(2n+1),$$

$$p^0 = \sqrt{m^2 + p_3^2 + eH(2n+1)},$$

$$N(1) = n_1N_0, \quad N_0 = \left(\frac{eH}{\pi}\right)^{1/4}\frac{1}{\sqrt{2|p^0|n!}}, \qquad (48)$$

$$n_1 = \frac{1}{\sqrt{2(n+1)m_\perp^2(m^2 + eHn)}},$$

$$\psi(2,x) = n_2 N_0 \begin{bmatrix} p_3 D_n(\zeta) \\ 0 \\ 0 \\ p^0 D_n(\zeta) \end{bmatrix} \exp[i(p_2x_2 + p_3x_3 - p^0 t)], \tag{49}$$

$$n_2 = \frac{1}{\sqrt{m_\perp^2}},$$

$$\psi(3,x) = n_3 N_0$$

$$\times \begin{bmatrix} \sqrt{2eH}\,p^0 D_n(\zeta) \\ i[-(m^2 + eHn)D_{n-1}(\zeta) + eHD_{n+1}(\zeta)] \\ (m^2 + eHn)D_{n-1}(\zeta) + eHD_{n+1}(\zeta) \\ \sqrt{2eH}\,p_3 D_n(\zeta) \end{bmatrix} \tag{50}$$

$$\times \exp[i(p_2x_2 + p_3x_3 - p^0 t)],$$

$$n_3 = \sqrt{\frac{n}{2m^2(m^2 + eHn)}}.$$

We separatare the scalar wave function normalization factor $N_0$ from the normalization factors $N(i)$ because we concentrate our attention on the differences from the scalar case. We also note that

$$N(3) \propto \Gamma^{-1/2}(n),$$

which vanishes for $n = 0$. For the state $\psi(3, x)$ only the values $n = 1, 2, 3, \ldots$ are therefore possible. The same follows from the fact that constraint (43) cannot be satisfied because it does not involve $c_1$ for $n = 0$.

We now construct the vector boson propagator. We start from the expression (which is a special case of a more general result derived in Section 6, see Eqs. (80) and (81))

$$G^{\mu\nu}(x,x') = i\int_{-\infty}^{\infty}\frac{dp_2}{2\pi}\int_{-\infty}^{\infty}\frac{dp_3}{2\pi}$$

$$\times \sum_{n=-1}^{\infty}\sum_{i=1}^{3}\begin{cases} {}_+\psi^\mu(i,x)_+\psi^{*\nu}(i,x'), & t > t' \\ {}_-\psi^\mu(i,x)_-\psi^{*\nu}(i,x'), & t < t'. \end{cases} \tag{51}$$

In what follows, we let $E_n$ denote the previous quantity $p^0$ and use the relations

$$-\frac{1}{2\pi i}\int_{-\infty}^{\infty} dp^0 \frac{\exp[-ip^0(t-t')]}{(p^0 - E_n + i\epsilon)(p^0 + E_n - i\epsilon)}$$

$$= \frac{1}{2E_n}\begin{cases} \exp[-iE_n(t-t')], & t > t' \\ \exp[iE_n(t-t')], & t < t', \end{cases} \tag{52}$$

$$\frac{1}{i(E_n^2 - (p^0)^2)} = \int_0^\infty ds \exp[-is(E_n^2 - p_0^2)]$$

to rewrite (51) as (with $p^0 = -p_0$)

$$G^{\mu\nu}(x,x') = i\sqrt{\frac{eH}{\pi}}\sum_{n=-1}^{\infty}\int_{-\infty}^{\infty}\frac{dp_2}{2\pi}\int_{-\infty}^{\infty}\frac{dp_3}{2\pi}\int_{-\infty}^{\infty}\frac{dp^0}{2\pi}$$

$$\times\int_0^\infty ds\, a^{\mu\nu}(x,x')\frac{1}{n!}\exp[-is(m_\perp^2 + p_3^2 - p_0^2) \tag{53}$$

$$+ i[p_2(x_2 - x_2') + p_3(x_3 - x_3') - p^0(t-t')]],$$

$$m_\perp^2 = m^2 + eH(2n+1).$$

We note that the lower line in the right-hand side of (52) is obtained from the upper line by the substitution $t \longleftrightarrow t'$, which does not change anything, because the right-hand side can be written as

$$(2E_n)^{-1}\exp[-iE_n|t-t'|].$$

The form of the left-hand side that is explicitly symmetric in $t$ and $t'$ is

$$\int_0^\infty ds \exp[-isE_n^2]\int_{-\infty}^{\infty}\frac{dp^0}{2\pi}\exp[isp_0^2 - ip^0(t-t')]$$

$$= \frac{e^{i\pi/4}}{2\sqrt{\pi}}\int_0^\infty\frac{ds}{\sqrt{s}}\exp\left[-isE_n^2 - i\frac{(t-t')^2}{4s}\right]. \tag{54}$$

We first obtain the scalar particle propagator in the proper-time representation [10]. We replace $a^{\mu\nu}(x, x')$ by $D_n(\zeta)D_n(\zeta')$ in (53). Using the formula

$$D_n(\zeta) = \sqrt{\frac{2}{\pi}}e^{\zeta^2/4}\int_0^\infty dy\, y^n e^{-y^2/2}\cos\left(\zeta y - \frac{n\pi}{2}\right), \tag{55}$$

we then find

$$\sum_{n=0}^{\infty}\frac{D_n(\zeta)D_n(\zeta')}{n!}\exp[-i\tau(2n+1)] = (2\sin 2\tau)^{-1/2}$$

$$\times\exp\left[-i\frac{\pi}{4} + i\frac{(\zeta-\zeta')^2}{8\tan\tau} - i\frac{(\zeta+\zeta')^2}{8\cot\tau}\right], \tag{56}$$

$$\tau = eHs, \quad \zeta' = \sqrt{2eH}\left(x_1' + \frac{p_2}{eH}\right).$$

Subsequent integration over $p_2$ gives

$$\int_{-\infty}^{\infty}\frac{dp_2}{2\pi}\sum_{n=0}^{\infty}\frac{D_n(\zeta)D_n(\zeta')}{n!}\exp[-i\tau(2n+1) + ip_2z_2]$$

$$
\begin{gathered}
= -i\sqrt{\frac{eH}{\pi}}(4\sin\tau)^{-1} \\
\times \exp\left[-i\frac{eHz_2(x_1+x_1')}{2} + i\frac{eH(z_1^2+z_2^2)}{4\tan\tau}\right], \\
z_\mu = x_\mu - x_\mu'.
\end{gathered}
\tag{57}
$$

Using

$$
\begin{gathered}
\int_{-\infty}^{\infty}\frac{dp_3}{2\pi}\int_{-\infty}^{\infty}\frac{dp^0}{2\pi} \\
\times \exp[i(p_3z_3 - p^0z^0) - is(p_3^2 - p_0^2)] \\
= \frac{1}{4\pi s}\exp\left[i\frac{z_3^2 - z_0^2}{4s}\right],
\end{gathered}
\tag{58}
$$

we find [6, 9–11]

$$
\begin{gathered}
G_{\text{spin0}}(x, x') = \frac{eH}{(4\pi)^2}\int_0^{\infty}\frac{ds}{s\sin(eHs)} \\
\times \exp\left[-i\frac{eHz_2(x_1+x_1')}{2}\right] \\
\times \exp\left[-ism^2 + i\frac{z_3^2 - z_0^2}{4s} + i\frac{(z_1^2+z_2^2)eH}{4\tan(eHs)}\right].
\end{gathered}
\tag{59}
$$

We now show how to obtain $a^{\mu\nu}(x, x')$ in (53) and how to turn it into a differential matrix that gives the vector boson propagator when inserted in the integrand in (59). As a preliminary step, we write two formulas directly related to (56):

$$
\begin{gathered}
\sum_{n=-1}^{\infty}\frac{D_{n+1}(\zeta)D_{n+1}(\zeta')}{(n+1)!}\exp[-i\tau(2n+1)] \\
= \exp(2i\tau)\sum_{n=0}^{\infty}\frac{D_n(\zeta)D_n(\zeta')}{n!}\exp[-i\tau(2n+1)],
\end{gathered}
\tag{60}
$$

$$
\begin{gathered}
\sum_{n=1}^{\infty}\frac{D_{n-1}(\zeta)D_{n-1}(\zeta')}{(n-1)!}\exp[-i\tau(2n+1)] \\
= \exp(-2i\tau)\sum_{n=0}^{\infty}\frac{D_n(\zeta)D_n(\zeta')}{n!}\exp[-i\tau(2n+1)].
\end{gathered}
\tag{61}
$$

We see that the expressions in (60) and (61) differ from the scalar case only by the factors $e^{2i\tau}$ and $e^{-2i\tau}$.

We now return to $a^{\mu\nu}(x, x')$. As seen from (51) and (53),

$$
a^{\mu\nu}(x, x') \propto \sum_{i=1}^{3}\psi^\mu(i, x)\psi^{\nu *}(i, x').
\tag{62}
$$

Taking, e.g., $a^{11}(x, x')$, we see from (49) that $\psi^1(2, x) = 0$, i.e., the term with $i = 2$ does not contribute to $a^{11}(x, x')$. In accordance with (47), the contribution of the term with $i = 1$ is

$$
\begin{gathered}
n_1^2 m_\perp^4 D_{n+1}(\zeta)D_{n+1}(\zeta'), \\
n_1^2 = \frac{1}{2(n+1)m_\perp^2(m^2+eHn)}.
\end{gathered}
\tag{63}
$$

The term with $i = 3$ gives

$$
\begin{gathered}
n_3^2[-(m^2+eHn)D_{n-1}(\zeta) + eHD_{n+1}(\zeta)] \\
\times[-(m^2+eHn)D_{n-1}(\zeta') + eHD_{n+1}(\zeta')], \\
n_3^2 = \frac{n}{2m^2(m^2+eHn)}.
\end{gathered}
\tag{64}
$$

We now have $a^{11}(x, x')$ as the sum of (63) and (64):

$$
\begin{gathered}
a^{11}(x, x') = \frac{1}{2(m^2+eHn)} \\
\times\left(\frac{m_\perp^2}{n+1} + \frac{(eH)^2 n}{m^2}\right)D_{n+1}(\zeta)D_{n+1}(\zeta') \\
+ \frac{n(m^2+eHn)}{2m^2}D_{n-1}(\zeta)D_{n-1}(\zeta') \\
- \frac{eHn}{2m^2}[D_{n+1}(\zeta)D_{n-1}(\zeta') + D_{n-1}(\zeta)D_{n+1}(\zeta')].
\end{gathered}
\tag{65}
$$

Next, we note that

$$
\frac{1}{m^2+eHn}\left(\frac{m_\perp^2}{n+1} + \frac{(eH)^2 n}{m^2}\right) = \frac{1}{n+1} + \frac{eH}{m^2},
\tag{66}
$$

i.e., the undesirable factor $1/(m^2 + eHn)$ contained in $n_1^2$ and $n_3^2$ in (63) and (64) disappears in the sum in Eq. (65).

In what follows, we use the relations

$$
\begin{gathered}
\left(\frac{d}{d\zeta} + \frac{\zeta}{2}\right)D_n(\zeta) = nD_{n-1}(\zeta), \\
\left(\frac{d}{d\zeta} - \frac{\zeta}{2}\right)D_n(\zeta) = -D_{n+1}(\zeta),
\end{gathered}
\tag{67}
$$

see Eqs. (8.2.15–16) in [7]. We also write the sum and the difference of these expressions:

$$
\begin{gathered}
2\frac{d}{d\zeta}D_n(\zeta) = nD_{n-1}(\zeta) - D_{n+1}(\zeta), \\
\zeta D_n(\zeta) = nD_{n-1}(\zeta) + D_{n+1}(\zeta).
\end{gathered}
\tag{68}
$$

It is then easy to verify that

$$a^{11}(x, x') = \frac{D_{n+1}(\zeta)D_{n+1}(\zeta')}{2(n+1)} + \frac{n}{2}D_{n-1}(\zeta)D_{n-1}(\zeta') + \frac{2eH}{m^2}\frac{\partial^2}{\partial\zeta\partial\zeta'}D_n(\zeta)D_n(\zeta'). \qquad (69)$$

The first term in the right-hand side of (69) is involved in Eq. (60) and the second term is used in (61); the necessary factor $n!$ comes from $N_0$ see Eq. (48). The third term can be written as

$$\frac{1}{m^2}\frac{\partial^2}{\partial x_1 \partial x'_1}D_n(\zeta)D_n(\zeta'). \qquad (70)$$

In a similar manner, we find the other components

$$a^{\mu\nu}(x, x') = a^{\nu\mu *}(x', x).$$

It is easy to verify that the differential operator $A^{\mu\nu}(x, x')$ corresponding to $a^{\mu\nu}(x, x')$ is given by

$$A^{\mu\nu} = B^{\mu\nu} + C^{\mu\nu}, \quad C^{\mu\nu} = \frac{1}{m^2}\Pi^{\mu}(x)\Pi^{\nu *}(x'),$$
$$\Pi_{\mu}(x) = -i\frac{\partial}{\partial x^{\mu}} + eA_{\mu}(x), \qquad (71)$$
$$\Pi^*_{\mu}(x') = i\frac{\partial}{\partial x'^{\mu}} + eA_{\mu}(x').$$

In our case,

$$A_{\mu}(x) = \delta_{\mu 2}Hx_1, \quad \Pi^0(x) = i\frac{\partial}{\partial t},$$
$$\Pi^{0*}(x') = -i\frac{\partial}{\partial t'}. \qquad (72)$$

The nonzero components $B^{\mu\nu}$ are

$$B^{11} = B^{22} = \cos\tau, \quad B^{21} = -B^{12} = \sin\tau,$$
$$B^{33} = -B^{00} = 1. \qquad (73)$$

The difference of $B^{\mu\nu}$ from $\eta^{\mu\nu}$ is due to the interaction of the boson magnetic moment with the magnetic field. We can say that $B^{\mu\nu}$ with $\mu, \nu = 1, 2$ are "dressed" by the magnetic field.

Thus,

$$G^{\mu\nu}(x, x') = \frac{eH}{(4\pi)^2}\int_0^{\infty}\frac{ds}{s\sin(eHs)}\exp(-ism^2)A^{\mu\nu}$$
$$\times \exp\left[-\frac{ieHz_2(x_1 + x'_1)}{2}\right] \qquad (74)$$
$$\times \exp\left[\frac{i(z_3^2 - z_0^2)}{4s} + \frac{i}{4}(z_1^2 + z_2^2)eH\cot(eHs)\right],$$
$$z_{\mu} = x_{\mu} - x'_{\mu}.$$

It is somewhat surprizing that this representation does not coincide with the Vanyashin–Terentyev representation [3] with the electric field switched off. Possibly, these are two different representations for the same propagator, and it would be interesting to verify this hypothesis.

## 6. THE VECTOR BOSON PROPAGATOR IN THE CONSTANT ELECTRIC FIELD

We first give the generalization of (51) for the case where the external field can create pairs [12, 6]. For this purpose, we write

$$G(x, x')_{\rm abs} = i\langle 0_{\rm out}|T(\Psi(x)\Psi^{\dagger}(x'))|0_{\rm in}\rangle$$
$$= \langle 0_{\rm out}|0_{\rm in}\rangle G(x, x'), \qquad (75)$$

where $T$ is the chronological ordering operator and

$$\Psi(x) = \sum_n [a_{n\,\rm out}\,{}^{+}\psi_n(x) + b^{+}_{n\,\rm out}\,{}^{-}\psi_n(x)],$$
$$\Psi^{\dagger}(x) = \sum_n [a^{\dagger}_{n\,\rm in}\,{}_{+}\psi^*_n(x) + b_{n\,\rm in}\,{}_{-}\psi^*_n(x)]. \qquad (76)$$

As usual, $a_n$ and $b_n$ are the particle and antiparticle destruction operators in a state with the quantum numbers $n$:

$$\Psi^{\dagger}(x')|0_{\rm in}\rangle = \sum_k {}_{+}\psi^*_k(x')a^{\dagger}_{k\,\rm in}|0_{\rm in}\rangle,$$
$$\langle 0_{\rm out}|\Psi(x) = \langle 0_{\rm out}|\sum_n a_{n\,\rm out}\,{}^{+}\psi_n(x). \qquad (77)$$

For $t > t'$, it follows from (75) and (77) that

$$G(x, x')_{\rm abs} = i\sum_{n,k} {}^{+}\psi_n(x)\,{}_{+}\psi^*_k(x')\langle 0_{\rm out}|a_{n\rm out}a^{\dagger}_{k\,\rm in}|0_{\rm in}\rangle,$$
$$t > t'. \qquad (78)$$

In our case, the Bogoliubov transformations have the form (cf. Eq (18)) [6]

$$a^{\dagger}_{n\,\rm out} = c^*_{1n}a^{\dagger}_{n\,\rm in} + c_{2n}b_{n\,\rm in},$$
$$b_{n\,\rm out} = c^*_{2n}a^{\dagger}_{n\,\rm in} + c_{1n}b_{n\,\rm in}. \qquad (79)$$

The first equation in (79) implies that

$$a^{\dagger}_{k\,\rm out}|0_{\rm in}\rangle = c^*_{1k}a^{\dagger}_{k\,\rm in}|0_{\rm in}\rangle.$$

Inserting $a^{\dagger}_{k\,\rm in}|0_{\rm in}\rangle$ from this relationship into (78) and using the commutation relation

$$[a_{k\,\rm out}, a^{\dagger}_{n\,\rm out}] = \delta_{kn},$$

we then obtain

$$G(x, x')_{\rm abs} = \langle 0_{\rm out}|0_{\rm in}\rangle i\sum_n {}^{+}\psi_n(x)\,{}_{+}\psi_n^*(x')\frac{1}{c_{1n}^*}, \quad t > t'. \tag{80}$$

Similarly, for $t < t'$, we find

$$G(x, x')_{\rm abs} = \langle 0_{\rm out}|0_{\rm in}\rangle i\sum_n {}_{-}\psi_n(x)\,{}^{-}\psi_n^*(x')\frac{1}{c_{1n}^*}, \quad t < t'. \tag{81}$$

If the external field does not create pairs, the expressions obtained become those in (51).

In terms of the states ${}_{+}\psi'$ and ${}_{+}\psi$ in (8), the transition current (41) becomes

$$\begin{aligned} J^0({}_{+}\psi', {}_{+}\psi) &= \sqrt{2eE}e^{\pi\lambda/4} \\ &\times[c_1'^*c_1 + c_2'^*c_2 + 2i({}_{+}c_-'^*\,{}_{+}c_+ - {}_{+}c_+'^*\,{}_{+}c_-)], \end{aligned} \tag{82}$$

where we used Eq. (8.2.11) in [7] (and its complex conjugate):

$$D_{\nu^*+1}(\tau^*)\frac{\overset{\leftrightarrow}{d}}{dt}D_{\nu-1}(\tau) = \sqrt{2eE}\exp\left(\frac{\pi\lambda}{4}\right) \tag{83}$$

$$= -D_{\nu^*-1}(\tau^*)\frac{\overset{\leftrightarrow}{d}}{dt}D_{\nu+1}(\tau) = D_{\nu^*}(\tau^*)i\frac{\overset{\leftrightarrow}{d}}{dt}D_\nu(\tau). \tag{84}$$

The constraint is given in (10). Using Eqs. (82), (8), and (10), we find the ${}_{+}\psi$ polarization states

$${}_{+}\psi(1, x) = N(1)\begin{bmatrix} p_2\sqrt{\frac{eE}{2}}e^{i\pi/4}[D_{\nu+1}(\tau) - \nu D_{\nu-1}(\tau)] \\ 0 \\ m_\perp^2 D_\nu(\tau) \\ p_2\sqrt{\frac{eE}{2}}e^{i\pi/4}[D_{\nu+1}(\tau) + \nu D_{\nu-1}(\tau)] \end{bmatrix} \times\exp[i\mathbf{p}\cdot\mathbf{x}],$$

$$N(i) = n_iN_0, \quad N(1) = n_1N_0, \tag{85}$$

$$n_1 = \sqrt{\frac{1}{m_\perp^2(m^2 + p_1^2)}}, \quad N_0 = (2eE)^{-1/4}\exp\left(-\frac{\pi\lambda}{8}\right),$$

$${}_{+}\psi(2, x) = N(2)\begin{bmatrix} D_{\nu+1}(\tau) + (1+\nu)D_{\nu-1}(\tau) \\ 0 \\ 0 \\ D_{\nu+1}(\tau) - (1+\nu)D_{\nu-1}(\tau) \end{bmatrix} \times\exp(i\mathbf{p}\cdot\mathbf{x}), \tag{86}$$

$$n_2 = \sqrt{\frac{eE}{2m_\perp^2}}, \quad m_\perp^2 = m^2 + p_1^2 + p_2^2,$$

$${}_{+}\psi(3, x) = N(3)\begin{bmatrix} p_1\sqrt{\frac{eE}{2}}e^{i\pi/4}[D_{\nu+1}(\tau) - \nu D_{\nu-1}(\tau)] \\ (m^2 + p_1^2)D_\nu(\tau) \\ p_1p_2D_\nu(\tau) \\ p_1\sqrt{\frac{eE}{2}}e^{i\pi/4}[D_{\nu+1}(\tau) + \nu D_{\nu-1}(\tau)] \end{bmatrix} \times\exp(i\mathbf{p}\cdot\mathbf{x}), \tag{87}$$

$$n_3 = \frac{1}{\sqrt{m^2(m^2 + p_1^2)}}.$$

The ${}^{+}\psi$ polarization states can be obtained from these ones using Eqs. (19) (see also (16)):

$${}^{+}\psi(1, x) = N(1) \times\begin{bmatrix} p_2\sqrt{\frac{eE}{2}}e^{-i\pi/4}[(1+\nu)D_{\nu^*-1}(-\tau^*) + D_{\nu^*+1}(-\tau^*)] \\ 0 \\ m_\perp^2 D_{\nu^*}(-\tau^*) \\ p_2\sqrt{\frac{eE}{2}}e^{-i\pi/4}[(1+\nu)D_{\nu^*-1}(-\tau^*) - D_{\nu^*+1}(-\tau^*)] \end{bmatrix} \times\exp(i\mathbf{p}\cdot\mathbf{x}), \tag{88}$$

$${}^{+}\psi(2, x) = N(2) \times\begin{bmatrix} i(1+\nu)\left[-D_{\nu^*-1}(-\tau^*) + \frac{1}{\nu}D_{\nu^*+1}(-\tau^*)\right] \\ 0 \\ 0 \\ i(1+\nu)\left[-D_{\nu^*-1}(-\tau^*) - \frac{1}{\nu}D_{\nu^*+1}(-\tau^*)\right] \end{bmatrix} \times\exp(i\mathbf{p}\cdot\mathbf{x}), \tag{89}$$

$${}^{+}\psi(3, x) = N(3) \times\begin{bmatrix} p_1\sqrt{\frac{eE}{2}}e^{-i\pi/4}[(1+\nu)D_{\nu^*-1}(-\tau^*) + D_{\nu^*+1}(-\tau^*)] \\ (m^2 + p_1^2)D_{\nu^*}(-\tau^*) \\ p_1p_2D_{\nu^*}(-\tau^*) \\ p_1\sqrt{\frac{eE}{2}}e^{-i\pi/4}[(1+\nu)D_{\nu^*-1}(-\tau^*) - D_{\nu^*+1}(-\tau^*)] \end{bmatrix} \times\exp(i\mathbf{p}\cdot\mathbf{x}). \tag{90}$$

In Eqs. (85)–(90), the states $\psi(i, x)$ are characterized by $p_1$, $p_2$, $p_3$, and $i$; $\nu$ and $\lambda$ are given in (7).

We note that the transition current $J^0({}^+\psi', {}^+\psi)$ expressed in terms of ${}^+c$ has the same form as $J^0({}_+\psi', {}_+\psi)$ expressed in terms of ${}_+c$, see Eq. (82). A similar statement is true for the negative-frequency states. Because $\nu + 1 = -\nu^*$ in accordance with Eq. (7), it follows from (19) that

$$\begin{aligned} {}^+c'^*_- \, {}^+c_+ &= {}_+c'^*_- \, {}_+c_+ \\ &= {}^-c'^*_- \, {}^-c_+ = {}_-c'^*_- \, {}_-c_+. \end{aligned} \tag{90a}$$

Therefore,

$$\begin{aligned} &J^0({}_+\psi(i, x), {}_+\psi(j, x)) \\ &= J^0({}^+\psi(i, x), {}^+\psi(j, x)) \propto \delta_{i, j}, \end{aligned} \tag{91}$$

and

$$\begin{aligned} &J^0({}_-\psi(i, x), {}_-\psi(j, x)) \\ &= J^0({}^-\psi(i, x), {}^-\psi(j, x)) = -J^0({}_+\psi(i, x), {}_+\psi(j, x)). \end{aligned} \tag{91a}$$

As previously, we focus our attention on the differences from the scalar case in expressions similar to (53). The proper time representation of the scalar particle propagator is given by [12]

$$\begin{aligned} G(x, x')_{\text{spin0}} &= \frac{eE}{(4\pi)^2}\exp\left[\frac{i}{2}eE(t + t')z_3\right] \\ &\times \int_0^\infty \frac{ds}{s\sinh(eEs)}\exp\left[-ism^2 + \frac{i}{4s}(z_1^2 + z_2^2)\right. \\ &\left.+ \frac{i}{4}eE(z_3^2 - z_0^2)\coth(eEs)\right]. \end{aligned} \tag{92}$$

This can be derived similarly to the magnetic case, but with the role of Eq. (52) played by the relation [12, 6]

$$\begin{aligned} &\sqrt{2}\int_0^\infty \frac{d\theta}{\sqrt{\sinh 2\theta}} \\ &\times \exp\left\{-i2\kappa\theta - \frac{i}{8}\left[\frac{(T + T')^2}{\coth\theta} + \frac{(T - T')^2}{\tanh\theta}\right]\right\} \\ &= \Gamma\left(i\kappa + \frac{1}{2}\right) \\ &\times \begin{cases} D_{-i\kappa - (1/2)}(\chi)D_{-i\kappa - (1/2)}(-\chi'), & T > T' \\ D_{-i\kappa - (1/2)}(-\chi)D_{-i\kappa - (1/2)}(\chi'), & T < T', \end{cases} \end{aligned} \tag{93}$$

where

$$\begin{aligned} &\theta = eEs, \quad T = \sqrt{2eE}\left(t - \frac{p_3}{eE}\right), \\ &T' = \sqrt{2eE}\left(t' - \frac{p_3}{eE}\right), \quad \chi = -\tau^* = e^{i\pi/4}T, \\ &\chi' = e^{i\pi/4}T', \quad \kappa = \frac{\lambda}{2} = \frac{m_\perp^2}{2eE}. \end{aligned} \tag{94}$$

The lower line in the right-hand side of (93) can be obtained from the upper line by the substitution $T \longleftrightarrow T'$. As seen from the left-hand side of (93), this does not change the value of (93), cf. the remark after Eq. (53).

By analogy with the magnetic case, we expect the appearance of the factors $e^{\pm 2\theta}$ in the integrand of (93), cf. Eqs. (73) and (60), (61). To make the insertion possible, we must rotate the integration contour clockwise by a certain angle. This is in line with the Vanyashin–Terentyev approach [3]. After the substitution $\kappa \longrightarrow \kappa + i$, it then follows from (93) that

$$\begin{aligned} &\sqrt{2}\int_C \frac{d\theta}{\sqrt{\sinh 2\theta}} \\ &\times \exp\left\{-i2\kappa\theta + 2\theta - \frac{i}{8}\left[\frac{(T + T')^2}{\coth\theta} + \frac{(T - T')^2}{\tanh\theta}\right]\right\} \\ &= \Gamma\left(i\kappa - \frac{1}{2}\right)\begin{cases} D_{-i\kappa + (1/2)}(\chi)D_{-i\kappa + (1/2)}(-\chi'), & T > T' \\ D_{-i\kappa + (1/2)}(-\chi)D_{-i\kappa + (1/2)}(\chi'), & T < T'. \end{cases} \end{aligned} \tag{95}$$

Similarly, substituting $\kappa \longrightarrow \kappa - i$ in (93), we obtain

$$\begin{aligned} &\sqrt{2}\int_0^\infty \frac{d\theta}{\sqrt{\sinh 2\theta}} \\ &\times \exp\left\{-i2\kappa\theta - 2\theta - \frac{i}{8}\left[\frac{(T + T')^2}{\coth\theta} + \frac{(T - T')^2}{\tanh\theta}\right]\right\} \\ &= \Gamma\left(i\kappa + \frac{3}{2}\right)\begin{cases} D_{-i\kappa - (3/2)}(\chi)D_{-i\kappa - (3/2)}(-\chi'), & T > T' \\ D_{-i\kappa - (3/2)}(-\chi)D_{-i\kappa - (3/2)}(\chi'), & T < T'. \end{cases} \end{aligned} \tag{96}$$

The integration over $p_3$ contained in the sum over $n$ in Eqs. (80) and (81) gives

$$\begin{aligned} &\int_{-\infty}^{\infty} \frac{dp_3}{2\pi}\exp\left[ip_3z_3 - \frac{i}{8}(T + T')\tanh\theta\right] \\ &= \frac{1}{2}e^{-i\pi/4}\sqrt{\frac{eE\coth\theta}{\pi}}\exp\left\{\frac{iz_3^2eE}{4\tanh\theta} + \frac{ieEz_3(t + t')}{2}\right\}, \end{aligned} \tag{97}$$

$$z_3 = x_3 - x_3',$$

where $T$ and $T'$ are functions of $p_3$, see (94). Further calculations leading to (92) are similar to those in the magnetic case.

We now consider the differences from the scalar case. We first rewrite relations (67) and (68) between the parabolic cylinder functions for the present case as

$$\begin{aligned}\left(\frac{d}{d\tau'^{*}}+\frac{\tau'^{*}}{2}\right)D_{\nu^*}(\tau'^{*}) &= \nu^* D_{\nu^*-1}(\tau'^{*}),\\ \left(\frac{d}{d\tau'^{*}}-\frac{\tau'^{*}}{2}\right)D_{\nu^*}(\tau'^{*}) &= -D_{\nu^*+1}(\tau'^{*}),\end{aligned} \tag{98}$$

$$\begin{aligned}2\frac{d}{d\tau'^{*}}D_{\nu^*}(\tau'^{*}) &= \nu^* D_{\nu^*-1}(\tau'^{*}) - D_{\nu^*+1}(\tau'^{*}),\\ \tau'^{*}D_{\nu^*}(\tau'^{*}) &= \nu^* D_{\nu^*-1}(\tau'^{*}) + D_{\nu^*+1}(\tau'^{*}).\end{aligned} \tag{99}$$

The other necessary relations are obtained from these by the substitution $\tau'^{*} \longrightarrow -\tau^{*}$.

Because

$$\begin{aligned}c_{1n} &= \frac{\sqrt{2\pi}}{\Gamma(-i\kappa+1/2)}\exp\left[-\frac{\pi\kappa}{2}+\frac{i\pi}{4}\right],\\ \frac{i}{c^*_{1n}}N_0^2 &= \frac{\exp[3i\pi/4]}{2\sqrt{\pi eE}}\Gamma\left(i\kappa+\frac{1}{2}\right),\end{aligned} \tag{100}$$

we can write the propagator as

$$\begin{aligned}G^{\mu\nu}(x, x') &= \frac{\exp[3i\pi/4]}{2\sqrt{\pi eE}}\\ &\times\int\frac{d^3p}{(2\pi)^3}a^{\mu\nu}(x, x')\exp[i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')].\end{aligned} \tag{101}$$

The scalar particle propagator can be obtained from the right-hand side of (101) if we replace $a^{\mu\nu}(x, x')$ with expression (93). As an example, we now calculate $a^{33}(x, x')$. For $t > t'$, we have

$$a^{33}(x, x') \propto \sum_{i=1}^{3} {}^{+}\psi^3(i, x)_{+}\psi^{3*}(i, x'). \tag{102}$$

The first term in the sum is

$$\begin{aligned}{}^{+}\psi^3(1, x)_{+}\psi^{3*}(1, x') &\propto -\frac{ieE}{2}\tau^*\\ &\times D_{\nu^*}(-\tau^*)\tau'^{*}D_{\nu^*}(\tau'^{*})\frac{p_2^2}{m_\perp^2(m^2+p_1^2)},\end{aligned} \tag{103}$$

where we used the second equation in (99) and the one obtained from it by the substitution $\tau'^{*} \longrightarrow -\tau^{*}$. Similarly,

$$\begin{aligned}{}^{+}\psi^3(3, x)_{+}\psi^{3*}(3, x') &\propto -\frac{ieE}{2}\tau^*\\ &\times D_{\nu^*}(-\tau^*)\tau'^{*}D_{\nu^*}(\tau'^{*})\frac{p_1^2}{m^2(m^2+p_1^2)}.\end{aligned} \tag{104}$$

Adding (103) and (104), we obtain

$$\begin{aligned}&-\frac{ieE}{2}\tau^* D_{\nu^*}(-\tau^*)\tau'^{*}D_{\nu^*}(\tau'^{*})\\ &\times\left[\frac{p_2^2}{m_\perp^2(m^2+p_1^2)}+\frac{p_1^2}{m^2(m^2+p_1^2)}\right].\end{aligned} \tag{105}$$

The expression in the square brackets can be simplified as

$$\frac{1}{m^2+p_1^2}\left(\frac{p_2^2}{m_\perp^2}+\frac{p_1^2}{m^2}\right) = \frac{1}{m^2}-\frac{1}{m_\perp^2}. \tag{106}$$

The undesirable factor $(m^2 + p_1^2)^{-1}$ involved in (103) and (104) disappears in sum (105). The first term in the right-hand side of (106) gives the following contribution to (105):

$$\begin{aligned}-\frac{ieE}{2m^2}\tau^* D_{\nu^*}(-\tau^*)\tau'^{*}D_{\nu^*}(\tau'^{*}) &= \frac{1}{m^2}\\ \times(p_3-eEt)(p_3-eEt')D_{\nu^*}(-\tau^*)&D_{\nu^*}(\tau'^{*}).\end{aligned} \tag{107}$$

This already has the desired form. We now rewrite the contribution of the second term in the right-hand side of (106) to (105) in the initial form (i.e., before using the second equation in (99)),

$$\begin{aligned}\frac{ieE}{2m_\perp^2}[&-(1+\nu)^2 D_{\nu^*-1}(-\tau^*)D_{\nu^*-1}(\tau'^{*})\\ &+(1+\nu)D_{\nu^*+1}(-\tau^*)D_{\nu^*-1}(\tau'^{*})\\ &+(1+\nu)D_{\nu^*-1}(-\tau^*)D_{\nu^*+1}(\tau'^{*})\\ &-D_{\nu+1}(-\tau^*)D_{\nu^*+1}(\tau'^{*})].\end{aligned} \tag{108}$$

This expression still contains the undesirable factor $1/m_\perp^2$. But we must take the contribution from the term with $i = 2$ in (102) into account:

$$\begin{aligned}&{}^{+}\psi^3(2, x)_{+}\psi^{3*}(2, x') \propto \frac{ieE}{2m_\perp^2}(1+\nu)\\ &\times\Big[-D_{\nu^*-1}(-\tau^*)D_{\nu^*+1}(\tau'^{*}) - \frac{1}{\nu}D_{\nu^*+1}(-\tau^*)D_{\nu^*+1}(\tau'^{*})\\ &-\nu D_{\nu^*-1}(-\tau^*)D_{\nu^*-1}(\tau'^{*}) - D_{\nu^*+1}(-\tau^*)D_{\nu^*-1}(\tau'^{*})\Big].\end{aligned} \tag{109}$$

It is easy to see that in the sum of (108) and (109), the undesirable terms are cancelled and the unpleasant denominator

$$m_{\perp}^2 = -ieE(1+2\nu)$$

disappears:

$$\begin{aligned}(108)+(109)\\ = \frac{1}{2}\Big[(1+\nu)D_{\nu^*-1}(-\tau^*)D_{\nu^*-1}(\tau'^*)\\ +\frac{1}{\nu}D_{\nu^*+1}(-\tau^*)D_{\nu^*+1}(\tau'^*)\Big].\end{aligned}\tag{110}$$

Thus, $a^{33}(x, x'))$ is given by the sum of expressions (107) and (110). The first term in the right-hand side of (110) is used in (96) and the second term in (95). In the same manner, we find all the other $a^{\mu\nu}(x, x')$ components. Similar to the magnetic case,

$$\begin{aligned}G^{\mu\nu}(x,x') = \frac{eE}{(4\pi)^2}\\ \times\int_C \frac{ds}{s\sinh(eEs)}A^{\mu\nu}\exp\left[\frac{ieE}{2}z_3(t+t')\right]\\ \times\exp\left[-ism^2+\frac{i}{4s}(z_1^2+z_2^2)+\frac{i}{4s}(z_3^2-z_0^2)eE\coth(eEs)\right],\end{aligned}\tag{111}$$

where $A^{\mu\nu}$ is given by (71), but the vector potential is

$$A_\mu(x) = -\delta_{\mu 3}Et.$$

The nonzero $B^{\mu\nu}$ components are

$$\begin{aligned}B^{11} = B^{22} = 1, \quad B^{33} = -B^{00} = \cosh(2eEs),\\ B^{30} = -B^{03} = \sinh(2eEs).\end{aligned}\tag{112}$$

We see that the electric field dresses $B^{\mu\nu}$ with $\mu, \nu = 3, 0$.

Proceeding to the case where $t < t'$, we note that in accordance with (19),

$${}_{-}c_{\pm} = -{}^{+}c_{\pm}, \quad {}^{-}c_{\pm} = -{}_{+}c_{\pm}. \tag{112a}$$

This implies that ${}_{-}\psi({}^{-}\psi)$ is obtained from ${}^{+}\psi({}_{+}\psi)$ by changing the sign of the arguments in the parabolic cylinder functions and the sign of $\psi^0$ and $\psi^3$. The overall change of sign of $\psi(2, x)$ does not affect the corresponding term in (102). In $\psi(1, x)$ and $\psi(3, x)$, changing the sign of $\psi^0$ and $\psi^3$ and of the arguments $\tau^*$ and $\tau'^*$ is equivalent to changing the sign of only the $D$-function arguments $\tau^*$ and $\tau'^*$ if $\psi^0$ and $\psi^3$ are expressed through the left-hand sides of (99). As expected, it now follows from (93)–(96) that $G^{\mu\nu}(x, x')$ retains the same form (111) for $t < t'$.

## 7. THE VECTOR BOSON PROPAGATOR IN THE CONSTANT ELECTROMAGNETIC FIELD

After we have considered the magnetic and electric fields separately, the construction of the vector boson propagator in both fields meets no new problems. We take the vector potential in the form

$$A_\mu(x) = \delta_{\mu 2}Hx_1 - \delta_{\mu 3}Et. \tag{113}$$

The transition current between the states ${}_{+}\psi'$ and ${}_{+}\psi$ is

$$\begin{aligned}J^0({}_{+}\psi', {}_{+}\psi)\\ = 2\Big\{[c_1'^*c_1D_{n-1}^2(\zeta)+c_2'^*c_2D_{n+1}^2(\zeta)]D_{\nu^*}(\tau^*)i\overset{\leftrightarrow}{\frac{d}{dt}}D_\nu(\tau)\\ -\Big[c_-'^*c_+D_{\nu^*-1}(\tau^*)i\overset{\leftrightarrow}{\frac{d}{dt}}D_{\nu+1}(\tau)\\ +c_+'^*c_-D_{\nu^*+1}(\tau^*)i\overset{\leftrightarrow}{\frac{d}{dt}}D_{\nu-1}(\tau)\Big]D_n^2(\zeta)\Big\}.\end{aligned}\tag{114}$$

Taking Eq. (84) into account and integrating over $x_1$, we obtain

$$\begin{aligned}\int_{-\infty}^{\infty}dx_1J^0({}_{+}\psi', {}_{+}\psi) = n!\sqrt{\frac{2\pi E}{H}}2e^{\pi\kappa/2}\\ \times\left[\frac{1}{n}c_1'^*c_1+(1+n)c_2'^*c_2+i(c_-'^*c_+-c_+'^*c_-)\right].\end{aligned}\tag{115}$$

The constraint is given by

$$\begin{aligned}\sqrt{2eH}[(1+n)c_2-c_1]\\ +\sqrt{2eE}e^{-i\pi/4}[{}_{+}c_- - (1+\nu)\ {}_{+}c_+] = 0.\end{aligned}\tag{116}$$

Using (115) and (116), we find the ${}_{+}\psi$ polarization states (in what follows, the factor $\exp[i(p_2x_2 + p_3x_3)]$ is omitted for brevity):

$$\begin{aligned}{}_{+}\psi(1,x) = N(1)\\ \times\begin{bmatrix}(1+n)\sqrt{e^2EH}e^{i\pi/4}[D_{\nu+1}(\tau)-\nu D_{\nu-1}(\tau)]D_n(\zeta)\\ im_\perp^2D_\nu(\tau)D_{n+1}(\zeta)\\ m_\perp^2D_\nu(\tau)D_{n+1}(\zeta)\\ (1+n)\sqrt{e^2EH}e^{i\pi/4}[D_{\nu+1}(\tau)+\nu D_{\nu-1}(\tau)]D_n(\zeta)\end{bmatrix},\end{aligned}\tag{117}$$

$$N(i) = n_i N_0, \quad N_0 = \left(\frac{H}{2\pi E}\right)^{1/4} \frac{e^{-\pi\kappa/4}}{\sqrt{n!}},$$
$$n_1 = \frac{1}{\sqrt{2m_\perp^2(m^2 + eHn)(1+n)}}, \tag{118}$$

$${}_{+}\psi(2, x) = N(2)$$
$$\times \begin{bmatrix} [D_{\nu+1}(\tau) + (1+\nu)D_{\nu-1}(\tau)]D_n(\zeta) \\ 0 \\ 0 \\ [D_{\nu+1}(\tau) - (1+\nu)D_{\nu-1}(\tau)]D_n(\zeta) \end{bmatrix}, \tag{119}$$
$$n_2 = \sqrt{\frac{eE}{2m_\perp^2}},$$

$${}_{+}\psi(3, x) = N(3)$$
$$\times \begin{bmatrix} \sqrt{e^2EH}[D_{\nu+1}(\tau) - \nu D_{\nu-1}(\tau)]D_n(\zeta) \\ e^{i\pi/4}D_\nu(\tau)[-(m^2 + eHn)D_{n-1}(\zeta) + eHD_{n+1}(\zeta)] \\ e^{-i\pi/4}D_\nu(\tau)[(m^2 + eHn)D_{n-1}(\zeta) + eHD_{n+1}(\zeta)] \\ \sqrt{e^2EH}[D_{\nu+1}(\tau) + \nu D_{\nu-1}(\tau)]D_n(\zeta) \end{bmatrix}, \tag{120}$$

$$n_3 = \sqrt{\frac{n}{2m^2(m^2 + eHn)}}. \tag{121}$$

To obtain the polarization states of ${}^{+}\psi$ (or of ${}_{-}\psi$ and ${}^{-}\psi$), we again use Eq. (19) (cf. Eqs. (88)–(90) and (47)–(50)). We then obtain

$${}^{+}\psi(1, x) = N(1)\begin{bmatrix} (1+n)\sqrt{e^2EH}e^{-i\pi/4}[(1+\nu)D_{\nu^*-1}(-\tau^*) + D_{\nu^*+1}(-\tau^*)]D_n(\zeta) \\ im_\perp^2 D_{\nu^*}(-\tau^*)D_{n+1}(\zeta) \\ m_\perp^2 D_{\nu^*}(-\tau^*)D_{n+1}(\zeta) \\ (1+n)\sqrt{e^2EH}e^{-i\pi/4}[(1+\nu)D_{\nu^*-1}(-\tau^*) - D_{\nu^*+1}(-\tau^*)]D_n(\zeta) \end{bmatrix}, \tag{122}$$

$${}^{+}\psi(2, x) = N(2)$$
$$\times \begin{bmatrix} i(1+\nu)\left[-D_{\nu^*-1}(-\tau^*) + \dfrac{1}{\nu}D_{\nu^*+1}(-\tau^*)\right]D_n(\zeta) \\ 0 \\ 0 \\ i(1+\nu)\left[-D_{\nu^*-1}(-\tau^*) - \dfrac{1}{\nu}D_{\nu^*+1}(-\tau^*)\right]D_n(\zeta) \end{bmatrix}, \tag{123}$$

$${}^{+}\psi(3, x) = N(3)$$
$$\times \begin{bmatrix} -i\sqrt{e^2EH}[-\nu^* D_{\nu^*-1}(-\tau^*) + D_{\nu^*+1}(-\tau^*)]D_n(\zeta) \\ e^{i\pi/4}D_{\nu^*}(-\tau^*)[-(m^2 + eHn)D_{n-1}(\zeta) + eHD_{n+1}(\zeta)] \\ e^{-i\pi/4}D_{\nu^*}(-\tau^*)[(m^2 + eHn)D_{n-1}(\zeta) + eHD_{n+1}(\zeta)] \\ -i\sqrt{e^2EH}[-\nu^* D_{\nu^*-1}(-\tau^*) - D_{\nu^*+1}(-\tau^*)]D_n(\zeta) \end{bmatrix}. \tag{124}$$

The first and the fourth lines in the right-hand sides of (122) and (124) can be written in a more compact form using relations that can be obtained from (99) by the substitution $\tau'^* \longrightarrow -\tau^*$.

Further calculations are quite similar to those in Sections 5 and 6. The result was of course evident in advance: $A^{\mu\nu}$ is now given by (71) with the vector potential (113) and all the nonzero $B^{\mu\nu}$ are “dressed”, see Eqs. (73) and (112). The scalar particle propagator is given by

$$G_{\text{spin}0}(x, x') = \frac{e^2EH}{(4\pi)^2}\int_0^\infty \frac{ds}{\sinh(eEs)\sin(eHs)}$$
$$\times \exp\left\{-ism^2 + \frac{i}{4}[(z_1^2 + z_2^2)eH\cot(eHs)\right.$$
$$+ (z_3^2 - z_0^2)eE\coth(eEs)] \tag{125}$$
$$\left. + \frac{i}{2}[eEz_3(t + t') - eHz_2(x_1 + x_1')]\right\},$$
$$z_\mu = x_\mu - x'_\mu.$$

This expression agrees with the calculations by Ritus [10, 11]. The overall phase factor $e^{-i\pi/2}$ is involved in his formulas because of a different definition of the propagator. We also note that Eq. (125) is symmetric in $t$ and $t'$ and that

$$G_{\text{spin}0}(x, x', e) = G_{\text{spin}0}(x', x, -e).$$

Therefore,

$$
\begin{aligned}
G^{\mu\nu}(x, x') &= \frac{e^2EH}{(4\pi)^2}\int_0^\infty \frac{ds}{\sinh(eEs)\sin(eHs)}A^{\mu\nu} \\
&\times \exp\left\{-ism^2 + \frac{i}{4}[(z_1^2 + z_2^2)eH\cot(eHs)\right. \\
&\quad + (z_3^2 - z_0^2)eE\coth(eEs)] \\
&\quad \left. + \frac{i}{2}[eEz_3(t + t') - eHz_2(x_1 + x'_1)]\right\}.
\end{aligned} \tag{126}
$$

## ACKNOWLEDGMENTS

The author is grateful to V.S. Vanyashin, V. Incera, and E. Ferrer for stimulating discussions. Special thanks to V.I. Ritus for suggesting valuable improvements in the manuscript. This work was supported in part by the Russian Foundation for Basic Research (project no. 00-15-96566).

## REFERENCES

1. V. V. Skalozub, Fiz. Élem. Chastits At. Yadra **16**, 1005 (1985) [Sov. J. Part. Nucl. **16**, 445 (1985)].
2. V. M. Mostepanenko, V. M. Frolov, and V. A. Sheluto, Yad. Fiz. **37**, 1261 (1983) [Sov. J. Nucl. Phys. **37**, 750 (1983)].
3. V. S. Vanyashin and M. V. Terentyev, Zh. Eksp. Teor. Fiz. **48**, 565 (1965) [Sov. Phys. JETP **21**, 375 (1965)].
4. M. S. Marinov and V. S. Popov, Yad. Fiz. **15**, 1271 (1972) [Sov. J. Nucl. Phys. **15**, 702 (1972)].
5. A. A. Grib, S. G. Mamaev, and V. M. Mostepanenko, *Vacuum Quantum Effects in Strong Fields* (Énergoatomizdat, Moscow, 1988).
6. A. I. Nikishov, Tr. Fiz. Inst. Akad. Nauk SSSR **111**, 152 (1979); J. Sov. Laser Res. **6**, 619 (1985).
7. *Higher Transcendental Functions (Bateman Manuscript Project),* Ed. by A. Erdelyi (McGraw-Hill, New York, 1953; Nauka, Moscow, 1974), Vol. 2.
8. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics* (Nauka, Moscow, 1980; Pergamon, Oxford, 1982).
9. J. Schwinger, Phys. Rev. **82**, 664 (1951).
10. V. I. Ritus, Zh. Eksp. Teor. Fiz. **75**, 1560 (1978) [Sov. Phys. JETP **48**, 788 (1978)]; Tr. Fiz. Inst. Akad. Nauk SSSR **111**, 134 (1979); J. Sov. Laser Res. **6**, 497 (1985).
11. V. I. Ritus, Tr. Fiz. Inst. Akad. Nauk SSSR **168**, 52 (1986); A. I. Nikishov and N. B. Narozhny, Tr. Fiz. Inst. Akad. Nauk SSSR **168**, 175 (1986); in *Issues in Intense-Field Quantum Electrodynamics*, Ed. by V. L. Ginzburg (Nova Science, Commack, 1987).
12. A. I. Nikishov, Zh. Éksp. Teor. Fiz. **57**, 1210 (1969) [Sov. Phys. JETP **30**, 660 (1969)].

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 211–220.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 242–251.*
*Original English Text Copyright © 2001 by Ritus.*

## NUCLEI, PARTICLES, AND THEIR INTERACTION

# Physical Properties of Scalar and Spinor Field States with the Rindler–Milne (Hyperbolic) Symmetry[¶]

**V. I. Ritus**
*Tamm Department of Theoretical Physics, Lebedev Physical Institute,*
*Russian Academy of Sciences, Moscow, 117924 Russia*
*e-mail: ritus@lpi.ac.ru*

Received March 19, 2001

**Abstract**—It is shown that right and left combinations of the positive- and negative-frequency hyperbolically symmetric solutions of the Klein–Fock–Gordon equation possess an everywhere timelike current density vector with a definite Lorentz-invariant sign of the charge density, and similar combinations of solutions to the Dirac equation possess the energy-momentum tensor with everywhere real eigenvalues and a definite Lorentz-invariant sign of the energy density. These right and left modes, just as their ±-frequency components, are eigenfunctions of the Lorentz boost generator with the eigenvalue $\kappa$. The sign of the charge (energy) density coincides with the sign of $\kappa$ for the right scalar (spinor) modes and is opposite to it for the left modes. It is then reasonable to assume that the particles (antiparticles) are precisely described by the right modes with $\kappa > 0$ ($\kappa < 0$) and by the left modes with $\kappa < 0$ ($\kappa > 0$). *© 2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

Three complete sets of solutions of the Klein–Fock–Gordon (KFG) and Dirac equations are usually considered in relation to the Unruh effect [1]. One of these solution sets is the usual planewave set and the other two are the sets of field modes with a hyperbolic symmetry. The hyperbolically symmetric modes radically differ from the planewave modes by singularities occurring on the light cone. As a result, the corresponding charge and energy densities oscillate with increasing the frequency at Compton distances near the cone and become infinite on the cone. It is not surprising that the charge density of the scalar field and the energy density of the spinor field can have either sign near the singularity. This means that these modes contain both particles and antiparticles near the light cone. It is then difficult to distinguish the hyperbolically symmetric field state created by external sources on the light cone from the state created by the measuring device itself. Nevertheless, there exist right and left states with hyperbolic symmetry for which the charge density of the scalar field and the energy density of the spinor field possess an everywhere definite Lorentz-invariant sign.

## 2. PLANE WAVES WITH DEFINITE MOMENTUM AND FREQUENCY

For scalar plane waves

$$\varphi_p^{(\pm)}(x) = \frac{1}{\sqrt{2E}}\exp[i(pz \mp Et)], \qquad E = \sqrt{m^2+p^2}, \quad x^\alpha = (t, z), \tag{1}$$

the current densities $j_p^{(\pm)\alpha}(x) = (\pm 1, p/E)$ are timelike vectors. The signs of the charge densities coincide with the frequency signs. The energy–momentum tensor $t_{\alpha\beta}$ has the components

$$t_{00}^{(\pm)}, t_{33}^{(\pm)}, t_{03}^{(\pm)} = E, p^2/E, \mp p, \tag{2}$$

with $\operatorname{sgn} t_{00}^{(\pm)} > 0$.

For spinor plane waves with definite momentum and frequency and with the double spin projection $s$,

$$\begin{gathered}\chi_{ps}^{(\pm)}(x) = \varphi_{\mathbf{p}}^{(\pm)}(x)\sqrt{m}u_s^{(\pm)}(\theta),\\ \bar{u}_s^{(\omega)}(\theta)u_{s'}^{(\omega')}(\theta) = 2\omega\delta_{\omega\omega'}\delta_{ss'}\end{gathered} \tag{3}$$

(the bispinors $u_s^{(\pm)}(\theta)$ are given in (41) in the chiral representation in the transposed form), the current densities $j_{ps}^{(\pm)\alpha}(x) = (1, p/E)$ are timelike vectors with positive time components. The energy–momentum tensor $t_{\alpha\beta}$ corresponding to (3) has the components

$$t_{00}^{(\pm)}, t_{33}^{(\pm)}, t_{03}^{(\pm)} = \pm E, \pm p^2/E, -p, \tag{4}$$

where $\operatorname{sgn} t_{00}^{(\pm)} \gtrless 0$.

A superposition of the scalar positive- (negative-) frequency plane waves, unlike the partial waves themselves, does not possess a definite positive (negative) charge density in general. Thus, if

$$\phi(x) = \int dp c_p \varphi_p^{(+)}(x), \tag{5}$$

the charge density may not be everywhere positive because of oscillations of the integrand in the representation

---

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

$$j^0(x) = i\phi^*(x)\overleftrightarrow{\partial}_t\phi(x) = \iint\frac{dpdp'(E+E')}{\sqrt{2E2E'}} \tag{6}$$
$$\times \exp\{i[(p'-p)z-(E'-E)t]\}c_p^*c_{p'}.$$

However, the total charge of the packet is positive and time-independent,

$$Q = \int dz j^0(x) = 2\pi\int dp|c_p|^2. \tag{7}$$

Similarly, a superposition of the spinor positive- (negative-) frequency plane waves does not possess an everywhere positive (negative) energy density in general. Thus, the positive-frequency wave packet

$$\chi_s(x) = \int dpc_p\chi_{ps}^{(+)}(x) \tag{8}$$

has the energy density

$$t_{00}(x) = \frac{1}{2}i\chi_s^+(x)\overleftrightarrow{\partial}_t\chi_s(x) = \frac{1}{2}\iint\frac{dpdp'(E+E')}{\sqrt{2E2E'}} \tag{9}$$
$$\times \exp\{i[(p'-p)z-(E'-E)t]\}mu_s^{(+)+}(\theta)u_s^{(+)}(\theta')c_p^*c_{p'}$$

that may not be everywhere positive, but the total energy of the packet is positive and conserved,

$$\mathscr{E} = \int dzt_{00}(x) = 2\pi\int dp|c_p|^2E(p). \tag{10}$$

The negative charge (energy) density for a positive-frequency scalar (spinor) wave packet can occur because the packet is nonstationary (cannot be represented as $\exp(-iEt)f(z)$, $E > 0$). Expressions (6) and (9) imply that the time-averaged values of the charge and energy densities are equal to zero at any point in space. This means that charge and energy come from infinity and go to infinity. In a finite region of space $\Delta z$, they can therefore reach perceptible values $\Delta Q$ and $\Delta\mathscr{E}$ only for a finite time interval $\Delta t$. In addition, each of the quantities

$$\Delta Q(t) = \int\limits_{\Delta z} dzj^0(x)$$

and

$$\Delta\mathscr{E}(t) = \int\limits_{\Delta z} dzt_{00}(x)$$

can also be negative. This indicates the appearance of the antiparticle in this space–time region.

## 3. POSITIVE- AND NEGATIVE-FREQUENCY SCALAR WAVES WITH HYPERBOLIC SYMMETRY

These scalar waves are defined by the integral representation [2]

$$\phi_\kappa^{(\pm)}(x) = \frac{1}{2}\int\limits_{-\infty}^{\infty}d\theta\exp[i(pz\mp Et)\mp i\kappa\theta], \tag{11}$$
$$p = m\sinh\theta,\quad E = m\cosh\theta,$$

where $\theta = \operatorname{arctanh}(p/E)$ is the rapidity. In the right and left sectors of the Minkowski plane, these functions can be represented by the Macdonald function of a real argument,

$$\begin{gathered}\exp(\pm\pi\kappa/2 - i\kappa\nu)K_{i\kappa}(\zeta),\\ \exp(\mp\pi\kappa/2 - i\kappa\nu)K_{i\kappa}(\zeta),\\ \zeta = m\sqrt{z^2-t^2},\quad \nu = \operatorname{arctanh}(t/z),\end{gathered} \tag{12}$$

and in the future and past sectors by the Macdonald function of an imaginary argument,

$$\begin{gathered}\exp(-i\kappa w)K_{i\kappa}(\pm i\tau),\quad \exp(-i\kappa w)K_{i\kappa}(\mp i\tau),\\ \tau = m\sqrt{t^2-z^2},\quad w = \operatorname{arctanh}(z/t).\end{gathered} \tag{13}$$

Using the Rindler metric

$$ds^2 = dz'^2 - (az')^2dt'^2$$

in the $R$ and $L$ sectors and the Milne metric

$$ds^2 = (at')^2dz'^2 - dt'^2$$

in the $F$ and $P$ sectors, we can write

$$\zeta = \pm mz',\quad \nu = \pm at',$$
$$\tau = \pm mt',\quad w = \pm az',$$

where $z'$ and $t'$ are space and time coordinates in the Rindler or Milne spaces, see [3]. It is essential that $\phi_\kappa^{(\pm)}$ includes plane waves with unlimited energy.

The scalar waves have the following properties.

(a) $\phi_\kappa^{(\pm)}(x)$ are analytical and finite functions in the lower/upper half-planes of the complex variables $x_+ = t + z$ and $x_- = t - z$.

(b) The hyperbolic symmetry implies that $\phi_\kappa^{(\pm)}$ are eigenfunctions of the Lorentz boost operator: under the transformation

$$z, t \longrightarrow z' = \frac{z-\beta t}{\sqrt{1-\beta^2}},\quad t' = \frac{t-\beta z}{\sqrt{1-\beta^2}}, \tag{14}$$

the variables $\zeta$ and $\tau$ remain invariant, while the cyclic variables $\nu$ and $w$ go to $\nu' = \nu - \alpha$ and $w' = w - \alpha$, where $\alpha = \operatorname{arctanh}\beta$ is the rapidity corresponding to the Lorentz transformation velocity $\beta$. Then

$$\phi_\kappa^{(\pm)}(z, t) \longrightarrow \phi_\kappa^{(\pm)}(z', t') = e^{i\alpha\kappa}\phi_\kappa^{(\pm)}(z, t), \tag{15}$$

and therefore, $e^{i\alpha\kappa}$ is an eigenvalue of the Lorentz boost operator $e^{-\alpha\partial_\nu}$ or $e^{-\alpha\partial_w}$; $\kappa$ is an eigenvalue of the Lorentz boost generator $i(t\partial_z + z\partial_t) = i\partial_\nu$ or $i\partial_w$ and is interpreted as the frequency for a Rindler observer or the momentum for a Milne observer.

(c) $\phi_\kappa^{(+)}$ and $\phi_\kappa^{(-)}$ are related by complex conjugation accompanied by changing the sign of $\kappa$,

$$\phi_\kappa^{(+)*}(x) = \phi_{-\kappa}^{(-)}(x), \quad \phi_{-\kappa}^{(\pm)}(z, t) = \phi_\kappa^{(\pm)}(-z, t). \qquad (16)$$

The complex conjugation is equivalent to time reflection. The last property is equivalent to space reflection.

(d) As a striking property of $\phi_\kappa^{(\pm)}$, we note that although the current density vectors corresponding to the plane wave components of $\phi_\kappa^{(\pm)}$ are everywhere timelike, the current densities $j_\kappa^{(\pm)\alpha}$ corresponding to $\phi_\kappa^{(\pm)}$ themselves are not timelike vectors in the entire Minkowski space: there are space–time regions inside the light cone where the current densities are spacelike.

The current density $j^\alpha$ for the Minkowski observer is related to the current density $J^\alpha$ for Rindler or Milne observers (more exactly, for local Lorentz observers momentarily comoving to them) by the Lorentz transformation

$$j^0 = \frac{J^0 + \beta J^3}{\sqrt{1-\beta^2}}, \quad j^3 = \frac{J^3 + \beta J^0}{\sqrt{1-\beta^2}}. \qquad (17)$$

For the Rindler observer with $\beta = t/z$ in the $R$ sector, we have

$$J_\kappa^{(\pm)0} = \frac{2m\kappa e^{\pm\pi\kappa}}{\zeta} K_{i\kappa}^2(\zeta), \quad J^{(\pm)3} = 0. \qquad (18)$$

For the $L$ sector, we must replace $e^{\pm\pi\kappa} \longrightarrow -e^{\mp\pi\kappa}$. The current density vector is timelike.

For the Milne observer with $\beta = z/t$, we have

$$\begin{aligned} J_\kappa^{(\pm)0} &= \pm\frac{\pi}{\sqrt{t^2-z^2}}, \\ J_\kappa^{(\pm)3} &= -\operatorname{sgn} t \frac{2m\kappa}{\tau}|K_{i\kappa}(i\tau)|^2. \end{aligned} \qquad (19)$$

The current density squared

$$(j_\kappa^{(\pm)})^2 = -\frac{\pi^2}{t^2-z^2}\left(1 - \frac{4\kappa^2}{\pi^2}|K_{i\kappa}(i\tau)|^4\right) \qquad (20)$$

can have either sign when $\tau = m\sqrt{t^2-z^2} \ll 1$, but is negative for $\tau \gtrsim 1$.

Thus, inside the light cone at invariant distances less than the Compton length from the cone, there are spacetime regions where the current densities $j_\kappa^{(\pm)\alpha}$ are spacelike vectors and the charge density $j_\kappa^{(\pm)0}$, $\kappa > 0$, is negative, while $j_\kappa^{(-)0}$, $\kappa > 0$, is positive. Because the current densities are timelike vectors for the real particles, we can relate the spacelike current density $j_\kappa^{(+)}$ to antiparticles of the virtual pairs created in regions with a very high energy concentration. The total charge of the $\phi_\kappa^{(+)}$ state on any spacelike surface in Minkowski space is positive and is equal to the charge on this surface entirely situated in the $P$, $L + R$ or $F$ sector. But the charge density $j^0$ for this state with $\kappa > 0$ is positive only in $R$ sector, is negative in the $L$ sector, and can have either sign in the $P$ and $F$ sectors.

Thus, unlike the sign of the total charge, the sign of the charge density is not well defined by the frequency sign of the $\phi_\kappa^{(+)}$ states. This situation occurs in external field problems due to a possible pair creation by the external field, or in problems of forming wave packets with a high energy density. The appearance of a negative charge density in the $P$, $F$, and $L$ sectors for the positive-frequency state $\phi_\kappa^{(+)}$ is a consequence of the hyperbolic symmetry of the state. The hyperbolic symmetry divides Minkowski space into spacelike and timelike subspaces with the Rindler and Milne metrics. These metrics have singularities on the light cone (which is their common boundary) and can be considered as a limiting case of a global nonsingular smooth metric of the space with a nonzero external field near the light cone. The pair creation by this field is then possible and the appearance of a negative charge density in the positive-frequency state $\phi_\kappa^{(+)}$ after switching the field off can be understood.

The states $\phi_\kappa^{(+)}$ and $\phi_\kappa^{(-)}$ possess, respectively, the positive and negative total charge but do not possess an everywhere positive and negative charge density. This means that both the particle and the antiparticle can be detected in any of these states.

## 4. RIGHT AND LEFT SCALAR MODES

In each of the $R$ and $L$ sectors, $\phi_\kappa^{(+)}$ and $\phi_\kappa^{(-)}$ differ only by factors. According to Unruh [1], one can find remarkable right and left combinations

$$\phi_\kappa^R = \alpha_\kappa\phi_\kappa^{(+)} + \beta_\kappa\phi_\kappa^{(-)}, \quad \phi_\kappa^L = \beta_\kappa\phi_\kappa^{(+)} + \alpha_\kappa\phi_\kappa^{(-)}, \qquad (21)$$

such that $\phi_\kappa^R = 0$ in the $L$ sector and $\phi_\kappa^L = 0$ in the $R$ sector. In these combinations,

$$\begin{gathered} \beta_\kappa = -\alpha_\kappa e^{-\pi\kappa}, \quad \alpha_\kappa = \frac{e^{\pi\kappa/2}}{\sqrt{2\sinh\pi\kappa}}, \\ |\alpha_\kappa|^2 - |\beta_\kappa|^2 = \varepsilon_\kappa \equiv \operatorname{sgn}\kappa. \end{gathered} \qquad (22)$$

For $\kappa < 0$, we have $\sqrt{\sinh\pi\kappa} = i\sqrt{\sinh\pi|\kappa|}$. The set $\phi_\kappa^{R,L}$ possesses the same hyperbolic symmetry as the set $\phi_\kappa^{(\pm)}$, but the striking property of these functions is that the corresponding current densities $j_\kappa^{R\alpha}$ and $j_\kappa^{L\alpha}$

are timelike vectors in the entire spacetime region where they are nonzero. The Lorentz transformation (17) again relates the current density $j^\alpha$ for the Minkowski observer to the current density $J^\alpha$ for the Rindler or Milne observers.

For the Rindler observer with $\beta = t/z$, we have

$$J_\kappa^{R0} = \frac{4m\kappa\sinh\pi|\kappa|}{\zeta}|K_{i\kappa}(\zeta)|^2, \quad J_\kappa^{R3} = 0. \tag{23}$$

The current density vector is then timelike.

For the Milne observer with $\beta = z/t$, we have

$$\begin{aligned} J_\kappa^{R,L0} &= \pm\frac{\operatorname{sgn}(\kappa t)\pi}{\sqrt{t^2-z^2}}, \\ J_\kappa^{R,L3} &= -\frac{\operatorname{sgn}(\kappa t)\pi^2\kappa}{\sqrt{t^2-z^2}\sinh\pi\kappa}|J_{i\kappa}(\tau)|^2. \end{aligned} \tag{24}$$

The Lorentz invariant current density squared is nonpositive,

$$\begin{aligned} (j_\kappa^{R,L})^2 &= -\frac{\pi^2}{t^2-z^2} \\ &\times\left[1-\left(\frac{\pi\kappa}{\sinh\pi\kappa}\right)^2|J_{i\kappa}(\tau)|^4\right] \le 0, \end{aligned} \tag{25}$$

for all real $\kappa$ and $\tau \ge 0$ [4]. The current density vector is timelike.

It is interesting to note that in the $R$ sector, the current density squared $(j_\kappa^R)^2$ tends to infinity as $\zeta \longrightarrow 0$, but in the $P$ or $F$ sectors, it is finite at $\tau = 0$:

$$(j_\kappa^R)^2\Big|_{\tau\to 0} = -\frac{\pi^2 m^2}{1+\kappa^2} + \ldots. \tag{26}$$

The state $\phi_\kappa^R$ ($\phi_\kappa^L$) describes a wave with hyperbolic symmetry and charge density that is only positive for $\kappa > 0$ ($\kappa < 0$) or only negative for $\kappa < 0$ ($\kappa > 0$). We can then say that the respective state describes the particle or the antiparticle. In other words, the state $\phi_\kappa^L$ describes the particle or the antiparticle with the sign of $\kappa$ that is opposite to the sign used in describing for the $\phi_\kappa^R$ state [4].

We note that complex conjugation (time reflection) of the functions $\phi_\kappa^{R,L}$ is equivalent to changing the sign of $\kappa$, while the space reflection is equivalent to changing the sign of $\kappa$ and replacing $R \rightleftarrows L$:

$$\begin{gathered} \phi_\kappa^{R*}(x) = \phi_{-\kappa}^R(x), \quad \phi_\kappa^{L*}(x) = \phi_{-\kappa}^L(x); \\ \phi_\kappa^R(-z, t) = i\phi_{-\kappa}^L(z, t). \end{gathered} \tag{27}$$

In the $R$ sector, where $\kappa$ is interpreted as energy by the Rindler observer and $\phi_\kappa^L = 0$, particles are described by the functions $\phi_\kappa^R$, $\kappa > 0$, and antiparticles by the complex conjugate functions, i.e., by $\phi_\kappa^R$, $\kappa < 0$. In the $F$ or $P$ sectors, where $\kappa$ is interpreted as momentum by the Milne observer, particles with the momentum $\kappa$ are described by the functions $\phi_\kappa^R$, $\kappa > 0$, and $\phi_\kappa^R$, $\kappa < 0$, while antiparticles with the same momenta are described by the complex conjugate functions $\phi_{-\kappa}^R$ and $\phi_{-\kappa}^L$.

The completeness of the sets $\phi_\kappa^{(\pm)}$ and $\phi_\kappa^{R,L}$ is expressed by

$$\begin{aligned} \Delta^{(\pm)}(x-x') &= \pm i\int_{-\infty}^{\infty}\frac{d\kappa}{2\pi^2}\phi_\kappa^{(\pm)}(x)\phi_\kappa^{(\pm)*}(x') \\ &= \frac{\pm i}{2\pi}K_0(m\sqrt{y^2}) \quad \text{if} \quad y^2 > 0, \\ &= \frac{1}{4}[\varepsilon(y^0)J_0(m\sqrt{|y^2|}) \mp iN_0(m\sqrt{|y^2|})] \\ &\quad \text{if} \quad y^2 < 0, y = x - x', \end{aligned} \tag{28}$$

$$\begin{aligned} \Delta(y) &= \sum_\pm \Delta^{(\pm)}(y) \\ &= i\int_{-\infty}^{\infty}\frac{d\kappa}{2\pi^2}\varepsilon_\kappa[\phi_\kappa^R(x)\phi_\kappa^{R*}(x') - \phi_\kappa^L(x)\phi_\kappa^{L*}(x')] \\ &= \frac{1}{2}\varepsilon(y^0)\theta(-y^2)J_0(m\sqrt{|y^2|}). \end{aligned} \tag{29}$$

It is interesting to note that the analytical properties of the functions $\phi_\kappa^R$ and $\phi_\kappa^L$ in each of the variables $u$ and $v$ are similar to the properties of the Pauli–Jordan function $\Delta(x)$ in $x^2$. Indeed, $\Delta(x)$ is also equal to the sum of the positive-frequency and negative-frequency functions $\Delta^\pm(x)$, which are boundary values of some function $F(x^2)$ that is analytical in the complex plane of $x^2$ cut along the real negative semi-axis $x^2 < 0$:

$$\Delta^{(\pm)}(x) = \pm F(x^2 \pm i\varepsilon\operatorname{sgn}x^0), \quad \varepsilon \longrightarrow +0.$$

It follows that $\Delta(x)$ differs from zero only for $x^2 < 0$ and is equal to the jump of $F(x^2)$ on the cut.

The solution of the Cauchy problem and the normalization condition are given by

$$\begin{gathered} \phi(y) = \int_S d\sigma^\alpha\Delta(y-x)\overleftrightarrow{\partial}_\alpha\phi(x), \\ i\int_S d\sigma^\alpha\phi_\kappa^{(\omega)*}\overleftrightarrow{\partial}_\alpha\phi_{\kappa'}^{(\omega')} = 2\pi^2\omega\delta(\kappa-\kappa')\delta_{\omega\omega'}, \end{gathered} \tag{30}$$

where $S$ is a spacelike surface in Minkowski space or in any of the $P$, $L + R$, $F$ sectors. For the functions $\phi_\kappa^a$ and $\phi_{\kappa'}^{a'}$, $a, a' \in R, L$, the right-hand side of the normalization condition is $2\pi^2\varepsilon_\kappa\varepsilon_a\delta(\kappa - \kappa')\delta_{aa'}$, where $\varepsilon_R = -\varepsilon_L = 1$. In accordance with the normalization condition, all the states have the same magnitude of the conserved total charge; the sign of the charge coincides with the frequency sign for the $\phi_\kappa^{(\pm)}$ states and with the sign of the product $\varepsilon_\kappa\varepsilon_a$ for the $\phi_\kappa^a$ states, $a \in R, L$.

An arbitrary solution of the KFG equation can be represented by the expansions

$$\phi(x) = \int_{-\infty}^{\infty} \frac{dp}{\sqrt{2E}} \times\left[c_p \exp[i(pz - Et)] + d_p^* \exp[i(pz + Et)]\right] \tag{31}$$

$$\begin{aligned} &= \int_{-\infty}^{\infty} \frac{d\kappa}{2\pi^2}[a_\kappa\phi_\kappa^{(+)}(x) + b_\kappa^*\phi_\kappa^{(-)}(x)] \\ &= \int_{-\infty}^{\infty} \frac{d\kappa}{2\pi^2}[r_\kappa\phi_\kappa^R(x) + l_\kappa^*\phi_\kappa^L(x)]. \end{aligned} \tag{32}$$

As an example, we consider

$$\phi(x) = \frac{1}{\sqrt{2E_1}}\exp[i(p_1 z - E_1 t)],$$

$$c_p = \delta(p - p_1), \quad d_p = 0.$$

It then follows that

$$a_\kappa = \frac{2\pi}{\sqrt{2E_1}}e^{i\kappa\theta_1}, \quad b_\kappa = 0, \quad \theta_1 = \operatorname{arctanh}(p_1/E_1), \tag{33}$$

$$\begin{aligned} r_\kappa &= \varepsilon_\kappa \frac{\pi\exp(\pi\kappa/2 + i\kappa\theta_1)}{(\sqrt{E_1}\sinh\pi\kappa)^*}, \\ l_\kappa^* &= \varepsilon_\kappa \frac{\pi\exp(-\pi\kappa/2 + i\kappa\theta_1)}{(\sqrt{E_1}\sinh\pi\kappa)^*}. \end{aligned} \tag{34}$$

The spectra are given by (with $g_1 = 2\pi^2/E_1$)

$$\begin{aligned} |r_\kappa|^2 &= \frac{g_1 e^{2\pi\kappa}}{e^{2\pi\kappa} - 1}, \quad \kappa > 0, \quad j^0 > 0; \\ &\frac{g_1}{e^{2\pi|\kappa|} - 1}, \quad \kappa < 0, \quad j^0 < 0, \end{aligned} \tag{35}$$

$$\begin{aligned} |l_\kappa|^2 &= \frac{g_1}{e^{2\pi\kappa} - 1}, \quad \kappa > 0, \quad j^0 < 0; \\ &\frac{g_1 e^{2\pi|\kappa|}}{e^{2\pi|\kappa|} - 1}, \quad \kappa < 0, \quad j^0 > 0. \end{aligned} \tag{36}$$

There are no reasons to associate these spectra with thermodynamic ones, especially for a uniformly moving Milne observer, for whom $\kappa$ is not the energy but the momentum, and all the more so for a Minkowski observer, for whom $\kappa$ is an eigenvalue of the Lorentz boost generator and is odd under space and time reflections. We have

$$\begin{aligned} \phi_\kappa^{(+)} &= \varepsilon_\kappa(\alpha_\kappa^*\phi_\kappa^R - \beta_\kappa^*\phi_\kappa^L), \\ \phi_\kappa^{(-)} &= \varepsilon_\kappa(\alpha_\kappa^*\phi_\kappa^L - \beta_\kappa^*\phi_\kappa^R), \end{aligned} \tag{37}$$

$$|\alpha_\kappa|^2 = \frac{e^{2\pi\kappa}}{e^{2\pi\kappa} - 1}, \quad |\beta_\kappa|^2 = \frac{1}{e^{2\pi\kappa} - 1}, \quad \kappa > 0, \tag{38}$$

$$|\alpha_\kappa|^2 = \frac{1}{e^{2\pi|\kappa|} - 1}, \quad |\beta_\kappa|^2 = \frac{e^{2\pi|\kappa|}}{e^{2\pi|\kappa|} - 1}, \quad \kappa < 0, \tag{39}$$

where $|\beta_\kappa/\alpha_\kappa|^2$ is the probability to find any nonzero number of pairs and $|\alpha_\kappa|^{-2}$ is the probability to find no pairs in the state $\phi_\kappa^{(+)}$, $\kappa > 0$, etc., cf. [5]. This interpretation follows from the none-one-particle consideration of the wave equation solutions and does not require transition to the secondary quantization, although it is confirmed by it [6].

We note that the modes $\phi_\kappa^{R,L}(x)$ with $\kappa = 0$ are not defined by Eq. (21) because the coefficients $\alpha_\kappa$ and $\beta_\kappa$ are infinite at $\kappa = 0$. The term with $\kappa = 0$ in expansions (32) of an arbitrary solution of the KFG equation is nevertheless finite and can be defined as the $\kappa \longrightarrow 0$ limit of

$$r_\kappa\phi_\kappa^R + l_\kappa^*\phi_\kappa^L \equiv a_\kappa\phi_\kappa^{(+)} + b_\kappa^*\phi_\kappa^{(-)}\Big|_{\kappa \to 0} = a_0\phi_0^{(+)} + b_0^*\phi_0^{(-)}.$$

A similar remark applies to the term with $\kappa = 0$ in expansion (29).

## 5. DIRAC EQUATION SOLUTIONS WITH THE HYPERBOLIC SYMMETRY

Solutions $\psi_{\kappa s}^{(\pm)}$ of the Dirac equation in the Rindler or Milne space are related to solutions $\chi_{\kappa s}^{(\pm)}$ of this equation in Minkowski space by the Lorentz transformation

$$\begin{aligned} \psi_{\kappa s}^{(\pm)}(x) &= e^{-\alpha\alpha_3/2}\chi_{\kappa s}^{(\pm)}(x), \quad \alpha = \operatorname{arctanh}\beta, \\ \alpha_3 &= \operatorname{diag}(\sigma_3, -\sigma_3), \end{aligned} \tag{40}$$

where $\beta = t/z$ or $z/t$ for the Rindler or the Milne space respectively. We use the chiral representation

$$\chi_{\kappa s}^{(\pm)}(x) = \frac{1}{2}\int_{-\infty}^{\infty} d\theta \exp[i(pz \mp Et) \mp i\kappa\theta] u_s^{(\pm)}(\theta),$$

$$\begin{gathered} p = m\sinh\theta, \quad E = m\cosh\theta, \\ \tilde{u}_1^{(\pm)}(\theta) = (e^{\pm\theta/2}, 0, \pm e^{\mp\theta/2}, 0), \\ \tilde{u}_{-1}^{(\pm)}(\theta) = (0, \pm e^{\mp\theta/2}, 0, e^{\pm\theta/2}), \end{gathered} \tag{41}$$

where $s = \pm 1$ are the eigenvalues of the matrix

$$\Sigma_3 = \mathrm{diag}(\sigma_3, \sigma_3).$$

This representation defines the bispinor $\chi_{\kappa s}^{(+)}(x)$ ($\chi_{\kappa s}^{(-)}(x)$) as an analytical function in the lower (upper) half-plane of the respective complex variable $x_+ = t + z$ and $x_- = t - z$.

Bispinor components of $\psi_{\kappa s}$ and $\chi_{\kappa s}$ can be expressed through the Macdonald functions with the indices $i\kappa \pm 1/2$. For example, in the $R$ and $F$ sectors, $\psi_{\kappa 1}^{(\pm)}$ can be represented by the respective expression

$$\begin{gathered} \exp(\pm\pi\kappa/2 \pm i\pi/4 - i\kappa v)\begin{pmatrix} K_{i\kappa - 1/2}(\zeta) \\ 0 \\ -iK_{i\kappa + 1/2}(\zeta) \\ 0 \end{pmatrix} \\ \text{and} \quad \exp(-i\kappa w)\begin{pmatrix} K_{i\kappa - 1/2}(\pm i\tau) \\ 0 \\ \pm iK_{i\kappa + 1/2}(\pm i\tau) \\ 0 \end{pmatrix}. \end{gathered} \tag{42}$$

In other sectors, these functions can be obtained using the symmetry relations

$$\begin{aligned} \psi_{\kappa s}^{(\pm)}(t, z) &= \alpha_3 \psi_{\kappa s}^{(\mp)}(-t, -z) \\ &= \pm\beta\psi_{-\kappa s}^{(\pm)}(t, -z) = \pm\beta\psi_{\kappa s}^{(\pm)*}(-t, z), \end{aligned} \tag{43}$$

where

$$\alpha_3 = \begin{pmatrix} \sigma_3 & 0 \\ 0 & -\sigma_3 \end{pmatrix}, \quad \beta = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{44}$$

The functions $\psi_{\kappa -1}^{(\pm)}$ with the opposite spin direction can be obtained from (42) by transposing the first row elements with the fourth row and the third row elements with the second row.

The orthogonality and normalization condition for $\psi_{\kappa s}^{(\pm)}$ is

$$\begin{aligned} &\int_S d\sigma_\alpha \bar{\psi}_{\kappa s}^{(\omega)}(x)\gamma^\alpha(x)\psi_{\kappa' s'}^{(\omega')}(x) \\ &= \frac{2\pi^2}{m}\delta_{\omega\omega'}\delta_{ss'}\delta(\kappa - \kappa'). \end{aligned} \tag{45}$$

This involves an oriented surface element $d\sigma_\alpha = n_\alpha d\sigma$, where $d\sigma$ is the invariant surface measure and $n_\alpha$ is the timelike normal to the surface. Because $\psi_{\kappa s}^{(\pm)}$ are solutions of the covariant Dirac equation with the coordinate-dependent metric $g_{\alpha\beta}$ and the matrices $\gamma^\alpha$ (see, e.g., § 3.8 in [3]), the normalization condition for these functions also contains $\gamma^\alpha(x)$ and it is convenient to choose the spacelike integration surface $S$ entirely in one of the $P$, $L + R$ or $F$ subspaces with either the Milne or the Rindler metric. For a constant $t'$ surface $S$, the surface element reduces to

$$d\sigma_0 = dz'\sqrt{\gamma}n_0, \quad n_0 = \sqrt{-g_{00}}$$

and $\gamma = |g_{33}|$ is the determinant of the space metric.

Because the Rindler and Milne spaces and the corresponding metrics only represent nonstandard coordinate forms of the flat space–time, the solutions $\psi_{\kappa s}^{(\pm)}$ must be related to the solutions $\chi_{\kappa s}^{(\pm)}$ of the usual Dirac equation in Minkowski space by a Lorentz transformation. These solutions satisfy the same symmetry relations (43) and orthogonality and normalization condition (45) with the standard $\gamma$ matrices. For a constant $t$ surface $S$, the surface element becomes $d\sigma_0 = dz$ and the right-hand side of (45) immediately follows when one uses integral representation (41) for $\chi_{\kappa s}^{(\pm)}$ and performs the integration over $z$ first.

In representation (42), the functions $\chi_{\kappa s}^{(\pm)}$ differ from $\psi_{\kappa s}^{(\pm)}$ by the factors $e^{v/2}$ and $e^{-v/2}$ of the first and the third bispinor elements in the $R$ sector and by $e^{w/2}$ and $e^{-w/2}$ in the $F$ sector.

Under Lorentz transformation (14), the functions $\chi_{\kappa s}^{(\pm)}$ go to

$$\begin{gathered} \chi_{\kappa s}^{(\pm)}(x') = \exp(i\alpha\kappa - \alpha\alpha_3/2)\chi_{\kappa s}^{(\pm)}(x), \\ \alpha = \mathrm{arctanh}\,\beta, \quad \alpha_3 = \mathrm{diag}(\sigma_3, -\sigma_3). \end{gathered} \tag{46}$$

The eigenvalues are again independent of the frequency sign. The current densities $j^\alpha$ and $J^\alpha$ for the Minkowski and Rindler or Milne observers are again related by (17).

For the Rindler observer with $\beta = t/z$, we have in the $R$ sector:

$$J_\kappa^{(\pm)0} = 2e^{\pm\pi\kappa}|K_{i\kappa-1/2}(\zeta)|^2, \quad J_\kappa^{(\pm)3} = 0. \tag{47}$$

For the $L$ sector, we must replace $e^{\pm\pi\kappa} \longrightarrow e^{\mp\pi\kappa}$.

For the Milne observer with $\beta = z/t$, we have

$$\begin{aligned} J_\kappa^{(\pm)0} &= |K_{i\kappa-1/2}(i\tau)|^2 + |K_{i\kappa+1/2}(i\tau)|^2, \\ J_\kappa^{(\pm)3} &= \pm|K_{i\kappa-1/2}(i\tau)|^2 \mp |K_{i\kappa+1/2}(i\tau)|^2. \end{aligned} \tag{48}$$

The current density is a timelike vector and its time component is positive (a well known fact for the spinor field). But the striking feature of $\chi_{\kappa s}^{(\pm)}$ is that the eigenvalues of the corresponding energy-momentum tensor $t_{\alpha\beta}^{(\pm)}$ are not everywhere real. There are some places inside the light cone where these eigenvalues are complex conjugate.

The energy–momentum tensors $t_{\alpha\beta}$ and $T_{\alpha\beta}$ for the Minkowski and Rindler or Milne observers are related by the Lorentz transformation

$$\begin{aligned} t_{00} &= \gamma^2(T_{00} - 2\beta T_{03} + \beta^2 T_{33}), \\ t_{33} &= \gamma^2(T_{33} - 2\beta T_{03} + \beta^2 T_{00}), \\ t_{03} &= \gamma^2[T_{03}(1+\beta^2) - \beta T_{00} - \beta T_{33}], \\ \gamma &= (1-\beta^2)^{-1/2}. \end{aligned} \tag{49}$$

For the Rindler observer with $\beta = t/z$ in the $R$ sector, we have

$$\begin{aligned} &T_{00}^{(\pm)}, T_{33}^{(\pm)}, T_{03}^{(\pm)} \\ &= \frac{2m\kappa e^{\pm\pi\kappa}}{\zeta}\left(|K_{i\kappa-1/2}(\zeta)|^2, \int_\zeta^\infty \frac{d\zeta}{\zeta}|K_{i\kappa-1/2}(\zeta)|^2, 0\right). \end{aligned} \tag{50}$$

For the $L$ sector, we replace $e^{\pm\pi\kappa} \longrightarrow -e^{\mp\pi\kappa}$.

For the Milne observer with $\beta = z/t$, we have

$$\begin{aligned} T_{00}^{(\pm)} &= \pm\frac{m\kappa}{\tau}\left(\int_\tau^\infty \frac{d\tau}{\tau}A(\tau) + \frac{\pi}{\kappa}\right), \\ T_{33}^{(\pm)} &= \pm\frac{m\kappa}{\tau}A(\tau), \quad T_{03}^{(\pm)} = \frac{m\pi\kappa}{\tau^2}, \\ A(\tau) &= |K_{i\kappa+1/2}(i\tau)|^2 - |K_{i\kappa-1/2}(i\tau)|^2. \end{aligned} \tag{51}$$

The eigenvalues (invariants) of the energy–momentum tensor,

$$\lambda_{1,2} = \frac{1}{2}(T_{33} - T_{00}) \pm \sqrt{\frac{1}{4}(T_{00} + T_{33})^2 - T_{03}^2}, \tag{52}$$

are real and have opposite signs in the Rindler space, while in the Milne space, they are complex conjugate for $\tau \ll 1$, when the momentum density (energy flux) is greater than half the sum of the energy density and the pressure:

$$\lambda_{1,2}(\tau) \approx \frac{R(\tau)}{\tau} \pm i\frac{\pi m\kappa}{\tau^2\cosh\pi\kappa} + \ldots, \quad \tau \ll 1. \tag{53}$$

As $\tau \longrightarrow 0$, $R(\tau)$ oscillates with a finite amplitude and an increasing frequency.

## 6. RIGHT AND LEFT SPINOR MODES

In the spinor case, the right and left superpositions of the positive- and negative-frequency modes are defined as in the scalar case, but the Dirac scalar product leads to different Bogoliubov coefficients,

$$\begin{aligned} \chi_{\kappa s}^R &= \alpha_\kappa \chi_{\kappa s}^{(+)} + \beta_\kappa \chi_{\kappa s}^{(-)}, \\ \chi_{\kappa s}^L &= \beta_\kappa \chi_{\kappa s}^{(+)} + \alpha_\kappa \chi_{\kappa s}^{(-)}, \\ \beta_\kappa &= i\alpha_\kappa e^{-\pi\kappa}, \quad \alpha_\kappa = \frac{e^{\pi\kappa/2}}{\sqrt{2\cosh\pi\kappa}}, \\ &|\alpha_\kappa|^2 + |\beta_\kappa|^2 = 1. \end{aligned} \tag{54}$$

Evidently, the right and left modes satisfy the orthogonality and normalization conditions

$$\int_S d\sigma_\alpha \bar{\chi}_{\kappa s}^a(x)\gamma^\alpha \chi_{\kappa' s'}^{\alpha'}(x) = \frac{2\pi^2}{m}\delta_{aa'}\delta_{ss'}\delta(\kappa - \kappa'), \tag{55}$$

where $a, a' \in L, R$ and $S$ is a spacelike surface as in (30) or (45).

The modes $\chi_{\kappa s}^{(\pm)}$ and $\chi_{\kappa s}^{L,R}$ form two complete sets of Dirac equation solutions and any other solution $\chi(x)$ can be decomposed into the corresponding integrals

$$\begin{aligned} \chi(x) &= \int_{-\infty}^{\infty} \frac{d\kappa}{2\pi^2}[a_{\kappa s}\chi_{\kappa s}^{(+)}(x) + b_{\kappa s}^* \chi_{\kappa s}^{(-)}(x)] \\ &= \int_{-\infty}^{\infty} \frac{d\kappa}{2\pi^2}[r_{\kappa s}\chi_{\kappa s}^R(x) + l_{\kappa s}^* \chi_{\kappa s}^L(x)], \end{aligned} \tag{56}$$

where summation over $s$ is assumed.

For example, for the positive-frequency plane wave solution with $s = 1$,

$$\begin{aligned} \chi_{p_1 1}^{(+)}(x) &= \frac{1}{\sqrt{2E_1}}\exp[i(p_1 z - E_1 t)]u_1^{(+)}(\theta_1), \\ \theta_1 &= \operatorname{arctanh}\frac{p_1}{E_1}, \end{aligned} \tag{57}$$

we have

$$a_{\kappa 1} = \frac{2\pi}{\sqrt{2E_1}}e^{i\kappa\theta_1}, \quad b_{\kappa 1}^* = 0, \tag{58}$$

$$r_{\kappa 1} = \frac{\pi\exp[\pi\kappa/2 + i\kappa\theta_1]}{\sqrt{E_1}\cosh\pi\kappa}, \qquad l^*_{\kappa 1} = \frac{-i\pi\exp[-\pi\kappa/2 + i\kappa\theta_1]}{\sqrt{E_1}\cosh\pi\kappa}. \tag{59}$$

For the spectra of the right and left modes, we then obtain (with $g_1 = 2\pi^2/E_1$)

$$|r_{\kappa 1}|^2 = \frac{g_1 e^{2\pi\kappa}}{e^{2\pi\kappa}+1}, \quad |l_{\kappa 1}|^2 = \frac{g_1}{e^{2\pi\kappa}+1}. \tag{60}$$

For the negative-frequency plane wave solution, the coefficients in expansions (56) are

$$a_{\kappa 1} = 0, \quad b^*_{\kappa 1} = \frac{2\pi}{\sqrt{2E_1}} e^{-i\kappa\theta_1}, \tag{61}$$

$$r_{\kappa 1} = \frac{-i\pi\exp[-\pi\kappa/2 - i\kappa\theta_1]}{\sqrt{E_1}\cosh\pi\kappa}, \qquad l^*_{\kappa 1} = \frac{\pi\exp[\pi\kappa/2 - i\kappa\theta_1]}{\sqrt{E_1}\cosh\pi\kappa}. \tag{62}$$

The spectra for the left and right modes then coincide with the respective expressions in (60).

Although these spectra resemble the thermal distribution of the Fermi-particle gas, this similarity seems to be artificial for the same reasons as in the scalar case. Moreover, decompositions (56) of the plane wave in the hyperbolic modes $\chi^{(\pm)}_{\kappa s}$ or $\chi^{R,L}_{\kappa s}$ and the inverse expansions of these modes in plane waves in Eqs. (41) and (54) confirm the completeness of these three sets and the absence of the loss of information or purity of states. We see that the hyperbolic symmetry and a definite frequency sign preserve the good analytical properties of the modes but lead to an indefinite sign of their charge density or energy density.

The "thermal" spectra appear when one preserves the hyperbolic symmetry of modes and requires the definiteness of the charge density or energy density signs in the entire Minkowski space. This can only be achieved at the expense of loosing good analytical properties of the modes and essentially consists in the transition from the boundary value of an analytical function on the cut to its jump on this cut. We have

$$\chi^{(+)}_{\kappa s} = \alpha^*_\kappa \chi^R_{\kappa s} + \beta^*_\kappa \chi^L_{\kappa s}, \quad \chi^{(-)}_{\kappa s} = \beta^*_\kappa \chi^R_{\kappa s} + \alpha^*_\kappa \chi^L_{\kappa s}, \tag{63}$$

$$|\alpha_\kappa|^2 = \frac{e^{2\pi\kappa}}{e^{2\pi\kappa}+1}, \quad |\beta_\kappa|^2 = \frac{1}{e^{2\pi\kappa}+1}, \tag{64}$$

where $|\alpha_\kappa|^2$ and $|\beta_\kappa|^2$ are the respective probabilities to find no pairs (one pair) and one pair (no pairs) in the state $\chi^{(+)}_{\kappa s}$, $\kappa > 0$ ($\kappa < 0$). This interpretation follows from the none-one-particle analysis of wave equation solutions and does not require the transition to the secondary quantization, although it is confirmed by it [5, 6].

For the Rindler observer with $\beta = t/z$, we have

$$T^R_{00}, T^R_{33}, T^R_{03} = \frac{4m\kappa\cosh\pi\kappa}{\zeta}\left(|K_{i\kappa-1/2}(\zeta)|^2, \int_\zeta^\infty \frac{d\zeta}{\zeta}|K_{i\kappa-1/2}|^2, 0\right) \tag{65}$$

and for Milne observer with $\beta = z/t$,

$$T^R_{00} = \frac{\pi m\kappa}{\tau^2}\left(1 + \frac{\pi\tau}{\cosh\pi\kappa}\int_0^\tau \frac{d\tau}{\tau}|J_{i\kappa+1/2}(\tau)|^2\right),$$

$$T^R_{33} = \frac{\pi m\kappa}{\tau^2}\left(1 - \frac{\pi\tau}{\cosh\pi\kappa}|J_{i\kappa+1/2}(\tau)|^2\right), \tag{66}$$

$$T^R_{03} = \frac{\pi m\kappa}{\tau^2}.$$

The energy density is greater than the pressure. As $\tau \longrightarrow 0$, we have

$$T^R_{00} \approx T^R_{33} \approx T^R_{03} = \pi m\kappa/\tau^2$$

similarly to the energy–momentum tensor of electromagnetic waves.

It is interesting to note that in the $R$ sector, the eigenvalue $\lambda^R_{1,2}$ tend to infinity as $\zeta \longrightarrow 0$, while in the $P$ or $F$ sectors, they are finite at $\tau = 0$,

$$\lambda^R_{1,2}\big|_{\tau\to 0} = -\frac{2\pi m\kappa}{1+4\kappa^2} \pm \frac{2\pi m\kappa}{1+4\kappa^2}\sqrt{\frac{1+4\kappa^2}{9+4\kappa^2}} + \ldots \tag{67}$$

The $\operatorname{sgn} t_{00}$ is relativistically invariant in only two cases:

1) the eigenvalues $\lambda_1$ and $\lambda_2$ are real and have opposite signs,

$$\lambda_1\lambda_2 = T^2_{03} - T_{00}T_{33} < 0, \tag{68}$$

2) the eigenvalues are real, have the same sign, and the energy density is greater than the pressure in magnitude:

$$(\lambda_1 - \lambda_2)^2 = (T_{00} + T_{33})^2 - 4T^2_{03} > 0, \qquad \lambda_1\lambda_2 > 0, \quad \operatorname{sgn}(T^2_{00} - T^2_{33}) > 0. \tag{69}$$

We note that $\operatorname{sgn}(t^2_{00} - t^2_{33})$ is relativistically invariant only if $\lambda_1$ and $\lambda_2$ are real, i.e., if $(\lambda_1 - \lambda_2)^2 > 0$. Then, if $\lambda_{1,2}$ are complex or if they are real and have the same sign, but $\operatorname{sgn}(T^2_{00} - T^2_{33}) < 0$, the $\operatorname{sgn} t_{00}$ can be changed by a Lorentz transformation.

The tensor $t^R_{\alpha\beta}$ possesses the first property in the $R$ sector and either the first or the second property depending on the value of $\tau$ in the $F$ and $P$ sectors.

In the $F$ and $P$ sectors, the eigenvalues $\lambda_1^R$ and $\lambda_2^R$ are real because of the inequality

$$\int_0^\tau \frac{d\tau}{\tau} |J_{i\kappa+1/2}(\tau)|^2 - |J_{i\kappa+1/2}(\tau)|^2 > 0. \qquad (70)$$

Inequalities (70) and (25) that are essential in this paper were not found in the mathematical literature.

## 7. CONCLUSION

Hyperbolic symmetry of scalar and spinor field states requires plane waves with unlimited frequencies to participate in the corresponding superpositions. For the scalar field, field states with the quantum number $\kappa$ that are formed as superpositions and are analytic in the coordinates $x_\pm = t \pm z$ do not possess an everywhere timelike current density, while for the spinor field, they do not possess the energy–momentum tensor with everywhere real eigenvalues. This means that these states describe both particles and antiparticles. Nevertheless, it is possible to construct hyperbolically symmetric right and left states that are not analytic in $x_\pm$ but possess an everywhere timelike current density and the energy–momentum tensor with everywhere real eigenvalues. Precisely these states describe the particle or the antiparticle.

This implies that the charge densities $j_\kappa^{R0}$ and $J_\kappa^{R0}$ for the scalar particle (antiparticle) states $\phi_\kappa^R$ and the energy densities $t_{\kappa 00}^R$ and $T_{\kappa 00}^R$ for the spinor particle (antiparticle) states $\chi_{\kappa s}^R$ are everywhere positive (negative) for $\kappa > 0$ ($\kappa < 0$) and are equal to zero in the $L$ sector. This assertion remains valid after replacing $R \rightleftarrows L$ and changing the sign of $\kappa$.

It is known [7] that if a wave packet is formed from plane waves and is localized in a region of the order of or less than the Compton wave length, it must contain both positive and negative frequencies. The superpositions $\phi_\kappa^{(+)}$ and $\chi_{\kappa s}^{(+)}$ do not contradict this assertion because each of them is localized in a region of the order of the Compton length only for $|t| \lesssim m^{-1}$, while for $|t| \gg m^{-1}$, each superposition consists of two waves that propagate along the light cone boundaries $z = \pm t$, exponentially decaying outside the cone for $\zeta = m\sqrt{z^2 - t^2} \gg 1$ and oscillating and falling off only as $\tau^{-1}$ inside the cone for $\tau = m\sqrt{t^2 - z^2} \gg 1$. Therefore, these two waves remain coherently connected in a single wave packet with the width $\approx 2|t|$.

In the well-known review [8], Pauli made the following remark about energy density in the Dirac electron field theory: "The concept of the energy density seems to be more problematic in this theory than that of the volume integrated total energy. The energy density is no longer positive definite for the theory of holes, in contradistinction to the case for the theories discussed in §§1 and 2. This is also shown in the $c$ number theory; even if limitation is made to wave packets in which the partial waves all have the same sign of the frequency in the phase $\exp i(\mathbf{k} \cdot \mathbf{x} - k_0 x_0)$ the energy density (as distinguished from the total energy) cannot be made positive definite." I do not know whether Pauli had some example of such a wave packet. In any case, each of the modes $\chi_{\kappa s}^{(\pm)}$ can serve as a specific illustration of his remark. The energy density for each of these modes can accept both signs near the light cone owing to singularities on the cone related to the hyperbolic symmetry of the modes. On the other hand, each of the modes $\chi_{\kappa s}^{R,L}$ is an example of such a superposition of positive- and negative-frequency spinor plane waves with a sign-definite energy density in the entire Minkowski space.

It is interesting that the scalar eigenfunctions of the Lorentz boost operator appear in the analysis of the photon wave function localized near the photon propagation plane in 3 + 1-space [9]. However, a scalar product different from (30) is used in this analysis.

## ACKNOWLEDGMENTS

I thank M.A. Soloviev for useful comments. The work was partly supported by the Russian Foundation for Basic Research (project nos. 00-15-96566 and 99-02-17916a).

## *APPENDIX*

The integral $J_{\kappa\kappa'}^R$ defined in [4] by Eq. (14), being the integral of a total differential, does not actually depend on the form of the spacelike surface over which it extends, but depends only on the parameters $mt$ and $\zeta$ fixing the coordinates of the left boundary of this surface. Namely, the $z$ coordinate of the left boundary is equal to $\sqrt{t^2 + \zeta^2/m^2}$, while the right boundary is at infinity. When the left boundary tends to zero at a fixed ratio $mt/\zeta$, we obtain the result (20) from [4] without any uncertainties related to the factor $\exp[i(\kappa - \kappa')\operatorname{arcsinh}(mt/\zeta)]$, which eventually turns into 1 at fixed $mt/\zeta$ and $\kappa = \kappa'$. Thus, the normalization integral (20) in [4] is correct for any spacelike surface lying in the $R$ sector with the left boundary at zero—not at $z = |t|$ as was assumed in [4].

Similarly, expression (28) for the normalization integral $J_{\kappa\kappa'}^L$ in [4] is correct for any spacelike surface lying in the $L$ sector with the right boundary at zero, rather than at $z = -|t|$ as was assumed in [4].

The integral $J_{\kappa\kappa'}^F$ defined by Eqs. (22) and (23) in [4] is justified for any spacelike surface lying inside the $F$

sector with the boundaries at the points defined by fixed values of $mt$ and $\tau = m\sqrt{t^2 - z^2}$. The $z$ coordinates of the left and right boundaries of this surface are then given by $z_{1,2} = \mp\sqrt{t^2 - \tau^2/m^2}$. As $t$ tends to infinity at fixed $\tau$, we obtain the result (25) from [4] without any ambiguity related to the factor inside the parentheses in Eq. (23) in [4], which turns into $\pi$ at fixed $\tau$ and $\kappa = \kappa'$. Thus, normalization integral (25) in [4] is correct for any spacelike integration surface lying in the $F$ sector and having the boundaries at $z_{1,2} = \mp\infty$ but not at $z_{1,2} = \mp|t|$, as was understood in [4]. A similar comment applies to the integral $J^P_{\kappa\kappa'}$.

On any spacelike surface entirely lying in the $P$, $L + R$ or $F$ sectors with the left and right boundaries at infinities, each of the states $\phi^{(\pm)}_{\kappa}$ has the same conserved total charge

$$Q^{(\pm)}_{\kappa} = Q^{(\pm)}_{\kappa P} = Q^{(\pm)}_{\kappa L} + Q^{(\pm)}_{\kappa R} = Q^{(\pm)}_{\kappa F} \gtrless 0. \qquad (71)$$

Therefore, the factor 1/2 in the right-hand sides of Eqs. (34) and (35) in [4] must be replaced by 1.

## REFERENCES

1. W. G. Unruh, Phys. Rev. D **14**, 870 (1976).
2. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **94** (7), 31 (1988) [Sov. Phys. JETP **67**, 1313 (1988)].
3. N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge Univ. Press, Cambridge, 1982).
4. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **114**, 777 (1998) [JETP **87**, 421 (1998)].
5. A. I. Nikishov, Nucl. Phys. B **21**, 346 (1970).
6. A. I. Nikishov, Tr. Fiz. Inst. Akad. Nauk SSSR **168**, 156 (1986); *Issues in Intense-Field Quantum Electrodynamics*, Proc. Lebedev Phys. Inst., Vol. 168. Ed. by V. L. Ginzburg (Nova Science, Commack, 1987).
7. J. D. Bjorken and S. D. Drell, *Relativistic Quantum Mechanics* (McGraw-Hill, New York, 1964; Nauka, Moscow, 1978).
8. W. Pauli, Rev. Mod. Phys. **13**, 203 (1941).
9. I. Bialynicki-Birula, Prog. Opt. **36**, 245 (1996).

Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 221–235.
Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 252–268.
Original Russian Text Copyright © 2001 by Zabolotskii.

# Dynamics of a Light Field in a Composite Integrable Model

**A. A. Zabolotskii**
*Institute of Automatics and Electrometry, Siberian Division, Russian Academy of Sciences,*
*Universitetskiĭ pr. 1, Novosibirsk, 630090 Russia*
*e-mail: zabolotskii@iae.nsk.su*

Received May 22, 2000

**Abstract**—The inverse scattering transform method is used to solve the model that describes the evolution of light pulses in an optical system that includes a set of media with different nonlinear optical properties. As a physical example, we analyze a model composed of the systems of equations that describe the resonant interaction of a very short light pulse with an energy transition of the medium and the ensuing propagation of the light field in an optical fiber. The constant boundary value of one of the fields is shown to result in an asymptotic quasi-radiative solution of the model. *© 2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

An analysis of soliton generation in nonlinear optics, including that in terms of completely integrable models [1], has been the subject of many theoretical papers (see, e.g., [2, 3]). To faithfully describe an experimental situation often requires solving an initial value–boundary value problem for systems of nonlinear evolution equations. In practice, this problem can currently be solved only in terms of completely integrable models. Therefore, finding and solving completely integrable models is of both theoretical and practical interest.

In this paper, we consider an integrable model composed of two or more models that describe the evolution of fields on nonoverlapping, joined intervals. We call such models composite to distinguish them from models combining integrable models on coincident intervals. The latter include, for example, a combination of the integrable system of Maxwell–Bloch equations for a two-level medium and the integrable nonlinear Schrödinger equation (NSE) for an infinite medium [3].

As far as we know, composite models of this kind and peculiarities of the pulse generation in them have not yet been studied. At the same time, the existence of such models is determined by actual physical situations. The experimental facilities used to generate very short optical pulses usually include different nonlinear and linear media. In some media, a pulse is generated; in other media, it is amplified; in still other media, the pulse is compressed and takes a shape convenient for subsequent applications, etc. An example of such a system can be a two-level laser amplifier supplemented with a nonlinear medium in the form of an optical fiber. Studies of composite models and the associated initial value–boundary value problems can reveal qualitatively new generation regimes under actual physical conditions.

A modified nonlinear Schrödinger equation (MNSE) with differential nonlinearity [4] is commonly used to analyze the propagation of very short light pulses in optical fibers. Stable soliton-like field pulses in such a medium result from the balance between dispersion and nonlinearity. In most studies of soliton effects in optical fibers, additional nonlinear interactions were taken into account as perturbations (see, e.g., [5]). At the same time, some combinations of nonlinear effects can be described in terms of integrable models that combine simpler integrable models in an infinite medium [3].

The example of a composite model analyzed here describes the propagation of very short pulses of a light (main) field in an optical fiber with allowance for the interaction with additional fields in a nonlinear resonant medium of finite length without invoking the perturbation theory. In an actual experiment, such a scheme can be used to study the generation of very short pulses in a resonant medium, which are subsequently injected into an optical fiber that serves to transmit information by means of these pulses. The model includes the MNSE and the system of reduced Maxwell equations that describes the resonant interaction of three light wave packets with a two-level medium. This integrable model is associated with a spectral problem related to the Wadati–Konno–Ichikawa (WKI) problem [6].

In general, to describe the generation of solitons and other wave packets in nonlinear optical problems in finite or semi-infinite media, an initial value–boundary value problem with nonzero fields at the boundaries should be solved. Such boundary conditions often correspond to more realistic and more easily realizable experimental conditions than the physical conditions for which a soliton “classical version” of the inverse scattering transform method (ISTM) was constructed [1]. Accordingly, determining the type and form of the solution associated with boundary conditions of this kind is important for applications. These boundary con-

1063-7761/01/9302-0221$21.00 © 2001 MAIK “Nauka/Interperiodica”

ditions can qualitatively change the generation conditions for optical solitons, breathers, and other types of solutions. At the same time, as was noted above, such a problem could be solved in practice only in terms of integrable models. In this case, solving the initial value–boundary value problem generally runs into serious mathematical difficulties. However, papers have recently appeared in which methods were developed for solving such problems for a number of models integrable by using the ISTM based on the solution of the Riemann–Hilbert problem [1]. Noteworthy are [7, 8], in which boundary value problems for the sine-Gordon equation, the NSE, and others were investigated. In particular, Kiselev [9] suggested a formal method of solving the Goursat problem for the system of Maxwell–Bloch equations for a two-level medium. Previously, the initial value–boundary value problem for the same system was considered by Gabitov *et al.* [10], who analyzed self-similar asymptotics. Fokas [11] proposed a new approach to using the Riemann–Hilbert problem to solve the initial value–boundary value problem on a finite interval. In [8], the Raman scattering model, which is related to the Maxwell–Bloch model for a single-frequency transition, with averaging over frequency mixing was solved in terms of the ISTM for a finite interval. The form and properties of the asymptotic solution generated by simple, but nontrivial boundary conditions have been discussed in several recent papers [12–14]. Isolated soliton-like solutions with variable parameters were found in [12–14], in which the Raman scattering model was solved for nonzero boundary values of the Stokes field and pumping. In [15], we proved for the same model that as the effective length of a nonlinear medium increased, the total contribution of these poles led to a quasi-self-similar, nonsoliton solution. The result proven theoretically for slowly changing boundary conditions is confirmed by numerical simulations [15].

In this paper, we use the version of ISTM based on the solution of the Riemann–Hilbert problem associated with the WKI spectral problem to solve the new integrable model. For simple, but nontrivial initial value–boundary value problems, we theoretically analyze the generation and dynamics of very short light pulses. The integrable model studied here also differs significantly from those considered, for example, in the review [3] in a nontrivial dependence of the scattering coefficient on one of the variables. In particular, this dependence can give rise to an infinite number of moving poles in the complex plane associated with soliton-like solutions.

In our physical example of two different nonlinear media, the boundary conditions for the second medium are determined by the field evolution in the first medium. This initial value–boundary value problem is generally unsolvable. However, an asymptotic solution can be found in terms of the new integrable model considered here. In this paper, we develop an approach that allows the field evolution in composite models to be asymptotically described for this kind of boundary conditions. It has been proven that for a sufficiently large effective length of the resonant medium, the solution reduces to a system of integral equations that asymptotically coincide with the equations describing the quasi-radiative solution of MNSE. This solution is determined only by the continuum of the problem ($\mathrm{Im}\lambda^2 = 0$) and can be expressed in terms of one of the Painleve transcendents [1]. Explicit asymptotics of the quasi-radiative solution for the MNSE were found by Kitaev and Vartanian [16]. Our results also provide the answer to the question of the form of the asymptotic solution for the problem of light interaction with a two-level medium under conditions of degenerate two-frequency resonance and weak excitation of the medium.

The paper was structured as follows. The general structure of composite integrable models is described in Section 2. The physical situations that lead to the composite model analyzed below are considered in Section 3. The method of model solution based on the Riemann–Hilbert problem is described in Section 4, and the dependence of scattering data on variable is derived in Section 5. In Section 6, we show, for simple initial value–boundary value problems, that the approximate quasi-radiative solution of MNSE describing the pulse dynamics in an optical fiber can be generated by the boundary conditions for an additional field in the resonant medium. The results and their possible generalization are discussed in the final section.

## 2. COMPOSITE INTEGRABLE MODELS

Let us describe the structure of an integrable model composed of $N \times M$ integrable models (some of them may coincide). Let the composite integrable model be represented as a compatibility condition for the following linear systems of equations:

$$\frac{\partial}{\partial \tau}\psi(\tau, z; \lambda) = \sum_{j=1}^{M} \beta_{j,\,j+1}(\tau) L_j(\tau, z; \lambda)\psi(\tau, z; \lambda) \equiv \mathcal{L}\psi, \tag{1}$$

$$\frac{\partial}{\partial z}\psi(\tau, z; \lambda) = \sum_{i=1}^{N} \alpha_{i,\,i+1}(z) A_i(\tau, z; \lambda)\psi(\tau, z; \lambda) \equiv \mathcal{A}\psi, \tag{2}$$

$$\beta_{j,\,j+1}(\tau) = [\theta(\tau - \tau_j)\theta(-\tau + \tau_{j+1})]\tilde{\beta}_j(\tau), \quad \tau_{j+1} > \tau_j,$$

$$\alpha_{i,\,i+1}(z) = [\theta(z - z_i)\theta(-z + z_{i+1})]\tilde{\alpha}_i(z), \quad z_{i+1} > z_i.$$

Here, $\theta(z)$ is the step function:

$$\theta(z) = \begin{cases} 0, & z < 0 \\ 1, & z \ge 0; \end{cases}$$

$\tilde{\beta}_i(\tau)$ and $\tilde{\alpha}_i(z)$ are smooth functions that do not become zero and infinite; $\alpha_{i,\,i+1}(z)/\tilde{\alpha}_i(z)$ and $\beta_{i,\,i+1}(\tau)/\tilde{\beta}_i(\tau)$ are the projectors:

$$\alpha^2_{i,\,i+1}(z)/\tilde{\alpha}^2_i(z) = \alpha_{i,\,i+1}(z)/\tilde{\alpha}_i(z)$$

etc.; $L_j$, $A_i$, $\mathscr{L}$, and $\mathscr{A}$ are matrix operators.

The compatibility condition for these linear systems is

$$\begin{gathered}\frac{\partial}{\partial z}\sum_{j=1}^{M}\beta_{j,\,j+1}(\tau)L_j - \frac{\partial}{\partial \tau}\sum_{i=1}^{N}\alpha_{i,\,i+1}(z)A_i \\ + \left[\sum_{j=1}^{M}\beta_{j,\,j+1}(\tau)L_j, \sum_{i=1}^{N}\alpha_{i,\,i+1}(z)A_i\right] = 0.\end{gathered} \tag{3}$$

Multiplying (3) by $a_{i,\,i+1}\beta_{j,\,j+1}\tilde{\alpha}_i^{-1}\tilde{\beta}_j^{-1}$ yields

$$\begin{gathered}\tilde{\beta}_j(\tau)\frac{\partial}{\partial z}L_j - \tilde{\alpha}_i(z)\frac{\partial}{\partial \tau}A_i + [\tilde{\beta}_j(\tau)L_j, \tilde{\alpha}_i(z)A_i] = 0, \\ z \in [z_i, z_{i+1}], \quad \tau \in [\tau_j, \tau_{j+1}].\end{gathered} \tag{4}$$

Thus, the evolution in the square $[z_i, z_{i+1}]$, $[\tau_j, \tau_{j+1}]$ is described by the system of equations with the Lax representation as linear systems:

$$\partial_\tau \psi = \tilde{\beta}_j L_j \psi, \quad \partial_z \psi = \tilde{\alpha}_i A_i \psi.$$

We do not known the ISTM apparatus for the spectral problem (1) in the case of several intervals $[\tau_j, \tau_{j+1}]$. In this paper, we study an example of a composite model with the Lax representation (1) and (2) for one semi-infinite interval $[\tau_1 = 0, \tau_2 = \infty)$ and two different media on intervals $[z_1 = 0, z_2]$ and $[z_2, z_3 = \infty]$. The problem for a finite or semi-infinite interval $[\tau_1, \tau_2]$ corresponding to a nonzero projector $\beta_{12}(\tau)$ was solved, as noted in the Introduction, for simple, but nontrivial initial-boundary conditions for one interval $[z_1, z_2]$(see, e.g., [8, 11]).

Note that here, the composite model including different models integrable on different intervals is assumed to be integrable in advance. This assumption generally imposes additional constraints on the physical parameters.

## 3. A PHYSICAL EXAMPLE OF A COMPOSITE MODEL

### *3.1. Statement of the Problem*

Let us consider a composite model that includes a model for the resonant interaction of the main field with an additional field in a two-level medium on interval $[z_0 = 0, z_1)$ and a model describing the evolution of the main field in a Kerr medium on a semi-infinite interval $z \geq z_1$. The field interaction in the resonant medium is assumed to generate pulses, which subsequently propagate through the Kerr medium. The initial-boundary conditions in the simplest case are as follows: the amplitude of the initial "additional" field in the resonant medium is nonzero, and the initial main field is zero. The pulse generation in this problem can be triggered by an arbitrarily small seed of the main field at the $z = 0$ boundary.

### *3.2. A Model of the Two-Wave Interaction in a Resonant Medium*

Let us first consider examples of the resonant field interaction with a two-level medium and the physical conditions that lead to the following system of evolution equations:

$$\partial_\tau q_1 = -2\bar{q}_1 q_2 + i\nu q_1, \tag{5}$$

$$\partial_z q_2 = q_1^2 - i2g^2 q_2|q_1|^2 \tag{6}$$

(the bar denotes a complex conjugate) and then to the system composed of model (5), (6), and MNSE.

### *3.3. Light Propagation under Conditions of a Degenerate Two-Photon Interaction*

The coherent interaction of a very short pulse with a two-photon-absorbing medium was considered in several papers, beginning with [17] (see also the review [3]). The effect of self-induced transparency was analyzed in terms of the system of Maxwell–Bloch equations with allowance for the change in level populations and for the nonlinear Stark effect. However, a significant (of the order of unity) change in level populations requires intense light fields. At the same time, for applications, for example, in microelectronics, it is important to find conditions for the existence of self-induced transparency for relatively weak fields that do not destroy the optical medium.

In this section, we derive the self-consistent integrable model of self-induced transparency that corresponds to weak excitation of the medium. The existence of multisoliton solutions for this model proves the possibility of observing self-induced transparency in this limiting case for much weaker (by orders of magnitude) fields than in the model noted above [17].

The general system of equations for the field and a two-level medium (with levels 1 and 2) under conditions of degenerate two-photon resonance was derived from the Schrödinger and wave equations in [17, 18]. The polarizability dynamics of the medium is described by the equation

$$\frac{\partial}{\partial t}R = -i\left(\frac{r_{22} - r_{11}}{4\hbar}\right)|E|^2 R - \frac{r^2}{2\hbar}E^2 R_3. \tag{7}$$

Here, $R_3$ is the difference between the upper- and lower-level populations, $R = iR_1 + R_2$ ($R_1$ and $R_2$ are the reactive and active polarization components of the medium), and $E$ is the field amplitude. The Bloch vector components are normalized to unity; i.e., the following equality holds:

$$R_3^2 + |R|^2 = 1. \tag{8}$$

Twice the field carrier frequency $\omega$ is assumed to be equal to the transition frequency $\omega_{12}$ of the medium. The standard derivation of the equations for a two-frequency interaction with a two-level medium involves an adiabatic elimination of the intermediate-level occupation probability amplitude (see, e.g., [17, 19]). This procedure leads to the following form of the coefficients:

$$r = \frac{2}{\hbar}\sum_{n>3}\frac{q_{2n}q_{n1}(\omega_{n2}+\omega_{n1})}{(\omega_{n2}+\omega)^2},$$

$$r_{mm} = \frac{1}{\hbar}\sum_{n>3}\frac{\omega_{nm}|q_{nm}|^2}{\omega_{nm}^2-\omega^2}, \quad m = 1, 2.$$

Here, $q_{mn}$ is the dipole moment between transition levels $m = 1, 2$ and intermediate levels $n$, and $\omega_{nm}$ is the corresponding transition frequency.

The Maxwell equations for a slowly changing envelope of field $E$ reduce to the equation [17]

$$\begin{gathered}\left(\frac{\partial}{\partial x}+\frac{1}{c}\frac{\partial}{\partial t}\right)E = -\frac{2\pi\omega N_0}{c}\\ \times\left[i\left(\frac{r_{22}-r_{11}}{4\hbar}\right)(R_3-R_3^{(0)})E+\bar{E}R\right],\end{gathered} \tag{9}$$

where $R_3^{(0)}$ is the initial difference between the level populations, and $c$ is the speed of light in the medium. In many known experiments on the observation of two-frequency coherent effects, the upper-level ($m = 2$) population during the field interaction with the two-level transition was found to remain low compared to the ground-state population. The ratio of the upper- and lower-level populations often does not exceed $10^{-5}$–$10^{-7}$ during the entire interaction [20]. In this case, i.e., in the limit of weak excitation of the medium, the following expansion holds:

$$R_3 \approx -1 + |R|^2/2 + \mathbb{O}(|R|^4) \tag{10}$$

(here, $R_3^{(0)} = -1$). To derive Eqs. (5) and (6), we use expansion (10) and disregard all terms of the fourth order or higher in field amplitude and polarization ($\sim|R|^2E^2$) in Eqs. (7) and (9). Note that including terms of the third order ($\sim|R|^2E$) results only in a renormalization of the constant $g^2$ in (5) and (6). Let us now change to variables $\tau$ and $z$:

$$\frac{\partial}{\partial\tau} = \frac{c}{\pi\omega N_0}\frac{\partial}{\partial x}, \quad \frac{\partial}{\partial z} = \frac{2\hbar}{r^{-2}}\frac{\partial}{\partial(t-xc^{-1})}.$$

System (7), (9) then takes the form (5), (6) after allowing for expansion (10), discarding terms higher than the third order, and the following substitutions:

$$E = q_1\exp\left(-ig_1\int_0^\tau|R|^2d\tau - i\nu\tau\right),$$

$$R = q_2\exp\left(-2ig_1\int_0^\tau|R|^2d\tau - 2i\nu\tau\right),$$

where

$$g_1 = \frac{r_{22}-r_{11}}{2\hbar}, \quad 2g^2 = \frac{r_{22}-r_{11}}{2r^2}+g_1.$$

When deriving (5) and (6), we used the equality

$$\frac{\partial|E|^2}{\partial\tau} = -2\frac{\partial}{\partial\tau}|R|^2, \tag{11}$$

which follows from (7) and (9).

### *3.4. Three-Wave Mixing in the Resonant Medium*

Let us describe yet another physical scheme for the interaction, which also leads to the system of equations (5) and (6). Let a three-frequency field be propagated through the medium:

$$\tilde{E}(x,t) = \sum_{j=1}^{3}\{P_j\exp[i(q_jx-\omega_jt)]+\text{c.c.}\}. \tag{12}$$

Here, $P_j$ are the slowly changing envelopes, $\omega_j$ are the carrier frequencies, and $q_j$ are the carrier wave vectors. The following conditions of two-photon resonance with the proper energy transition of the medium with frequency $\omega_0$ are assumed to be satisfied:

$$\omega_1+\delta_1\omega_2 = \omega_0+\nu_1, \quad \omega_3+\delta_2\omega_1 = \omega_0+\nu_2. \tag{13}$$

Here, $\delta_1 = \pm1$, $\delta_2 = \pm1$, and the detunings $\nu_k$ satisfy the condition

$$\nu_k \ll \omega_i, \quad k = 1, 2, \quad i = 0, \ldots, 3.$$

The resonance conditions (13) not only allow one to enhance significantly (by orders of magnitude) the effect of nonlinear mixing, but also to eliminate the terms describing the cubic (in amplitude) field self-action from the equations, i.e., the terms of type $P_j|P_j|^2$. In that case, these terms breaks the model integrability. As a result, the model includes only the cubic terms that correspond to the two-photon-induced Kerr nonlinearity (see, e.g., [21]).

Thus, assuming that the polarization of the medium follows the field variations and using the standard procedure of adiabatic elimination [2], we obtain the Maxwell equations in which the nonlinear field mixing is described by the two-photon-induced Kerr nonlinearity. The reduced Maxwell equations for the resonance conditions (13) at $\delta_1 = \delta_2 = -1$ are

$$\left(\frac{\partial}{\partial x} + \frac{1}{v_1}\frac{\partial}{\partial t}\right)P_1 = i\frac{4\pi\omega_1}{c^2}[\alpha_{22}P_1|P_2|^2$$

$$+ \alpha_{32}P_1|P_3|^2 + 2\mathrm{Im}\,\alpha_{23}P_2\bar{P}_1P_3\exp(-i\Delta x)],$$

$$\left(\frac{\partial}{\partial x} + \frac{1}{v_2}\frac{\partial}{\partial t}\right)P_2 = i\frac{4\pi\omega_2}{c^2} \tag{14}$$

$$\times[\alpha_{22}P_2|P_1|^2 + \alpha_{23}P_1^2\bar{P}_3\exp(i\Delta x)],$$

$$\left(\frac{\partial}{\partial x} + \frac{1}{v_3}\frac{\partial}{\partial t}\right)P_3 = i\frac{4\pi\omega_3}{c^2}$$

$$\times[\alpha_{33}P_3|P_1|^2 + \alpha_{32}\bar{P}_2P_1^2\exp(i\Delta z)],$$

where $\Delta = 2q_1 - q_2 - q_3$ and $v_i$ are the group velocities of the fields with envelopes $P_i$. The Maxwell equations corresponding to this interaction scheme and the coefficients $a_{23}$, $a_{32}$, $a_{22}$, and $a_{33}$ are given in [2, Section 10.4]. However, in contrast to the equations in [2], system (14) includes a time dependence of the field amplitudes and uses the approximation of slow envelopes. We also assume that the time scales of amplitude variations are much shorter than the time $\Gamma^{-1}$, where $\Gamma$ is the minimum relaxation constant. Similar equations emerge for other $\delta_i$, which we do not provide here to save space.

In several cases, when waves were mixed in a medium with Kerr nonlinearity, the constancy condition for one of the fields was satisfied with good accuracy [20]. This condition also significantly simplifies the synchronization of field pulses, which is required to observe the effect. The constancy condition for one of the fields is satisfied, for example, when $|P_2| \gg |P_3|$ or $|P_2| \gg |P_3|$. Let us choose the latter case where the variations in field $P_3$ may be ignored. In this case, only the first two equations remain in system (14). Let $P_3(x, t) \equiv A$, where $A$ is a real constant (for a complex $A = A_0e^{i\phi_0}$, the phase factor $e^{i\phi_0}$ is removed by shifting the phase of field $P_2$). In this approximation, the first two equations of system (14) are reduced to system (5), (6) by the following substitutions:

$$P_1 = q_1\exp\left[-i\frac{\Delta}{2}x + ig_1\int_0^\tau|q_2|^2d\tau\right],$$

$$P_2 = iq_2\exp\left[ig_1\int_0^\tau|q_2|^2d\tau\right],$$

$$g^2 = -\frac{1}{2}\left(\frac{\alpha_{ss}}{\alpha_{sa}A} + \frac{1}{2}g_1\right), \quad g_1 = \frac{\alpha_{ss}}{\alpha_{pa}A},$$

$$\nu = \frac{c^2\Delta}{8\pi\omega_1\alpha_{pa}} + \frac{\alpha_{as}}{\alpha_{pa}}|A|^2,$$

$$\frac{\partial}{\partial\tau} = \frac{c^2}{4\pi\alpha_{pa}\omega_1}\left(\frac{\partial}{\partial x} + \frac{1}{v_1}\frac{\partial}{\partial t}\right),$$

$$\frac{\partial}{\partial z} = \frac{c^2}{4\pi\alpha_{sa}\omega_2}\left(\frac{\partial}{\partial x} + \frac{1}{v_2}\frac{\partial}{\partial t}\right).$$

The integrable system of equations (5) and (6) is given for the first time. This system, as well as a similar integrable system of equations built previously [22], is associated with a WKI-type spectral problem. However, applying the ISTM apparatus to system (5), (6) results in new singularities related to peculiarities of the boundary conditions for this model. The practical significance of model (5), (6) may stem from the fact that it describes the degenerate two-frequency interaction of light with a medium for minimum field intensities and simultaneously has nonsingular soliton solutions. This distinguishes it from the standard model of second-harmonic generation, which is formally equivalent to (5), (6) at $g = 0$. In the latter case, this system of equations is similar to the model of explosive instability [1] without stable soliton solutions. Some self-similar solutions for the system of equations (5) and (6) can coincide with the solutions of the second-harmonic generation model or the model of a two-wave interaction in media with quadratic nonlinearity with allowance for cubic nonlinearity. Accordingly, analyzing system (5), (6) is of interest in its own right, but here, it will be analyzed as an element of the composite model.

### *3.5. The Composite Model*

In actual experimental facilities, it is not uncommon for light pulses to pass through different optical media located sequentially. Consider a situation when a very short pulse first passes through a resonant medium on interval $[z_1, z_2]$ and then propagates in an optical fiber on interval $[z_2, \infty)$.

In the resonant medium, the main field interacts with two additional fields. Let the light pulse $\mathcal{E}$ (main field) on interval $[z_1 = 0, z_2]$ interact with two additional fields with envelopes $\mathcal{G}$ and $\mathcal{U}$ according to the scheme described in the preceding subsection. In the notation of this model,

$$\mathcal{E} \equiv P_2, \quad \mathcal{G} \equiv P_3, \quad \mathcal{U} \equiv P_1.$$

We assume that the change in field $\mathcal{G}$ during the interaction may be disregarded. Let us describe the mechanism for energy conversion of the additional fields into pulses of the main field and determine the form of the corresponding solution. To this end, we construct a composite model built from model (5), (6) on interval $[0, z_2]$, $\alpha_{1,2} \neq 0$, and the MNSE on interval $[z_2, \infty)$, $\alpha_{2,3} \neq 0$ ($\alpha_{i,i+1}$ are defined in Section 2).

The MNSE includes differential cubic nonlinearity [see the left-hand side of Eq. (20) below], which produces a nonlinear phase modulation of the pulse and which should be taken into account when describing the evolution of sufficiently short light pulses in optical fibers [23]. A linearly polarized optical wave obeys the Maxwell equation in a one-dimensional medium:

$$\frac{\partial^2 E}{\partial x^2} - \frac{1}{c^2}\frac{\partial^2 D}{\partial t^2} = \frac{2\pi n_2 n_0}{c^2}\frac{\partial^2}{\partial t^2}(|E|^2 E). \qquad (15)$$

Here, $E$ is the electric field and $D = E + 4\pi P_L$ is the field in the medium. The refractive index is

$$n(\omega, E) = n(\omega) + n_2|E|^2, \quad n_0 = n(\omega_0), \qquad (16)$$

where $\omega_0$ is the carrier frequency of the field with a slow envelope $\mathcal{E}$:

$$E(x, t) = \mathcal{E}\exp\left[i\omega_0\left(\frac{n_0}{c}x - t\right)\right]. \qquad (17)$$

Substituting (17) in (15), passing to the frame of reference of the pulse, and retaining terms with the first derivative with respect to nonlinearity, we obtain the modified nonlinear Schrödinger equation

$$\begin{gathered} i\left(\frac{\partial}{\partial x} + \frac{n_0\partial}{c\partial t}\right)\mathcal{E} - \frac{1}{2}k''\frac{\partial^2\mathcal{E}}{\partial\tau^2} + \frac{\pi n_0\omega_0}{c}|\mathcal{E}|^2\mathcal{E} \\ + i\frac{2\pi n_2}{c}\frac{\partial}{\partial\tau}(|\mathcal{E}|^2\mathcal{E}) = 0 \end{gathered} \qquad (18)$$

(the prime denotes a derivative with respect to $\omega$ at point $\omega_0$).

For very short pulses, the dispersion term $ik'''E_{\tau\tau\tau}/6$ [24] must generally be added to the left-hand side of (18). Marcuse [25] found the following empirical ratio of the coefficients for single-mode optical fibers, which holds near the minimum of $|k''|$:

$$-\frac{k'''}{3|k''|} \approx 7 \times 10^{-4}\frac{\lambda[\mu\mathrm{m}] - 0.86}{1.27 - \lambda[\mu\mathrm{m}]}. \qquad (19)$$

It follows from (19) that the third-order dispersion may be ignored near $\lambda[\mu\mathrm{m}] \approx 0.86$. At the same time, according to (19), this dispersion must be taken into account near $\lambda_0 = 1.27$, because $k'' = 0$ at this wavelength [26]. For $\lambda$ far enough from $\lambda_0$, the nonlinear waves, the solutions of Eq. (18), result from the balance between cubic differential nonlinearity, cubic nonlinearity, and quadratic dispersion. These robust nonlinear modes (solitons and other self-similar solutions) are stable against small perturbations. Under these conditions, the cubic dispersion ($\sim ik'''E_{ttt}$), according to (19), may be treated as a small perturbation whose allowance causes the nonlinear-mode parameters to change only slightly. Therefore, the cubic dispersion was disregarded here.

It can be shown that, given the interaction with additional fields and in the approximations used, the system of Maxwell equations for $\mathcal{E}$, the field envelope of the pulse propagating in an optical fiber, reduces to the system

$$\begin{gathered} \partial_z\mathcal{E} + \theta(z - z_2)[iD\partial_{tt}^2\mathcal{E} + \tilde{\alpha}|\mathcal{E}|^2\mathcal{E} + \tilde{\beta}\partial_t(|\mathcal{E}|^2\mathcal{E})] \\ = \theta(-z + z_2)(-2\tilde{\gamma}_1\mathcal{U}^2\mathcal{G} + 2i\tilde{g}_1^2\mathcal{E}|\mathcal{U}|^2), \\ z \geq 0, \end{gathered} \qquad (20)$$

$$\begin{gathered} \theta(-z + z_2) \\ \times(\partial_t\mathcal{U} - \tilde{\gamma}_2\mathcal{E}\overline{\mathcal{U}} + 2i\tilde{g}_2^2\mathcal{U}|\mathcal{E}|^2 + i\tilde{\nu}\mathcal{U}) = 0, \\ z \geq 0, \end{gathered} \qquad (21)$$

where $D$, $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\gamma}_{1,2}$, and $\tilde{g}_{1,2}$ are real coefficients.

Since the interaction of field $\mathcal{E}$ with field $\mathcal{U}$ in the resonant medium and the self-action in the Kerr medium are spatially separated, the coefficients on the right-hand side of Eq. (20) can be arbitrary. In general, the interaction in a resonant medium is characterized by nonlinearities with coefficients larger than those in optical fibers by several orders of magnitude. At the same time, the nonlinear interaction and dispersion effectively manifest themselves for actual fibers whose lengths are several orders of magnitude larger than the length of the resonant medium. Therefore, to justify the physical applicability of system (20), (21), it will suffice to verify the presence of terms on the right- and left-hand sides of Eq. (20) separately.

The differential cubic nonlinearity, just as the second term on the right-hand side of Eq. (20), produces a nonlinear phase modulation, in contrast to the remaining terms of the equation. As was shown above, this nonlinear frequency modulation is resonant in nature and, therefore, determines the dominant cubic term on the right-hand side of Eq. (20). We ignore any change in $\mathcal{G}$.

Next, we assume the detuning $\tilde{\nu}$ of field $\mathcal{U}$ to be large enough to disregard the nonlinear frequency shift, i.e.,

$$|\tilde{g}_2^2\mathcal{U}||\mathcal{E}|^2 \ll |\tilde{\nu}\mathcal{U}|. \qquad (22)$$

Let us change from variables $\mathcal{E}$, $\mathcal{U}$, $t$, and $x$ to variables $Q$, $U$, $\tau$, and $z$:

$$\begin{gathered} \mathcal{E} = \tilde{\rho}Q\exp(-2i\tilde{\nu}t), \quad \mathcal{U} = \tilde{\delta}U\exp(-\tilde{\nu}t), \\ t - x/V_0 = \tilde{\kappa}\tau, \quad z = \zeta_0 x, \end{gathered} \qquad (23)$$

where

$$\tilde{\rho} = \sqrt{\frac{2\zeta_0}{\tilde{\alpha} + 2\tilde{\nu}}}, \quad \tilde{\kappa} = \sqrt{\frac{D}{\zeta_0}},$$

$$\tilde{\delta}^2 = \frac{\sqrt{2\zeta_0}}{\tilde{\gamma}_1 G\sqrt{\tilde{\alpha} + 2\tilde{\nu}}},$$

$$\frac{1}{v_0} = \frac{4\tilde{\beta}D\tilde{\nu}^2}{\tilde{\alpha}+2\tilde{\nu}} - D\tilde{\nu}, \quad \zeta_0 = \frac{\tilde{g}_1^2\sqrt{\tilde{\alpha}+2\tilde{\nu}}}{\tilde{\beta}\sqrt{2D}\gamma_1 G}.$$

The final system (20), (21) can then be written as

$$\begin{aligned}&\partial_z Q - \alpha_{2,3}\{\partial_{\tau\tau}^2 Q - 2i(|Q|^2 - |Q^{\pm}|^2)Q \\ &\quad + g^2\partial_\tau[(|Q|^2 - |Q^{\pm}|^2)Q]\} \\ &= \alpha_{1,2}(2U^2 + 2ig^2Q|U|^2), \quad z \ge 0,\end{aligned} \tag{24}$$

$$\alpha_{1,2}(\partial_\tau U + hQ\bar{U}) = 0, \quad z \ge 0, \tag{25}$$

where

$$\begin{gathered}h = \frac{\tilde{\gamma}_2 G\sqrt{2D}}{\sqrt{\tilde{\alpha}+2\tilde{\nu}}}, \quad |Q^{\pm}|^2 = \frac{2D\tilde{\nu}^2}{\zeta_0}, \\ g^2 = \frac{2\tilde{\beta}\sqrt{\zeta_0}}{\sqrt{D}(\tilde{\alpha}+2\tilde{\nu})}.\end{gathered} \tag{26}$$

Below, $\alpha_{1,2} = \theta(z)\theta(-z+z_2)$ and $\alpha_{2,3} = \theta(z-z_2)$. For the ISTM to be applicable to (24) and (25), the following constraint must be imposed on the model coefficient:

$$h = 1. \tag{27}$$

Condition (27) can be satisfied for a certain choice of the detuning $\tilde{\nu}$ and/or the amplitude of additional field $\mathcal{G}$.

System (24), (25) is invariant with respect to the following simultaneous transformations:

$$\begin{aligned} z &\longrightarrow \tilde{z} = z + (1-f^2)\tau/g^2, \\ Q &\longrightarrow \tilde{Q} = Q\exp[-2i(1-f^2)z], \\ U &\longrightarrow \tilde{U} = U\exp[-i(1-f^2)z], \\ Q^{\pm} &\longrightarrow \tilde{Q}^{\pm} = fQ^{\pm}, \end{aligned} \tag{28}$$

where $f$ is an arbitrary real constant. Clearly, the conditions used above to derive the integrable system (24), (25) limit its applicability. However, the model integrability allows the form of the nonlinear field dynamics to be analyzed analytically, which cannot be done for nonintegrable models. Studying the model allows qualitative characteristics of the generated pulse packets in close systems to be determined analytically, because the self-similar solutions of integrable models generally retain their characteristic singularities in the case of a small departure from the integrability conditions. These conditions can also be used as test conditions in numerical simulations. The effects of departures from the integrability conditions on the form of the solutions can be analyzed by using the perturbation theory. Note that such an analysis was performed for the model combining the system of Maxwell–Bloch equations and the nonlinear Schrödinger equation by Doktorov and Prokopenya [27]. These authors showed that when an integrability condition similar to condition (27) was slightly violated, the basic properties of solitons did not change, to a first approximation; in particular, their amplitudes and velocities did not change. One might also expect similar results for the composite model considered here.

Let us now formulate the initial value–boundary value problem for the composite model (24), (25), which is solved below. Since the composite model is assumed to be integrable, it will suffice to formulate the conditions for $\tau = 0$ and $z = z_1 = 0$. Indeed, the condition of integrability through the ISTM suggests the existence of Lax representation in the form (1), (2), which is analogous to the separation of variables. In this case, the $z$ dependence of the solution is entirely determined by the $z$ dependence of scattering data, which, for our problem, reduces to an ordinary differential equation for the scattering matrix [see Eq. (44) below]. The solution of this equation has the following obvious property: the value of the scattering matrix at point $z = z_2$ (i.e., at the end of the resonant medium) is the boundary value for the scattering matrix that describes the evolution in the second (Kerr) medium. Thus, it will suffice to specify the boundary condition at point $z = z_1$, because in this case, the boundary value for the Kerr medium at point $z = z_2$ is determined by the field evolution in the resonant medium; i.e., in the integrable model, by the solution of Eq. (44) at this point. We emphasize that this property results from the integrability of the entire composite model. The initial conditions for the field must be common to both media, because the spectral problem is common to both models of the field interaction in the resonant and Kerr media.

For simplicity, we choose the initial conditions that correspond to trivial asymptotics at $\tau \longrightarrow \infty$ and $\tau = 0$. In general, the ISTM apparatus presented in the next two sections corresponds to the following initial-boundary conditions:

$$\begin{gathered} q^{\pm} = 0, \quad q(\tau, 0) = q_0(\tau), \quad \infty > \tau > 0, \\ q_0(0) = 0, \quad \lim_{\tau\to\infty} q(\tau, z) = 0, \ \forall z, \\ U(0, z) = U_0(z), \quad q(0, z) = 0, \quad 0 \le z \le z_2, \\ U(0, z) = 0, \quad q(0, z) = q_1(z), \quad z > z_2, \end{gathered} \tag{29}$$

where

$$q = \frac{Q}{ig}, \quad q^{\pm} = \frac{Q^{\pm}}{ig}.$$

The constant $g$, such that $\mathrm{Im}g^2 = 0$, is defined in (24) and (25). The function $U_0(z)$ is specified on a finite interval, is limited, and rapidly decreases at infinity. The functions $q_0(\tau)$ and $q_1(z)$ are smooth and limited (see below). This apparatus allows soliton and other related solutions that become zero at infinity, as well as the radiative solution, to be found. In Section 6, we solve a special boundary-value problem with the conditions

$$q_0 \equiv 0, \quad q_1 \equiv 0, \quad \partial z(U_0/|U_0|) = 0.$$

## 4. THE RIEMANN–HILBERT PROBLEM FOR THE WKI SPECTRAL PROBLEM

The Lax representation for the integrable composite model (24), (25) is

$$\partial_\tau \Phi = \begin{pmatrix} -i\lambda^2 - ig^{-2} & \lambda q \\ \lambda\bar{q} & i\lambda^2 + ig^{-2} \end{pmatrix} \Phi \equiv \mathcal{L}_c \Phi, \qquad (30)$$

$$\tau \geq 0,$$

$$\partial_z \Phi = \frac{i\alpha_{1,2}(z)\lambda g^3}{\lambda^2 g^2 + 1} \begin{pmatrix} \lambda g|U|^2 & iU^2 \\ i\bar{U}^2 & -\lambda g|U|^2 \end{pmatrix} \Phi$$
$$+ \alpha_{2,3}(z) \begin{pmatrix} -iH_{11} & H_{12} \\ H_{21} & iH_{11} \end{pmatrix} \Phi \equiv \mathcal{A}_c \Phi, \qquad (31)$$

here, $\lambda$ is the spectral parameter,

$$\begin{aligned} H_{11} &= [2(\lambda^2 + g^{-2})^2 - |q^{\pm}|^2(2\lambda^2 + g^{-2}) - \lambda^2|q|^2], \\ H_{12} &= \lambda[2q(\lambda^2 + g^{-2} - |q^{\pm}|^2) + i\partial_\tau q - |q|^2 q], \\ H_{21} &= \lambda[2\bar{q}(\lambda^2 + g^{-2} - |q^{\pm}|^2) - i\partial_\tau \bar{q} - |q|^2\bar{q}], \end{aligned} \qquad (32)$$

and the bar denotes a complex conjugate. The integrable system (24), (25) and its Lax representation are given for the first time. At $g = 0$, this system reduces to the standard system combining the second-harmonic generation equations without Kerr nonlinearity and the nonlinear Schrödinger equation.

Below, we provide basic information from the ISTM apparatus for the WKI spectral problem (30) for a semi-infinite interval $[\tau_1 = 0, \tau_2 = \infty)$.

For $q(\tau, z)$ decreasing rapidly enough as $\tau \longrightarrow \infty$ (this function and all its derivatives decrease more rapidly than any positive power of $\tau$), we define the vector functions $\psi^{\pm}(\tau, z; \lambda)$ and $\phi^{\pm}(\tau, z; \lambda)$ as the Jost functions of the system

$$(\partial_\tau - \mathcal{L}_c(\tau, z; \lambda))\psi^{\pm}(\tau, z; \lambda) = 0,$$
$$(\partial_\tau - \mathcal{L}_c(\tau, z; \lambda))\phi^{\pm}(\tau, z; \lambda) = 0$$

[see (30)] with the following values at the ends of the semi-infinite interval:

$$\lim_{\tau \to +\infty} \psi^+(\tau, z; \lambda) = (0, 1)^T e^{i\Lambda^2\tau},$$
$$\lim_{\tau \to +\infty} \psi^-(\tau, z; \lambda) = (1, 0)^T e^{-i\Lambda^2\tau}, \qquad (33)$$

$$\phi^+(0, z; \lambda) = (1, 0)^T, \quad \phi^-(0, z; \lambda) = (0, -1)^T, \qquad (34)$$

where $\Lambda^2 = \lambda^2 + g^{-2}$; the superscripts ± mean that the function is analytic at $\mathrm{Im}\lambda^2 \gtrless 0$, respectively; and $T$ denotes transposition. It follows from the Cauchy theorem that condition (34) is satisfied for $\forall\tau < 0$.

The Jost functions are related by

$$\psi^{\pm}(\tau, z; \lambda) = \mp a^{\pm}(z; \lambda)\phi^{\mp}(\tau, z; \lambda)$$
$$+ b^{\mp}(z; \lambda)\phi^{\pm}(\tau, z; \lambda), \quad \mathrm{Im}\lambda^2 = 0, \qquad (35)$$

where $a^{\pm}$ and $b^{\pm}$ are the complex scalar functions of $\lambda$ and $z$. The superscript means that the function is analytically continuable to the corresponding half plane of the complex $\lambda^2$ plane. For example, $a^{\pm}(\lambda)$ is analytically continuable to the domain where $\mathrm{Im}\lambda^2 \gtrless 0$.

For $\lambda \longrightarrow \infty$ and $\mathrm{Im}\lambda^2 > 0$, we have

$$a^+(\lambda) = 1 + \mathbb{O}(\lambda^{-2}),$$

$$\psi^+(\tau, z; \lambda)e^{-i\Lambda^2 z} = \left(\frac{q(\tau, z)}{2\lambda}, 1\right)^T + \mathbb{O}(\lambda^{-2}),$$

$$\phi^+(\tau, z; \lambda)e^{i\Lambda^2 z} = \left(1, \frac{\bar{q}(\tau, z)}{2\lambda}\right)^T + \mathbb{O}(\lambda^{-2}),$$

for $\lambda \longrightarrow \infty$ and $\mathrm{Im}\lambda^2 < 0$,

$$a^-(z; \lambda) = 1 + \mathbb{O}(\lambda^{-2}),$$

$$\psi^-(\tau, z; \lambda)e^{i\Lambda^2 z} = \left(1, \frac{\bar{q}(\tau, z)}{2\lambda}\right)^T + \mathbb{O}(\lambda^{-2}),$$

$$\phi^-(\tau, z; \lambda)e^{-i\Lambda^2 z} = -\left(\frac{q(\tau, z)}{2\lambda}, 1\right)^T + \mathbb{O}(\lambda^{-2}),$$

and for $\mathrm{Im}\lambda^2 = 0$,

$$\begin{gathered} a^+(\lambda)a^-(\lambda) + b^+(\lambda)b^-(\lambda) = 1, \\ a^+(\lambda) = \overline{a^-(\bar{\lambda})}, \\ b^+(\lambda) = -\overline{b^-(\bar{\lambda})}, \quad a^{\pm}(-\lambda) = a^{\pm}(\lambda), \\ b^{\pm}(\lambda) = -b^{\pm}(-\lambda), \\ a^+(\lambda)a^-(\lambda) + b^+(\lambda)b^-(\lambda) = 1, \\ a^+(\lambda) = \overline{a^-(\bar{\lambda})}, \quad b^+(\lambda) = -\overline{b^-(\bar{\lambda})}, \\ a^{\pm}(-\lambda) = a^{\pm}(\lambda), \quad b^{\pm}(\lambda) = -b^{\pm}(-\lambda). \end{gathered} \qquad (36)$$

Expressions (36) are valid in the $\lambda$ plane on contour $\hat{\Gamma} = \{\lambda; \mathrm{Im}\lambda^2 = 0\}$, which is oriented as shown in Fig. 1. As above, the bar denotes a complex conjugate.

Assume that

$$\psi(\tau, z; \lambda) \equiv \mu(\tau, z; \lambda)\exp(-i\Lambda^2\tau\sigma_3),$$

$\sigma_3 = \mathrm{diag}(1, -1)$ is the Pauli matrix. Let us define

$$\mu^+(\tau, z; \lambda) \equiv \begin{pmatrix} \dfrac{\phi_1^+(\tau, z; \lambda)}{a^+(\lambda, t)} & \psi_1^+(\tau, z; \lambda) \\ \dfrac{\phi_2^+(\tau, z; \lambda)}{a^+(\lambda, t)} & \psi_2^+(\tau, z; \lambda) \end{pmatrix} \exp(i\Lambda^2\tau\sigma_3),$$

for $\mathrm{Im}\lambda^2 > 0$ and

$$\mu^-(\tau, z; \lambda) \equiv \begin{pmatrix} \psi_1^-(\tau, z; \lambda) & -\dfrac{\phi_1^-(\tau, z; \lambda)}{a^-(\lambda, t)} \\ \psi_1^-(\tau, z; \lambda) & -\dfrac{\phi_2^-(\tau, z; \lambda)}{a^-(\lambda, t)} \end{pmatrix} \exp(i\Lambda^2\tau\sigma_3).$$

for $\mathrm{Im}\lambda^2 < 0$.

Introduce the scattering coefficient

$$\rho^\pm(\lambda) = b^\pm(\lambda)/a^\pm(\lambda),$$

whose analytic properties follow from the above formulas. Next, in order not to write the superscripts, we designate $\rho(\lambda) \equiv \rho^-(\lambda)$.The $2 \times 2$ matrix function $\mu(\tau, z; \lambda)$ [$\det\mu(\tau, z; \lambda) = 1$] is then the solution of the following Riemann–Hilbert problem:

(i) $\mu(\tau, z; \lambda)$ is holomorphic for $\forall\lambda \in \mathbb{C}\backslash\hat{\Gamma}$;

(ii) $\mu(\tau, z; \lambda)$ satisfies the condition

$$\mu^+(\tau, z; \lambda) = \mu^-(\tau, z; \lambda)E(\tau; \lambda)^{-1}G(\lambda)E(\tau; \lambda), \quad \lambda \in \hat{\Gamma}, \tag{37}$$

where

$$G(\lambda) = \begin{pmatrix} 1 - \overline{\rho(\bar{\lambda})}\rho(\lambda) & \rho(\lambda) \\ -\overline{\rho(\bar{\lambda})} & 1 \end{pmatrix},$$

$$E(\tau; \lambda) = \exp(i\Lambda^2\tau\sigma_3),$$

with $\rho(\lambda)$ defined for $\lambda \in \hat{\Gamma}$ and having the property $\rho(-\lambda) = -\rho(\lambda)$;

(iii) for $\lambda \longrightarrow \infty$, $\lambda \in \mathbb{C}\backslash\hat{\Gamma}$,

$$\mu(\tau, z; \lambda) = \mathbf{I} + \mathbb{O}(\lambda^{-1}),$$

**I** is a unit matrix. These properties follow from the definition of $\mu(\tau, z; \lambda)$ for $\mathrm{Im}\lambda^2 \gtrless 0$ and from the analytic properties of the Jost functions. Let the integral of the modulus of $\rho$ as a function of $\lambda$ over all values of variable $\lambda \in \hat{\Gamma}$ be limited. The Riemann–Hilbert problem formulated above is then uniquely solvable:

$$\psi(\tau, z; \lambda) \equiv \mu(\tau, z; \lambda)\exp(-i\Lambda^2\tau\sigma_3)$$

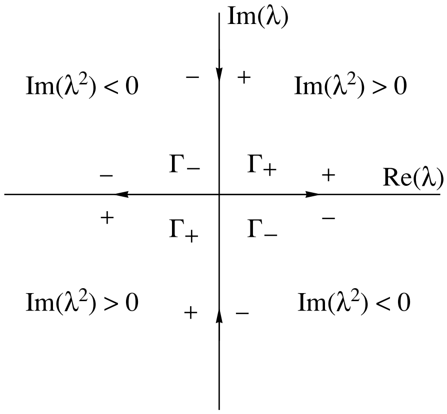

**Fig. 1.** $\hat{\Gamma}$ continuum: $\{\lambda, \mathrm{Im}\lambda^2 = 0\}$ and integration contours $\Gamma_\pm$.

is the solution of system (30), (31),

$$\begin{aligned} q(\tau, z) &\equiv 2i\lim_{\lambda \to \infty}(\lambda\mu(\tau, z; \lambda))_{12}, \\ \bar{q}(\tau, z) &\equiv -2i\lim_{\lambda \to \infty}(\lambda\mu(\tau, z; \lambda))_{21}, \end{aligned} \tag{38}$$

and $\mu(\tau, z; \lambda)$ has the following symmetry properties:

$$\mu(\tau, z; -\lambda) = \sigma_3\mu(\tau, z; \lambda)\sigma_3,$$

$$\mu(\tau, z; \lambda) = \sigma_1\overline{\mu(\tau, z; \bar{\lambda})}\sigma_1,$$

where $\sigma_1$ is the Pauli matrix. The solvability of the Riemann–Hilbert problem for $\forall z$ follows from the integral of the modulus of $\rho$ over $\lambda \in \hat{\Gamma}$ being limited (see [16, 28–31] for details). Conditions (38) follow from (30) and from the expansion of $\Phi_{1,2}$ in terms of $\lambda^{-1}$ for $\lambda \longrightarrow \infty$.

The solution of the Riemann–Hilbert problem (37) for the case under consideration is well known. It reduces to the solution of the Volterra integral equations. For $a(\lambda) \neq 0$, we derive the Volterra equations from (37) (see, e.g., [11, 30, 31]):

$$\bar{\psi}_1^+(\tau, z; \lambda) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \frac{1}{2i\pi}\int_{\Gamma_+}\bar{\rho}(z; \lambda)\exp(2i\tau\Lambda^2)\bar{\psi}_2^-(\tau, z; \lambda)\frac{d\zeta}{\zeta - \lambda}, \tag{39}$$

$$\bar{\psi}_2^-(\tau, z; \lambda) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{2i\pi}\int_{\Gamma_-}\rho(z; \lambda)\exp(-2i\tau\Lambda^2)\bar{\psi}_1^+(\tau, z; \lambda)\frac{d\zeta}{\zeta - \lambda}, \tag{40}$$

where $\Lambda^2 = \zeta^2 + g^{-2}$. $\Gamma_+$ is the integration contour that combines the paths along the axes in the first and third quadrants of the $\lambda$ plane (see Fig. 1) and the arcs at infinity in these quadrants that connect them; $\Gamma_-$ is a similar contour in the second and fourth quadrants of the $\lambda$ plane. Thus, the problem has been reduced to solving the integral equations (39) and (40). Let the coefficient $a(z; \lambda_k) = 0$ and $\mathrm{Im}\lambda^2 \neq 0$ at a finite number of isolated nondegenerate poles $\lambda_k$, $k = 1, 2, \ldots$. The sum over the residues at all these poles must then be added to the integrals on the right-hand sides of (39) and (40).

To solve the composite evolutionary model (24), (25) for field $q(\tau, z)$ by the ISTM, we must solve Eqs. (39) and (40) by taking into account the $z$ dependences of $\rho = \bar{b}/a$ and $\bar{\rho} = b/a$ and then restore $q(\tau, z)$ using Eq. (38).

## 5. THE $z$ DEPENDENCE OF COEFFICIENT $\rho$

The next step of the ISTM procedure is to find the $z$ dependence of the spectral data required for the problem to be solved. Below, we omit the superscripts ±. Let us first consider the general case and find the $\lambda$ dependence of coefficients $a$ and $\bar{b}$ for different optical media located on $N$ intervals $[z_i, z_{i+1}]$. To this end, we rewrite (35) in matrix form:

$$\psi(\tau, z; \lambda) = \phi(\tau, z; \lambda)T(z; \lambda). \tag{41}$$

It follows from the symmetry properties of the specific spectral problem (30) that the matrix $T$ is

$$T(z, \lambda) = \begin{pmatrix} a(\lambda) & \bar{b}(\lambda) \\ b(\lambda) & \bar{a}(\lambda) \end{pmatrix}. \tag{42}$$

Substituting (41) in (30) yields

$$\frac{\partial}{\partial z}T(z; \lambda) = \tilde{A}^{+}T(z; \lambda) - T(z; \lambda)\tilde{A}, \tag{43}$$

where

$$\tilde{A}^{+}(z; \lambda) = \lim_{\tau \to \infty} E^{-1}(\tau; \lambda)\mathcal{A}(\tau, z; \lambda)E(\tau, \lambda),$$

$$\tilde{A}(z; \lambda) = E^{-1}(0, \lambda)\mathcal{A}(0, z, \lambda)E(0, \lambda),$$

$$E(\tau, \lambda) = \exp[i\sigma_3\tau(\lambda^2 + g^{-2})],$$

here, $\mathcal{A}$ is an arbitrary $2 \times 2$ matrix and $\sigma_3$ is the Pauli matrix.

Consider one finite interval $[\tau_1 = 0, \tau_2)$. In this case, $\beta_{1,2} \neq 0$; i.e., $L_j \equiv 0$ in (1) for $j > 1$. In what follows, we pass to the limit $\tau_2 \longrightarrow \infty$. We consider $N$ intervals $[z_i, z_{i+1}]$ in variable $z$. In this case, $\alpha_{i,i+1}(z) \neq 0$. The $z$ dependence of matrix $T$ is then given by the equation

$$\begin{aligned} \frac{\partial}{\partial z}T(z; \lambda) &= E^{-1}(\tau_2; \lambda)\mathcal{A}(\tau_2, z; \lambda)E(\tau_2; \lambda)T(z; \lambda) \\ &\quad - T(z; \lambda)E^{-1}(0; \lambda)\mathcal{A}(0, z; \lambda)E(0; \lambda). \end{aligned} \tag{44}$$

Let us introduce the functions

$$\mathcal{B}(y, z) = \int_0^z E^{-1}(y; \lambda)\sum_{i=1}^{N}\alpha_i(z)A_i(y, z; \lambda)E(y; \lambda)d\lambda,$$

$$B_i(y, z) = \int_{z_i}^{z} E^{-1}(y; \lambda)A_i(y, z; \lambda)E(y; \lambda)d\lambda,$$

$$z_i \leq z < z_{i+1}.$$

The formal solution of (44) is then

$$\begin{aligned} T_{i+1}(z; \lambda) &= \exp[\mathcal{B}(\tau_2, z)]T_1(0, \lambda)\exp[-\mathcal{B}(0, z)] \\ &= \exp[\mathcal{B}_2(\tau_2, z)]T_i(z_i, \lambda)\exp[-\mathcal{B}(0, z)] \\ &= \exp[\mathcal{B}_i(\tau_2, z)]\ldots\exp[\mathcal{B}_1(\tau_2, z)]T_1(z_1; \lambda) \\ &\quad \times \exp[-\mathcal{B}_1(0, z)]\ldots\exp[-\mathcal{B}_i(0, z)]. \end{aligned} \tag{45}$$

We see from this solution that the evolution of $T(z, \lambda)$ on interval $[z_i, z_{i+1}]$ is described by formula (45) with the boundary condition $T(z_i, \lambda)$.

For $\tau_2 \longrightarrow \infty$, the following relation holds:

$$\mathcal{A}_{12}(\tau_2, z; \lambda) = \mathcal{A}_{21}(\tau_2, z; \lambda) = 0$$

($\mathcal{A}_{12, 21}$ are the nondiagonal elements of matrix $\mathcal{A}$; it follows from Eq. (44) that

$$\frac{\partial}{\partial z}\mathcal{T}_{1,k}(z; \lambda) = -\mathcal{T}_{1,k}(z; \lambda)\tilde{A}_k(0, z; \lambda), \tag{46}$$

where

$$\mathcal{T}_{1,k} = \exp\left[-i\sigma_3\int_0^z\sum_{i=1}^{k}\alpha_{i,i+1}(s)\lim_{\tau_2 \to \infty}(A_i)_{11}(\tau_2, s; \lambda)ds\right]T_{1,k},$$

$$\tilde{A}_k(0, z; \lambda) = E^{-1}(0; \lambda)A_k(0, z; \lambda)E(0; \lambda).$$

Formula (46) describes the dynamics of matrix $\mathcal{T}_{1,k}$ on interval $[z_k, z_{k+1}]$ with the boundary value $\mathcal{T}_{1,k-1}$, $k > 2$.

The formal solution of (46) is

$$\begin{aligned} \mathcal{T}_{1,N} &= \mathcal{T}_{1,2}\mathcal{T}_{2,3}\ldots\mathcal{T}_{N-1,N}, \\ \mathcal{T}_{i,i+1}(z; \lambda) &= \exp\left[\int_{z_i}^{z_{i+1}}\alpha_{i,i+1}(s)\tilde{A}_i(0, s; \lambda)ds\right]. \end{aligned} \tag{47}$$

Let us turn to the case of two different media located on intervals $[z_1, z_2]$ and $[z_2, z_3]$. Below, we provide the solution of (44) for the elements of matrix $T$ in the case of an arbitrary $2 \times 2$ matrix $\mathcal{A}(0, z; \lambda)$ such that $\det\mathcal{A}(0, z; \lambda) \neq 0$ for $\forall z$, which is constant on intervals $[0, z_2]$ and $[z_2, \infty)$ and changes abruptly at point $z = z_2$.

Without loss of generality, we choose $\tilde{\alpha}_i(z) \equiv 1$, $i = 1, 2$. For conditions (54), the solution of (44) for $z \geq 0$ is

$$\begin{aligned} a(z; \lambda) &= \frac{a_0}{2\Omega}\{[(\Omega + \tilde{A}_{11}) + \rho_0\tilde{A}_{21}] \\ &\times \exp(-z\Omega - \Theta(z)) \\ &+ [(\Omega - \tilde{A}_{11}) - \rho_0\tilde{A}_{21}]\exp(z\Omega - \Theta(z))\}, \end{aligned} \tag{48}$$

$$\begin{aligned} \bar{b}(z; \lambda) &= \frac{a_0}{2\Omega}\{[(\Omega - \tilde{A}_{11})\rho_0 + \tilde{A}_{12}] \\ &\times \exp(-z\Omega - \Theta(z)) \\ &+ [(\Omega + \tilde{A}_{11})\rho_0 - \tilde{A}_{12}]\exp(z\Omega - \Theta(z))\}, \end{aligned} \tag{49}$$

where $\tilde{A}_{ij}$ are the elements of matrix $\tilde{A}\,(\tau = 0)$ [see (43)], which also change abruptly at point $z = z_2$, $\tilde{A}_{11} = -\tilde{A}_{22}$,

$$\Omega^2(z) = \tilde{A}_{11}^2(0, z; \lambda) + \tilde{A}_{12}(0, z; \lambda)\tilde{A}_{21}(0, z; \lambda),$$

$$\rho_0(\lambda) = \rho(z = 0, \lambda) = \frac{\bar{b}_0}{a_0},$$

$$\Theta(z) = \int_0^z \tilde{A}_{11}^+(s; \lambda)ds.$$

Finally, we obtain for the coefficient $\rho(z; \lambda)$

$$\begin{aligned} \rho(z; \lambda) &= \frac{\bar{b}}{a} \\ &= \frac{[(\Omega - \tilde{A}_{11})\rho_0 + \tilde{A}_{12}]e^{-2\Omega z} + [(\Omega + \tilde{A}_{11})\rho_0 - \tilde{A}_{12}]}{[(\Omega + \tilde{A}_{11}) + \rho_0\tilde{A}_{21}]e^{-2\Omega z} + [(\Omega - \tilde{A}_{11}) - \rho_0\tilde{A}_{21}]}. \end{aligned} \tag{50}$$

The coefficient $\rho_0(\lambda)$ can be determined from the spectral problem (30) for a given potential $q(\tau, 0)$. It is easy to show that $\rho_0 = 0$ if $q(\tau, 0) \equiv 0$. In the latter case, the solution of the model is determined only by the boundary conditions.

For the physical problem (24), (25) under consideration and specific initial-boundary conditions (54), the nondiagonal elements of matrix $\mathcal{A}_c$ [see the linear system (31)] are zero at $z > z_2$ (see the next section). It follows from the above formulas that the $z$ dependence of $T$ at $z > z_2$ is given by

$$T(z; \lambda) = T_+(z_2, z)T(z_2; \lambda)T_+^{-1}(z_2, z), \tag{51}$$

where $T_+(z_2, z)$ is the solution of the linear system (31) for zero nondiagonal elements of matrix $\mathcal{A}_c(\tau_2, z; \lambda)$:

$$T_+(z_2, z; \lambda) = e^{-i2\sigma_3\Lambda^4 z}, \quad z > z_2. \tag{52}$$

For the coefficients of matrix $T$ (42), we derive

$$\frac{\bar{b}}{a}(z, \lambda) = \frac{\bar{b}_2}{a_2}\exp(4i\Lambda^4 z), \quad z > z_2, \tag{53}$$

where $a_2$ and $\bar{b}_2$ are the elements of matrix $T(0, z_2, \lambda)$. It follows from the latter formula that the boundary condition for the medium on interval $[z_2, z_3]$ in our example of a composite model is determined by the $z$ dependence of $\rho$ in the medium on interval $[z_1, z_2]$.

## 6. THE QUASI-RADIATIVE SOLUTION OF THE COMPOSITE MODEL

Let us show that the quasi-radiative solution for the main field $q = Q/ig$ in the composite model (24), (25) can be generated by an additional field $U$ for zero initial and boundary values of the main field. To this end, we choose the following initial-boundary conditions:

$$\begin{gathered} q^{\pm} = 0, \quad q(\tau, 0) = 0, \\ \tau \geq 0, \quad q(0, z) = 0 \quad \forall z, \\ U(0, z) = U_0 \neq 0, \quad \partial_z\frac{U_0}{|U_0|} = 0, \\ z_1 = 0 \leq z \leq z_2, \quad U(0, z) = 0, \quad z > z_2. \end{gathered} \tag{54}$$

Since a linear analysis of the stability of the solution $q(\tau, z) \equiv 0$ of Eq. (24) for $z > z_2$ shows this solution to be stable, we assume, in view of (54), that the following relation holds

$$\lim_{\tau \to \infty} q(\tau, z) = 0 \quad \forall z.$$

Below, we show that the initial-boundary conditions (54) lead to the quasi-radiative solution of the model for a sufficiently large effective length of the resonant nonlinear medium.

Next, for the application of Eqs. (39) and (40) in terms of the ISTM, it remains to determine $\rho(z; \lambda)$ and the positions of zeros of $a(\lambda)$. Using solution (50), we obtain for model (24), (25) and conditions (54)

$$T(z; \lambda) = s_0\begin{pmatrix} \left[\coth(\Omega\tau) - \frac{(\mathcal{A}_0)_{11}}{\Omega} + \rho_0\frac{(\mathcal{A}_0)_{21}}{\Omega}\right]e^{i\Theta} & \left\{\rho_0\left[\coth(\Omega\tau) + \frac{(\mathcal{A}_0)_{11}}{\Omega}\right] - \frac{(\mathcal{A}_0)_{12}}{\Omega}\right\}e^{i\Theta} \\ \left\{\bar{\rho}_0\left[\coth(\Omega\tau) - \frac{(\mathcal{A}_0)_{11}}{\Omega}\right] - \frac{(\mathcal{A}_0)_{21}}{\Omega}\right\}e^{-i\Theta} & \left\{\left[\coth(\Omega\tau) + \frac{(\mathcal{A}_0)_{11}}{\Omega}\right] - \bar{\rho}_0\frac{(\mathcal{A}_0)_{12}}{\Omega}\right\}e^{-i\Theta} \end{pmatrix}. \tag{55}$$

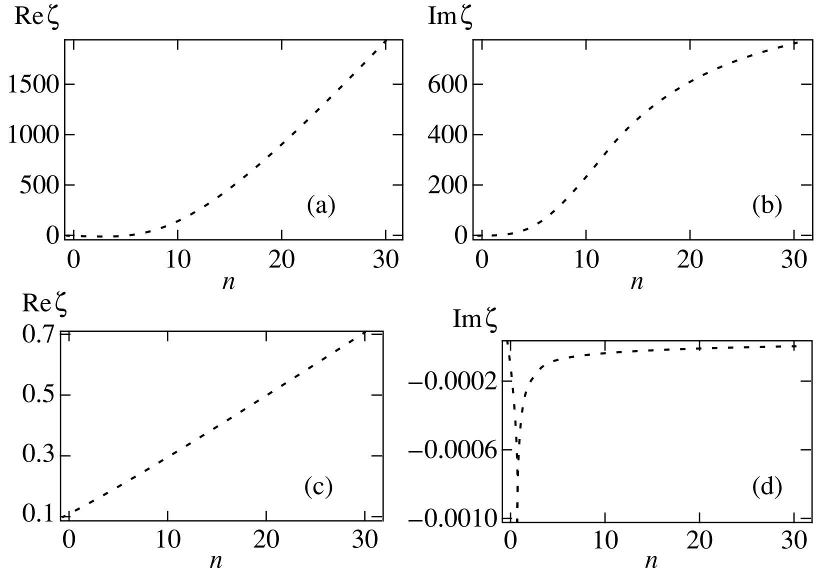

**Fig. 2.** Imaginary and real parts of ζ versus $n$; $L_0 = 0.01$ (a, b) and 50 (c, d).

Here,

$$s_0 = a_0\sinh(\Omega\tau), \quad \Theta = \int_0^z (\mathcal{A}_\infty)_{11}(s;\lambda)ds,$$

$(\mathcal{A}_0)_{ij}$ and $(\mathcal{A}_\infty)_{11}$ are the elements of matrix $\mathcal{A}_c(z, \tau; \lambda)$ (31), respectively, at $\tau = 0$ and $\tau = \infty$,

$$\Omega = \frac{i}{\Lambda^2}\{\alpha_{1,2}(z)[(1-g^2\Lambda^2)|U_0|^2 + \alpha_{2,3}(z)2\Lambda^6]^2 - \alpha_{1,2}(z)(1-g^2\Lambda^2)|U_0|^4\}^{1/2},$$

$$(\mathcal{A}_0)_{11} = -\alpha_{1,2}(z)i|U_0|^2\frac{1-g^2\Lambda^2}{\Lambda^2} - 2\alpha_{2,3}(z)i\Lambda^4.$$

We do not provide the expression for $(\mathcal{A}_0)_{12}$, because it is not used explicitly below.

The equation that gives the positions of poles, i.e.,

$$T_{11}(z; \lambda) = 0,$$

takes the form

$$a_0\left[\coth(\Omega\tau) - \frac{(\mathcal{A}_0)_{11}}{\Omega}\right] + \bar{b}_0\frac{(\mathcal{A}_0)_{21}}{\Omega} = 0. \tag{56}$$

Those solutions that lie in the first and third quadrants (see Fig. 1) correspond to soliton solutions. It follows from (56) that at $\bar{b}_0(\mathcal{A}_0)_{21} \neq 0$, the positions of poles are determined both by the solution of spectral problem (30) and by the solution of the additional problem (31). Since $q(0, z) \equiv 0$ in (54), we obtain $\rho_0 = 0$. In this case, the positions of poles in the complex domain at point $z$ is given by

$$\coth(\Omega z) - \frac{(\mathcal{A}_0)_{11}}{\Omega} = 0, \tag{57}$$

in which

$$(\mathcal{A}_0)_{ij}(z) \equiv (\mathcal{A}_0)_{ij}(z_2), \quad \forall z \in [z_1, z_2].$$

For the $z = z_2$ boundary of the resonant medium, Eq. (57) takes the form

$$\coth(i\zeta L_0) = \zeta, \tag{58}$$

where

$$\zeta = \pm\frac{\sqrt{g^2\Lambda^2 - 1}}{g\Lambda},$$

$L_0 = z_2|U_0|^2g^2$ is the effective length of the resonant medium. This equation has a countable set of solutions $\zeta(n)$ ($n$ is an integer) satisfying the equation

$$\zeta(n) + \frac{i}{L_0}\ln\frac{\zeta(n)+1}{\zeta(n)-1} = \frac{\pi n}{L_0}. \tag{59}$$

A numerical solution of Eq. (59) is shown in Fig. 2. We found the imaginary part of $\zeta(n)$ to approach a constant value as $n$ increases for $L_0 \ll 1$. At $L_0 \gg 1$, it may be assumed, with accuracy $\mathcal{O}(1/(L_0\ln L_0))$, that

$$\mathrm{Im}\zeta(n) = 0, \quad \mathrm{Re}\zeta(n) = \zeta(0) + n\pi/L_0. \tag{60}$$

Thus, the solution of integral equations (39) and (40) for $z \in [0, z_2]$ describes the field dynamics in the resonant medium. The $z$ dependence of ρ on this interval results in an infinite series of poles (59). Outside this interval, the $z$ dependence of $\rho$ has a simple form [see (53)]:

$$\begin{aligned}\rho(z;\lambda) &= \rho(z_2)\exp[4i(\lambda^2+g^{-2})^2 y] \\ &= \rho(z_2)\exp(4i\Lambda^4 y), \quad y = z - z_2, \quad z > z_2.\end{aligned} \tag{61}$$

To solve Eqs. (39) and (40), let us first calculate the sum over the residues at poles $\lambda_n$. These poles are related to $\zeta_n$ in Eq. (59) by

$$\lambda_n^2 = \frac{g^{-2}\zeta^2(n)}{1-\zeta^2(n)}.$$

Rewrite the right-hand side of Eq. (39) as

$$\begin{aligned}F_1(\tau, z; \xi) &= \sum_{n=-\infty}^{\infty}\frac{\exp[-i\Lambda_n^2\tau + 4i\Lambda_n^4(z-z_2)\theta(z-z_2)]}{\xi-\lambda_n} \\ &\times\frac{(c_1(\lambda_n)+c_2(\lambda_n))\exp[2iL_0\Omega(\lambda_n)]}{c_4(\lambda_n)\exp[2iL_0\Omega(\lambda_n)]2iL_0\Omega'(\lambda_n)}\psi_2(\tau;\lambda_n),\end{aligned} \tag{62}$$

where $\xi$ is an arbitrary complex variable,

$$\Omega'(\lambda_n) = \lim_{\lambda\to\lambda_n}\frac{\partial}{\partial\lambda}\Omega(\lambda),$$

$$\exp(2iL\Omega(\lambda_n)) = -\frac{c_3(\lambda_n)}{c_4(\lambda_n)}, \quad \Lambda_n^2 = \lambda_n^2 + g^{-2}.$$

The $\lambda$ dependences of $c_i$, $i = 1, \ldots, 4$, at $b_0(z;\lambda) = 0$ are

$$c_1(\lambda) = (\mathcal{A}_1)_{12} = -c_2(\lambda), \quad c_3(\lambda) = \Omega + (\mathcal{A}_1)_{11},$$

$$c_4(\lambda) = \Omega - (\mathcal{A}_1)_{11}$$

[these coefficients are the result of substituting solution (50)]; $(\mathcal{A}_1)_{ij}$ is the element of matrix $\mathcal{A}_c$ (31) at point $\tau = 0$, $z = z_1$. Thus, the first factor in sum (62) corresponds to the evolution of field $\mathcal{U}$ in the second medium ($z > z_2$), and the second factor is associated with its evolution in the first medium.

Consider a large effective length of the resonant medium, $L_0 \gg 1$. We substitute the $n$ dependence of $\lambda_n^2$ that follows from (60) in (62). Considering that all coefficients in sum (62) change monotonically at large $L_0$ and that each term of the sum is proportional to $L_0^{-1}$, we pass in (62) from the sum over $n$ to an integral over the continuous variable $\mu = n\pi/L_0$ in the limit $L_0 \longrightarrow \infty$:

$$\begin{aligned}&F_1(\tau, z; \xi) \\ &= g\int_{-+}\frac{\exp(2iX^2\tau + 4iX^4 y)}{\xi-\lambda}\frac{\psi_1(\tau;\mu)\mu d\mu}{\pi(1-\mu^2)^3} \\ &= \frac{g^3}{2\pi}\int_{-+}\frac{\exp(2i\tau X^2 + 4iX^4 y)}{\xi-\chi}\psi_1(\tau;\chi)dX^2, \\ &z > z_2.\end{aligned} \tag{63}$$

Here,

$$X^2 = g^{-2}(1-\mu^2)^{-2} + g^{-2} = \chi^2 + g^{-2},$$

$$y = z - z_2 > 0.$$

Similarly, we calculate the integral in Eq. (40). As a result, we obtain approximate, valid at $z > z_2$ with accuracy $\mathbb{O}(1/L_0\ln L_0)$, singular integral equations:

$$\begin{aligned}\bar{\psi}_1^+(\tau, y; \lambda) &= \begin{pmatrix}0\\1\end{pmatrix} + \frac{1}{2i\pi}\int_{\Gamma_+}\bar{\rho}_{\text{eff}}(\chi) \\ &\times\exp(2i\tau X^2 + 4iX^4 y)\bar{\psi}_2^-(\tau, y; \chi)\frac{d\chi}{\lambda-\chi},\end{aligned} \tag{64}$$

$$\begin{aligned}\bar{\psi}_2^-(\tau, y; \lambda) &= \begin{pmatrix}1\\0\end{pmatrix} + \frac{1}{2i\pi}\int_{\Gamma_-}\rho_{\text{eff}}(\chi) \\ &\times\exp(-2i\tau X^2 - 4iX^4 y)\bar{\psi}_1^+(\tau, y; \chi)\frac{d\chi}{\lambda-\chi},\end{aligned} \tag{65}$$

where $\rho_{\text{eff}}(\chi) = 2ig^3\chi$.

These equations describe the radiative solution of MNSE with effective scattering coefficients $\rho_{\text{eff}}$, which is associated with the continuum of the WKI spectral problem. This solution reduces to Painleve transcendent $P_{IV}$ [32].

Let us find the solution for small $\tau$. Using (38), we obtain for the limit $(\bar{\psi}_1^+)_2 \approx 1$, $(\bar{\psi}_2^-)_2 \ll 1$,

$$\begin{aligned}q(\tau, y) &\approx \frac{\sqrt{i}g^3\exp(i\tau^2/8y)}{2\pi\sqrt{2y}} \\ &\times\left[D_{-1}\left(\frac{\sqrt{i}\tau}{\sqrt{2y}}\right) + D_{-1}\left(-\frac{\sqrt{i}\tau}{\sqrt{2y}}\right)\right], \\ &y = z - z_2 > 0,\end{aligned} \tag{66}$$

where $D_{-1}$ is the function of a parabolic cylinder [33]. Below, we provide the asymptotics of the solution for $y \longrightarrow \infty$ by using the results of [16]. The asymptotic solution for $q$ is

$$\begin{aligned}q(\tau, y) &\approx \sqrt{\pm\frac{\mu(\lambda_0)}{2\lambda_0^2 y}}\exp\Bigg\{i\Bigg[4\lambda_0^4 y \mp \mu(\lambda_0)\ln|y| \\ &+ \phi^{\pm}(\lambda_0) + \hat{\Phi}^{\pm}(\lambda_0) - 2\vartheta^{\pm}(\lambda_0) + \frac{\pi}{2}\Bigg]\Bigg\} + \mathbb{O}\left(\frac{C_2(\lambda_0)\ln|y|}{\lambda_0^2 y}\right).\end{aligned} \tag{67}$$

Here, we introduced the variable $\lambda_0 = (1/2)\sqrt{\tau/y}$ and the functions of it:

$$\mu(\lambda_0) = -\frac{1}{2\pi}\ln(1-|r(\lambda_0)|^2),$$

$$\phi^-(\lambda_0) = \frac{1}{\pi}\int_{\lambda_0}^{\infty} \ln\left|\xi^2 - \lambda_0^2\right| d\ln(1 - |r(\xi)|^2)d\xi,$$

$$\phi^+(\lambda_0) = \frac{1}{\pi}\int_{0}^{\lambda_0} \ln\left|\xi^2 - \lambda_0^2\right| d\ln(1 - |r(\xi)|^2)d\xi$$

$$-\frac{1}{\pi}\int_{0}^{\infty} \ln\left|\xi^2 + \lambda_0^2\right| d\ln(1 + |r(i\xi)|^2)d\xi,$$

$$\hat{\Phi}^{\pm}(\lambda_0) = \pm\arg\Gamma(i\mu(\lambda_0))$$

$$+\arg r(\lambda_0) \mp 3\mu(\lambda_0)\ln 2 + \frac{(2\pm 1)\pi}{4},$$

$$\vartheta^+(\lambda_0) = -\int_{0}^{\lambda_0} \frac{\ln(1 - |r(\xi)|^2)}{\xi}\frac{d\xi}{\pi} + \int_{0}^{\infty} \frac{\ln(1 + |r(i\xi)|^2)}{\xi}\frac{d\xi}{\pi},$$

$$\vartheta^-(\lambda_0) = -\int_{\lambda_0}^{\infty} \frac{\ln(1 - |r(\xi)|^2)}{\xi}\frac{d\xi}{\pi},$$

$$r(\lambda_0) = \rho_{\text{eff}}(\lambda_0) = 2ig^3\lambda_0,$$

$C_2(s) \sim \mathcal{O}(1)$, and $\Gamma$ is the Gamma function.

## 7. CONCLUSION

We have analyzed a composite model that can be used to describe the generation of pulses in a system composed of a resonant two-level medium and an optical fiber. A nonzero field $U(0, z)$ was shown to produce a wave packet $Q(\tau, z)$. This packet is described by a solution that asymptotically approaches the quasi-radiative solution of MNSE as the effective length of the first medium increases ($L_0 \longrightarrow \infty$).

The approach used here is universal. The boundary conditions that lead to an infinite sequence of poles in the complex plane whose positions depend on variable quite often arise in problems of coherent nonlinear optics. For example, when solving the Maxwell–Bloch equations, which describe the dynamics of pulses in a two-level semi-infinite medium under conditions of single-frequency resonance, infinite series of moving poles also emerge for incomplete initial inversion of the medium [8]. Simulating the interaction of light with a two-level transition under conditions of two-frequency resonance and four-wave mixing in a semi-infinite Kerr medium [15] can lead to the same boundary conditions. For a degenerate two-frequency resonance of this kind, the boundary conditions that generate an infinite series of poles in the complex plane arise for a nonzero seed field slowly changing with time variable. For our example of a model of the field interaction with a composite medium, a nonzero field $U(0, z)$ results in nontrivial field dynamics for all $z$. Applying the ISTM apparatus to this model allowed the form of the asymptotic solution to be determined.

An asymptotic solution of the initial value–boundary value problem for the model that describes a degenerate two-frequency field interaction with a two-level medium (7) and (9) for a constant boundary value of the field amplitude without regard for phase modulation, i.e., for

$$\tilde{\beta} = \tilde{g}_1 = \tilde{g}_2 = 0$$

can be found in a similar way. In this case, a nontrivial asymptotic solution arises under the initial-boundary conditions (54) where

$$U_0 = \exp(ik_0z)|U_0(z)|, \quad k_0 \neq 0,$$

for an arbitrary finite amplitude $|U_0(z)|$.

One of our results is the proof that the asymptotics of the solution of the composite model for the initial-boundary conditions specified above is described by a self-similar solution and contains no true solitons (in the sense [1]) for any intensity $|U_0|^2$ of the additional field and a large effective length ($L_0 \gg 1$) of the resonant medium.

Our analytic apparatus can also be used for other composite models, while our solutions can be used as test ones for nearly integrable models and as a zero-order approximation in constructing the perturbation theory.

As was noted above, the laser facilities used to generate light usually include different nonlinear media, linear modulation of losses, etc. More complex composite integrable models can be employed in analyzing the evolution of light pulses in idealized models of the optical facilities that use different forms of nonlinear interaction between fields and media at different times. Combined schemes of the interaction with different media are known to offer additional possibilities for controlling the nonlinear interaction, field conversion, and pulse generation. As we showed here, using composite integrable models allows one to analyze such schemes in terms of exactly solvable models and simultaneously to extend the range of applications of integrable models.

## ACKNOWLEDGMENTS

This study was supported in part by the Russian Foundation for Basic Research (project no. 98-02-17904).

## REFERENCES

1. S. P. Novikov, S. V. Manakov, L. P. Pitaevskii, and V. E. Zakharov, *Theory of Solitons: the Inverse Scattering Method* (Nauka, Moscow, 1980; Consultants Bureau, New York, 1984).

2. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984; Nauka, Moscow, 1989).
3. A. I. Maimistov and A. M. Basharov, *Nonlinear Optical Waves* (Kluwer, Dordrecht, 1999).
4. Y. Kodama, J. Stat. Phys. **39**, 597 (1985); Y. Kodama and A. Hasegawa, IEEE J. Quantum Electron. **23**, 510 (1987); G. P. Agrawal, *Nonlinear Fiber Optics* (Academic, New York, 1989; Mir, Moscow, 1996).
5. *Optical Solitons—Theory and Experiment*, Ed. by J. R. Taylor (Cambridge Univ. Press, Cambridge, 1992), Cambridge Studies in Modern Optics, Vol. 10; H. A. Haus, Proc. IEEE **81**, 970 (1993); A. Hasegawa and Y. Kodama, *Solitons in Optical Communications* (Oxford Univ. Press, Oxford, 1995), Oxford Series in Optical and Imaging Sciences, No. 7.
6. M. Wadati, K. Konno, and Y. Ichikawa, J. Phys. Soc. Jpn. **46**, 1965 (1979); D. J. Kaup and A. C. Newell, J. Math. Phys. **19**, 798 (1978); H. Eichhorn, Inverse Probl. **1**, 193 (1985); J. H. B. Nijhof and G. H. M. Roelofs, J. Phys. A **25**, 2403 (1992).
7. E. K. Sklyanin, Funkt. Anal. Ego Prilozh. **21**, 86 (1987); A. S. Fokas and A. R. Its, Teor. Mat. Fiz. **92**, 387 (1992); A. S. Fokas and A. R. Its, Phys. Rev. Lett. **68**, 3117 (1992); E. D. Belokos, in *Algebraic and Geometrical Methods in Mathematical Physics* (Kluwer, Amsterdam, 1996), p. 263.
8. J. Leon and A. V. Mikhailov, Phys. Lett. A **53A**, 33 (1999); M. Boiti, J.-G. Caputo, J. Leon, and F. Pempinelli, Inverse Probl. **16**, 303 (2000).
9. O. M. Kiselev, Teor. Mat. Fiz. **98**, 29 (1994).
10. I. R. Gabitov, V. E. Zakharov, and A. V. Mikhaĭlov, Teor. Mat. Fiz. **63**, 11 (1985).
11. A. S. Fokas and C. R. Menyuk, J. Nonlinear Sci. **9**, 1 (1999).
12. D. J. Kaup and C. R. Menyuk, Phys. Rev. A **42**, 1712 (1990).
13. C. R. Menyuk, Phys. Rev. A **47**, 2235 (1993).
14. H. Steudel and D. J. Kaup, J. Phys. A **32**, 6219 (1999).
15. A. A. Zabolotskii, Zh. Éksp. Teor. Fiz. **115**, 1168 (1999) [JETP **88**, 642 (1999)].
16. A. V. Kitaev and A. H. Vartanian, Inverse Probl. **13**, 1311 (1997).
17. É. M. Belenov and I. A. Poluéktov, Zh. Éksp. Teor. Fiz. **56**, 1407 (1969) [Sov. Phys. JETP **29**, 754 (1969)].
18. I. A. Poluéktov, Yu. M. Popov, and V. S. Roĭtberg, Pis'ma Zh. Éksp. Teor. Fiz. **18**, 638 (1973) [JETP Lett. **18**, 373 (1973)].
19. H. Steudel, Physica D (Amsterdam) **6**, 155 (1983).
20. D. S. Butylkin, A. E. Kaplan, Yu. G. Khronopulo, and E. I. Yakubovich, *Resonant Interaction of Light with Matter* (Nauka, Moscow, 1977).
21. V. E. Zakharov and A. V. Mikhaĭlov, Pis'ma Zh. Éksp. Teor. Fiz. **45**, 279 (1987) [JETP Lett. **45**, 349 (1987)].
22. A. A. Zabolotskii, Phys. Lett. A **124**, 500 (1987).
23. N. Tzoar and M. Jain, Phys. Rev. A **23**, 1266 (1981).
24. K. J. Blow, N. J. Doran, and E. Commins, Opt. Commun. **48**, 770 (1983).
25. D. Marcuse, Appl. Opt. **19**, 1653 (1980).
26. M. Klauder, E. W. Laedke, K. H. Spatschek, and S. K. Turitsyn, Phys. Rev. E **47**, R3844 (1993).
27. E. V. Doktorov and I. N. Prokopenya, Zh. Éksp. Teor. Fiz. **100**, 1129 (1991) [Sov. Phys. JETP **73**, 623 (1991)].
28. V. S. Gerdzhikov, M. I. Ivanov, and P. P. Kulish, Teor. Mat. Fiz. **44**, 342 (1980).
29. J.-H. Lee, Trans. Am. Math. Soc. **314**, 107 (1989).
30. X. Zhou, Commun. Pure Appl. Math. **42**, 895 (1989); X. Zhou, SIAM (Soc. Ind. Appl. Math.) J. Math. Anal. **20**, 966 (1989).
31. X. Zhou, J. Diff. Eqns. **115**, 277 (1995).
32. M. J. Ablowitz, A. Ramani, and H. Segur, J. Math. Phys. **21**, 1006 (1980).
33. *Higher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1953; Nauka, Moscow, 1974), Vol. 2.

*Translated by V. Astakhov*

Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 236–246.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 269–279.
Original Russian Text Copyright © 2001 by Voronkov, Sazonov.

# Self-Induced Acoustic Transparency in the Long-Short-Wave Resonance Mode

**S. V. Voronkov and S. V. Sazonov***
*Kaliningrad State Technical University, Kaliningrad, 236000 Russia*
**e-mail: nst@alg.kaliningrad.ru*

Received January 29, 2001

**Abstract**—The effect of self-induced acoustic transparency for transverse–longitudinal pulses propagating along an external magnetic field in a system of resonance paramagnetic impurities with the effective spin $S = 1/2$ is theoretically investigated. In this case, the short-wave transverse component of the pulse causes quantum transitions, and the longitudinal long-wave component dynamically shifts the frequency of those transitions. When the speeds of the longitudinal and transverse acoustic waves in the crystal matrix are close to each other, both components interact in the mode of the long-short-wave resonance, which is described by a system of nonlinear integro-differential equations. It is shown that this interaction results, in particular, in the modulation of the carrier frequency of the circular-polarized component of the pulse. More precisely, the frequency in the neighborhood of the signal's maximum is less than in the vicinity of its edges. Solutions in the form of traveling $2\pi$-pulses are analyzed analytically and numerically. It is shown that there exist solutions that include a longitudinal component and cannot be reduced to well-known transverse solitons of the sinus–Gordon equation. © *2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

After the discovery of the optical self-induced transparency (SIT) in [1], experimental [2, 3] and theoretical [2, 4] search for the corresponding acoustic resonance effect (ASIT) in low-temperature ($T \approx 4$ K) crystal samples containing paramagnetic impurities proved fruitful. For example, in [2, 3], ASIT was investigated for $Fe^{2+}$ impurities, which possess the effective spin $S = 1$ and are interstitial in the crystal matrices MgO and $LiNbO_3$, when a microsecond longitudinal acoustic pulse propagates at an angle to the external magnetic field **B**. In [4], a theoretical treatment of ASIT is given for the transverse pulse propagating along **B** in the system of spins $S = 1/2$.

In the general case, the almost independent propagation of the longitudinal and transverse components of the acoustic pulse in a solid body is ensured by a considerable detuning of their linear speeds $a_\parallel$ and $a_\perp$. For the majority of crystals, such a situation actually takes place [5]. However, if the Cauchy relation [6] holds for the components of the elastic constant tensor in an elastically isotropic crystal (the speeds of the longitudinal and transverse components of the elastic field are independent of the direction), the speeds of the longitudinal, $a_\parallel$, and transverse, $a_\perp$, acoustic waves are equal to each other. These conditions are best satisfied with ionic crystals of halides of alkaline metals [6], for which the interaction force between atoms is central [7]. NaBr is a representative of this group of crystals [6].

When propagating in the crystal, both components of the acoustic pulse interact with quantum paramagnetic impurities, through which an effective nonlinear interaction between the two components can realize provided that $a_\parallel$ and $a_\perp$ are close enough. This interaction can substantially influence the resonance excitation of paramagnetic ions by the acoustic field, which affects the manifestation of specific features of ASIT, which were not considered earlier and are discussed in Section 4 of this paper.

In [8], soliton-like propagation modes of transverse–longitudinal pulses along **B** under the condition of a weak (the spectrum of the pulse does not overlap the quantum transitions) excitation of paramagnetic impurities with the effective spin $S = 1$ are investigated. The system of coupled equations for the longitudinal and transverse components of the elastic pulse with weak (power) nonlinearity derived in [8] describes the quasi-soliton dynamics in the long-short-wave resonance mode when the approximation of the slowly varying amplitudes and phases (SVAP), which is standard for quasi-monochrome pulses, is inapplicable. The role of the long-wave component of the elastic field is played by the longitudinal component of the acoustic pulse, and the role of the short-wave component is played by the transverse component. If the SVAP approximation is used, this system takes the form of the long-short-wave resonance systems of the type of the Zakharov [9] and Yadjima–Oikawa [10] equations.

In this paper, we consider the ASIT effect for quasi-monochrome acoustic pulses with longitudinal–transverse structure that propagate along **B** in a system of paramagnetic impurities with the effective spin $S = 1/2$ under the condition of long-short-wave resonance ($a_\parallel \approx a_\perp$).

In this case, the frequency of the transverse component of the monochromatic pulse $\omega$ coincides with the frequency $\omega_0$ of the Zeeman split of the Kramers doublet.

The paper is organized as follows. In Section 2, the semiclassic approach is used to derive self-consistent wave and constitutive equations describing the nonlinear dynamics of longitudinal–transverse acoustic pulses that propagate along an external magnetic field in a system of paramagnetic impurities. In Section 3, an asymptotic method for solving constitutive equations is suggested when neglecting the nonuniform broadening. This asymptotics is used as a basis for deriving a system of equations for the long-short-wave resonance in the strong nonlinearity mode. In the following section, solutions to this system in the form of stationary traveling pulses are analyzed analytically and numerically, and the main specific features of the acoustic self-induced transparency for longitudinal–transverse solitons are formulated. In conclusions, the main results obtained in this paper are presented, and the most interesting unsolved problems are formulated.

## 2. SELF-CONSISTENT EQUATIONS OF MOTION

The Hamiltonian of the Zeeman interaction of the $\alpha$ magnetic moment $\hat{\boldsymbol{\mu}}^{(\alpha)}$ with the magnetic field $\mathbf{B}$ has the form $\hat{H}^z_\alpha = -\hat{\boldsymbol{\mu}}^{(\alpha)} \cdot \mathbf{B}$. In turn, the components $\hat{\mu}^{(\alpha)}_j$ ($j = x, y, z$) of the vector operator $\hat{\boldsymbol{\mu}}^{(\alpha)}$ are related to the corresponding components of the spin $\hat{\mathbf{S}}^{(\alpha)} \equiv \hat{\mathbf{S}}(\mathbf{r}_\alpha)$ ($\mathbf{r}_\alpha$ is the radius vector of the spin $\alpha$) by the gyromagnetic equation $\hat{\mu}^{(\alpha)}_j = -\sum_k \mu_B g_{jk} \hat{S}^{(\alpha)}_k$, where $\mu_B$ is the Bohr magneton and $g_{jk}$ are the components of the Landé tensor $\hat{\mathbf{g}}$. Then,

$$\hat{H}^z = \sum_{\alpha=1}^{N} \hat{H}^z_\alpha = \mu_B \sum_\alpha \sum_{j,k} B_j g_{jk} \hat{S}^\alpha_k. \tag{1}$$

Here $N$ is the total number of spins. In terms of the basis consisting of the eigenfunctions of $\hat{S}^\alpha_z$, the components of the spin $S = 1/2$ are written in terms of the Pauli matrices as

$$\hat{S}^\alpha_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{S}^\alpha_y = \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$
$$\hat{S}^\alpha_z = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2}$$

We orient the coordinate axes $x$, $y$, and $z$ along the principal axes of the Landé tensor, which coincide with the symmetry axes of the crystal. Then, in the absence of strains, the $\hat{\mathbf{g}}$ tensor has a diagonal form with $g_{jk} = g^{(0)}_{jk} = g_{jj}\delta_{jk}$ ($\delta_{jk}$ is the Kronecker symbol). When propagating in the crystal, the acoustic wave distorts the components of the Landé tensor as

$$g_{jk} = g^{(0)}_{jk} + \sum_{p,q} \left(\frac{\partial g_{jk}}{\partial \mathcal{E}_{pq}}\right)_0 \mathcal{E}_{pq} + \ldots, \tag{3}$$

where $\mathcal{E}_{pq}$ is the elastic strain tensor of the crystal at the point where the paramagnetic ion is located. The strain tensor depends on the components of its displacement $\mathbf{U} = (U_x, U_y, U_z)$ as

$$\mathcal{E}_{pq} = \frac{1}{2}\left(\frac{\partial U_p}{\partial x_q} + \frac{\partial U_q}{\partial x_p}\right).$$

The subscript “0” of the terms under the summation sign means differentiation at $\mathcal{E}_{pq} = 0$.

Hence, the Hamiltonian can be written as the sum

$$\hat{H}^z = \hat{H}_s + \hat{H}_{\text{int}}, \tag{4}$$

where the Hamiltonians of the spin subsystem and spin–phonon interaction have the form

$$\hat{H}_s = \sum_\alpha \sum_j \mu_B g_{jj} B_j \hat{S}^{(\alpha)}_j, \tag{5}$$

$$\hat{H}_{\text{int}} = \sum_\alpha \sum_{j,k,p,q} \mu_B B_j F_{jkpq} \mathcal{E}_{pq} \hat{S}^{(\alpha)}_k. \tag{6}$$

Here $F_{jkpq} \equiv (\partial g_{jk}/\partial \mathcal{E}_{pq})_0$ are the constants of the spin–phonon coupling [11, 12].

We supplement Hamiltonians (5) and (6) with the Hamiltonian of the acoustic field

$$H_a = \int\left(\frac{1}{2\rho}\sum_j p_j^2 + \frac{1}{2}\sum_{j,k,l,m} \lambda_{jklm} \frac{\partial U_j}{\partial x_k}\frac{\partial U_l}{\partial x_m}\right) d^3\mathbf{r}, \tag{7}$$

where $\rho$ is the average density of the crystal, $p_j$ ($j = x, y, z$) are the components of the momentum due to dynamic shifts, and $\lambda_{jklm}$ is the tensor of the crystal elasticity moduli [13]. The integration in (7) is performed over the entire crystal volume. Here, following [8], we use the semiclassical approach in which the spin dynamics is described in the framework of quantum mechanics, and the elastic momentum field is described classically. For this reason, Hamiltonian (7), in contrast to (5) and (6), is a complex-valued functional rather than an operator.

From the microscopic point of view, the spin–phonon coupling in the case $S = 1/2$ appears due to the modulation of the spin–orbit interaction by the elastic field under the condition of “freezing” the orbital momentum [14]. In systems with higher spins, the dominant contribution to the spin–phonon interaction is made by quadratic (with respect to spin operators) terms $\sim \hat{S}_j \hat{S}_k + \hat{S}_k \hat{S}_j$, which correspond to the Waller and Van Vleck mechanisms

[12, 14]. For $S = 1/2$, these terms vanish due to anticommutativity of the Pauli operators [11, 14].

According to the semiclassical approach [8, 15], the following equation holds for the evolution of the density operator $\hat{\rho}^{(\alpha)}$ of the α spin:

$$i\hbar\frac{\partial\hat{\rho}^{(\alpha)}}{\partial t} = [\hat{H}^{z}, \hat{\rho}^{(\alpha)}], \tag{8}$$

where $\hat{\rho}^{(\alpha)}$ is written in the form

$$\hat{\rho}^{(\alpha)} = \begin{pmatrix} \rho_{22} & \rho_{21} \\ \rho_{12} & \rho_{11} \end{pmatrix};$$

the elastic momentum field obeys the classical Hamiltonian equations for continuous media:

$$\frac{\partial\mathbf{U}}{\partial t} = \frac{\delta H}{\delta\mathbf{p}}, \quad \frac{\partial\mathbf{p}}{\partial t} = -\frac{\delta H}{\delta\mathbf{U}}, \tag{9}$$

where $H = H_a + \langle\hat{H}_{\text{int}}\rangle$ and $\langle\ldots\rangle$ is the operation of quantum mechanical averaging. Using (9), the classical interaction Hamiltonian $\langle\hat{H}_{\text{int}}\rangle$ can be conveniently written in the form

$$\langle\hat{H}_{\text{int}}\rangle = \sum_{j,k,p,q}\int\mu_B n B_j F_{jkpq}\mathcal{E}_{pq}(\mathbf{r})\langle\hat{S}_k(\mathbf{r})\rangle d^3\mathbf{r}. \tag{10}$$

Here $n = \sum_\alpha \delta(\mathbf{r}-\mathbf{r}_\alpha)$ is the concentration of paramagnetic ions and $\delta(\mathbf{r}-\mathbf{r}_\alpha)$ is the Dirac delta function.

Let a transverse–longitudinal acoustic pulse be propagating in a cubic crystal along **B** and one of the fourth-order symmetry axes, which coincides with the axis $z$. Consider the one-dimensional case when all dynamical variables depend only on $z$ and $t$. In this case, rotation by 90° about the axis ($x \longrightarrow y$, $y \longrightarrow -x$, and $z \longrightarrow z$) and the reflections $x \longrightarrow -x$ and $y \longrightarrow -y$ are symmetry transformations. Taking into account the axial property of the vectors **B** and **S** (when one of the coordinate axes is inverted, the corresponding components of **B** and **S** remain unchanged, and the two others change their sign), we rewrite $\hat{H}_s$ and $\hat{H}_{\text{int}}$ as

$$\begin{gathered}\hat{H}_s = \sum_\alpha \hbar\omega_0\hat{S}_z^{(\alpha)}, \\ \hat{H}_{\text{int}} = \sum_\alpha\frac{\hbar\omega_0}{g}\{F_{11}\mathcal{E}_{zz}\hat{S}_z^{(\alpha)} + F_{44}(\mathcal{E}_{xz}\hat{S}_x^{(\alpha)} + \mathcal{E}_{yz}\hat{S}_y^{(\alpha)})\},\end{gathered} \tag{11}$$

where $\omega_0 = g\mu_B B/\hbar$ is the frequency of the Zeeman splitting of the Kramers doublet, $g = g_{xx} = g_{yy} = g_{zz}$, and $\hbar$ is the Planck constant.

Under our assumptions, the Hamiltonian $H_a$ has the form

$$\begin{aligned}H_a = \frac{1}{2}\int\Bigg\{&\frac{p_x^2 + p_y^2 + p_z^2}{\rho} + \lambda_{11}\left(\frac{\partial U_z}{\partial z}\right)^2 \\ &+ \lambda_{44}\left[\left(\frac{\partial U_x}{\partial z}\right)^2 + \left(\frac{\partial U_y}{\partial z}\right)^2\right]\Bigg\}d^3\mathbf{r}.\end{aligned} \tag{12}$$

In (10) and (11), we use the Focht notation for the indices of fourth-order tensors: $xx \longrightarrow 1$, $yy \longrightarrow 2$, $zz \longrightarrow 3$, $yz \longrightarrow 4$, $xz \longrightarrow 5$, and $xy \longrightarrow 6$.

Now we have from (8)–(12) and (2)

$$\frac{\partial^2\mathcal{E}_\perp}{\partial t^2} - a_\perp^2\frac{\partial^2\mathcal{E}_\perp}{\partial z^2} = \frac{nG_\perp}{\rho}\frac{\partial^2 S_\perp}{\partial z^2}, \tag{13}$$

$$\frac{\partial^2\mathcal{E}_\parallel}{\partial t^2} - a_\parallel^2\frac{\partial^2\mathcal{E}_\parallel}{\partial z^2} = \frac{nG_\parallel}{2\rho}\frac{\partial^2 W}{\partial z^2}, \tag{14}$$

$$\frac{\partial S_\perp}{\partial t} = i\left(\omega_0 + \frac{G_\parallel\mathcal{E}_\parallel}{\hbar}\right)S_\perp - i\frac{G_\perp\mathcal{E}_\perp}{\hbar}W, \tag{15}$$

$$\frac{\partial W}{\partial t} = -\mathrm{Im}\left(\frac{G_\perp\mathcal{E}_\perp}{\hbar}S_\perp^*\right), \tag{16}$$

where $G_\parallel = \hbar\omega_0 F_{11}/g$, $G_\perp = \hbar\omega_0 F_{44}/g$, $\mathcal{E}_\parallel = \mathcal{E}_{zz}$, $\mathcal{E}_\perp = \mathcal{E}_{xz} + i\mathcal{E}_{yz}$, $a_\parallel = \sqrt{\lambda_{11}/\rho}$, and $a_\perp = \sqrt{\lambda_{44}/\rho}$ are the velocities of the longitudinal and transverse acoustic waves, respectively, $W = (\rho_{22} - \rho_{11})/2$ is the inversion of population in the system of the Kramers doublets, $S_\perp$ is expressed in terms of the transverse components of the Bloch vector $U = (\rho_{21} + \rho_{12})/2$ and $V = (\rho_{21} - \rho_{12})/2i$ as $S_\perp = U + iV = \rho_{21}$.

It follows from the constitutive Bloch-type equations (15), (16) that the transverse component of the strain causes quantum transitions inside the Zeeman doublet as the pulse propagates in the Faraday geometry; the longitudinal component causes a shift in the frequency of those transitions.

The system of Eqs. (13)–(16) describes the self-consistent dynamics of paramagnetic impurities and acoustic pulses when the latter propagate along the fourth-order axis of the cubic crystal in parallel to the external magnetic field.

## 3. EQUATIONS OF SELF-INDUCED TRANSPARENCY FOR TRANSVERSE–LONGITUDINAL PULSES

Let the circular polarized transverse component of the acoustic field be a quasi-monochrome pulse of the duration $\tau_p$ with the filling frequency $\omega$ such that $\omega\tau_p \gg 1$. Then, this component is written in the form

$$\frac{G_\perp\mathcal{E}_\perp}{\hbar} = \Omega_\perp\exp[i(\omega t - kz)], \tag{17}$$

where $k$ is the wave number and $\Omega_\perp(z, t)$ is the slowly varying complex amplitude

$$\left|\frac{\partial\Omega_\perp}{\partial t}\right| \ll \omega|\Omega_\perp|, \quad \left|\frac{\partial\Omega_\perp}{\partial z}\right| \ll k|\Omega_\perp|. \tag{18}$$

In the constitutive equations, we pass to the representation of the rotating wave

$$S_\perp = R\exp[i(\omega t - kz)]. \tag{19}$$

Substituting (17) and (19) into (13), (15), and (16), neglecting (by virtue of (18)) the second derivatives with respect to $\Omega_\perp$, i.e., using the SVAP, and assuming that $k = \omega/a_\perp$, we obtain

$$\frac{\partial\Omega_\perp}{\partial t} + a_\perp\frac{\partial\Omega_\perp}{\partial z} = i\frac{nG_\perp^2\omega}{8\hbar\rho a_\perp^2}R, \tag{20}$$

$$\begin{aligned}\frac{\partial R}{\partial t} &= i(\Delta + \Omega_\parallel)R - i\Omega_\perp W,\\ \frac{\partial W}{\partial t} &= -\mathrm{Im}(\Omega_\perp R^*),\end{aligned} \tag{21}$$

where $\Delta = \omega_0 - \omega$ is the detuning of the pulse carrier frequency from the resonance frequency of the quantum transition, $\Omega_\parallel = G_\parallel\varepsilon_\parallel/\hbar$.

As before, the longitudinal component of the acoustic pulse is described by Eq. (14).

We will seek solutions to the constitutive Eqs. (21) under the exact resonance condition ($\Delta = 0$). In general, the existence of inhomogeneous broadening violates this condition. Assuming that $\Delta \sim 1/T_2^*$ ($T_2^*$ is the time of invertible relaxation in the system of Zeeman's doublets due to inhomogeneous broadening), we see from (21) that the inhomogeneous broadening can be neglected under the conditions $T_2^*/\tau_p \sim \Omega_\parallel T_2^* \sim |\Omega_\perp|T_2^* \gg 1$; hence, we can set $\Delta = 0$ in (21). Setting $T_2^* \sim 10^{-8}$ s (see [2, 12]), we find that $\tau_p \sim 10^{-9}$ s and $\Omega_\parallel$, $|\Omega_\perp| \sim 10^9$ s$^{-1}$. On the other hand, $|\Omega_\perp| \sim \omega_0 F\mathscr{E}$, where $F$ and $\mathscr{E}$ are the mean values of the constants of the spin–phonon interaction and relative strain. Assuming that $\omega_0 \sim 10^{11}$–$10^{12}$ s$^{-1}$ and $F \sim 10$–$10^2$ [11, 12], we find that $\mathscr{E} \sim 10^{-4}$, which corresponds to the intensity of acoustic pulses $I \approx \rho a^3\mathscr{E}^2 \approx 10$–$10^2$ W/cm$^2$ (here we assumed that $\rho \sim 2$ g/cm$^3$ and $a \sim 3 \times 10^5$ cm/s).

In order to solve system (21) under the conditions formulated above, it is convenient to rewrite this system in the matrix form

$$\frac{\partial\mathbf{Y}}{\partial t} = i\hat{A}\mathbf{Y}, \tag{22}$$

where

$$\mathbf{Y} = \begin{pmatrix} R \\ R^* \\ W \end{pmatrix},\qquad \hat{A} = \begin{pmatrix} \Omega_\parallel & 0 & -\Omega_\perp \\ 0 & -\Omega_\parallel & \Omega_\perp^* \\ -\Omega_\perp^*/2 & \Omega_\perp/2 & 0 \end{pmatrix}. \tag{23}$$

The values of the components of the vector $\mathbf{Y}$ are bounded: $0 \le |R|^2$ and $|W| \le 1/2$. Hence, it is seen from (22) that, as the time of the pulse action decreases, the nonzero components of the matrix $\hat{A}$ increase (in the limit, $\Omega_\parallel, |\Omega_\perp| \longrightarrow \infty$ as $\tau_p \longrightarrow 0$). Hence, the components of $\hat{A}$ are large dynamical parameters; thus, the Wentzel–Brillouin–Kramers–Jeffry method (WBKJ) is applicable [16]. It is seen from (23) that the matrix $\hat{A}(t)$ in the case under consideration does not commute with itself at different time moments; i.e. $[\hat{A}(t), \hat{A}(t')] \ne 0$. Hence, the solution to (22) can be symbolically written in the form of a chronological exponent. However, in the limit $t \longrightarrow t_0$ and $\|\hat{A}\| \longrightarrow \infty$ (where $\|\ldots\|$ is the operator norm), the elements of $\hat{A}$ do not noticeably change within the time $\Delta t = t - t_0$. Therefore, as $\Delta t \longrightarrow 0$ and $\|\hat{A}\| \longrightarrow \infty$, we have $[\hat{A}(t), \hat{A}(t')] \longrightarrow 0$ and

$$\mathbf{Y}(t) = \hat{U}(t, t_0)\mathbf{Y}(t_0), \tag{24}$$

where

$$\hat{U}(t, t_0) = \lim_{\substack{t \to t_0 \\ \|\hat{A}\| \to \infty}} \exp\left(i\int_{t_0}^{t}\hat{A}(t')dt'\right) \tag{25}$$

is the evolution operator [17]. Under the condition that all eigenvalues $\lambda_j$ ($j = 1, \ldots, N$) of the matrix

$$\hat{\theta} \equiv \int_{t_0}^{t}\hat{A}(t')dt'$$

are different, we can use the Sylvester formula for computing the exponent in (25) [18, 19]:

$$e^{i\hat{\theta}} = \sum_j\prod_{k \ne j}\frac{\hat{\theta} - \lambda_k\hat{I}}{\lambda_j - \lambda_k}e^{i\lambda_j}, \tag{26}$$

where $\hat{I}$ is the identity matrix.

The eigenvalues $\lambda$ of the matrix $\hat{\theta}$ are determined from the equation

$$\det(\theta_{jk} - \lambda\delta_{jk}) = 0. \tag{27}$$

The elements $\theta_{jk}$ of $\hat{\theta}$ are expressed in terms of the corresponding elements of $\hat{A}$ as

$$\theta_{jk} = \int_{t_0}^{t} A_{jk} dt' \approx A_{jk}\Delta t$$

(as $\Delta t \longrightarrow 0$). Setting

$$\lambda \approx p\Delta t \approx \int_{t_0}^{t} p dt',$$

in (27), we obtain for the matrix $\hat{A}$ of order $N \times N$

$$\det(\theta_{jk} - \lambda\delta_{jk}) \approx \det(A_{jk}\Delta t - p\delta_{jk}\Delta t)$$
$$= (\Delta t)^N \det(A_{jk} - p\delta_{jk}) = 0.$$

Comparing this equation with (27), we see that if $\{\lambda_j\}$ is the set of eigenvalues of the matrix

$$\hat{\theta} = \int_{t_0}^{t} \hat{A} dt',$$

then $\{p_j\}$ is the set of eigenvalues of the matrix $\hat{A}$ in the limit $\Delta t \longrightarrow t_0$ and $\|\hat{A}\| \longrightarrow \infty$; in addition,

$$\{\lambda_j\} = \left\{\int_{t_0}^{t} p_j dt'\right\}.$$

Setting

$$\hat{\theta} \approx \hat{A}\Delta t, \quad \lambda_j \approx p_j \Delta t \approx \int_{t_0}^{t} p_j dt'$$

in (26) and using (25), we find the evolution operator

$$\hat{U}(t, t_0) = \sum_{j=1}^{N} \exp\left(i\int_{t_0}^{t} p_j dt'\right) \prod_{k \neq j} \frac{\hat{A} - p_k \hat{I}}{p_j - p_k}. \qquad (28)$$

Thus, (24) and (28) determine a solution to system (22) for an arbitrary order $N$ if the WBKJ approximation is valid, i.e., when $\Delta t \longrightarrow 0$ and $\|\hat{A}\| \longrightarrow \infty$.

In our case, $N = 3$, and it is seen from (23) that

$$p_1 = 0, \quad p_2 = -p_3 = \sqrt{\Omega_\parallel^2 + |\Omega_\perp|^2}.$$

From the physical point of view, the limit $\Delta t \longrightarrow 0$ means that $\tau_p / T_2^* \ll 1$ (see above), which corresponds to neglecting the inhomogeneous broadening. In this case, the evolution operator for system (22) (or (21 for $\Delta = 0$) is written as

$$\hat{U}(t, t_0) = \hat{I} - 2\frac{\hat{A}^2}{\Omega^2}\sin^2\frac{\theta}{2} + i\frac{\hat{A}}{\Omega}\sin\theta, \qquad (29)$$

where

$$\Omega = \sqrt{\Omega_\parallel^2 + |\Omega_\perp|^2}, \quad \theta = \int_{t_0}^{t} \Omega dt'. \qquad (30)$$

Tending formally $t_0 \longrightarrow -\infty$ and taking into account that $\mathbf{Y}^\dagger(-\infty) = (0, 0, W_\infty)$, we obtain from (29), (30), and (24) that

$$R = -iW_\infty \frac{\Omega_\perp}{\Omega}\sin\theta + 2W_\infty \frac{\Omega_\parallel \Omega_\perp}{\Omega^2}\sin^2\frac{\theta}{2}, \qquad (31)$$

$$W = W_\infty \left(1 - 2\frac{|\Omega_\perp|^2}{\Omega^2}\sin^2\frac{\theta}{2}\right). \qquad (32)$$

The direct substitution of (31) and (32) into (21) at $\Delta = 0$ shows that solutions (31), (32) are valid if the derivatives of $\Omega_\perp/\Omega$ and $\Omega_\parallel\Omega_\perp/\Omega^2$ are neglected compared to the derivatives of $\theta$. The slow change of the coefficients of the periodic functions $\sin\theta$ and $\sin^2(\theta/2)$ in comparison with the periodic functions themselves is consistent with the basic assumptions of the WBKJ method [16].

For $\Omega_\parallel = 0$, Eqs. (21) turn into the system used in [20]. Setting $\Omega_\parallel = 0$ ($\Omega = |\Omega_\perp|$) in (31) and (32), we obtain the solutions that were obtained in [20] in a different way. This fact provides an important argument in favor of the general Eqs. (24), (28), which determine an asymptotic solution to system (22) as $\|\hat{A}\| \longrightarrow \infty$.

Substituting (31) and (32) into (20) and (14), we obtain

$$\frac{\partial\Omega_\perp}{\partial t} + a_\perp \frac{\partial\Omega_\perp}{\partial z} = -\alpha_\perp \frac{\Omega_\perp}{\Omega}\left(\sin\theta + 2i\frac{\Omega_\parallel}{\Omega}\sin^2\frac{\theta}{2}\right), \qquad (33)$$

$$\frac{\partial^2\Omega_\parallel}{\partial t^2} - a_\parallel^2 \frac{\partial^2\Omega_\parallel}{\partial z^2} = \alpha_\parallel \frac{\partial^2}{\partial z^2}\left(\frac{|\Omega_\perp|^2}{\Omega^2}\sin^2\frac{\theta}{2}\right), \qquad (34)$$

where

$$\alpha_\perp = -\frac{W_\infty n G_\perp^2 \omega_0}{8\hbar\rho a_\perp^2}, \quad \alpha_\parallel = -\frac{2W_\infty n G_\parallel^2}{\hbar\rho},$$

and $\Omega$ and $\theta$ are determined by (30).

Since $||\Omega_\perp| \sim 1/\tau_p$, we have $||\partial\Omega_\perp/\partial t| \sim \tau_p^{-2}$. Hence, it is seen from (33) that under the condition $\alpha_\perp \tau_p^2 \ll 1$ the pulse speed $v$ is close to $a_\perp$. If, in addition, the long-short-wave resonance condition ($a_\| = a_\perp = a$) is strictly satisfied and $\alpha_\| a_\|^{-2}\tau_p \ll 1$, we can perform a reduction with respect to the derivatives on both sides of (34) with the help of the approximation of the one-directional propagation [21] (or the slowly changing profile). In the thermodynamically equilibrium case, the initial inversion $\Omega_\infty = -0.5\tanh(\hbar\omega_0/2k_BT)$, where $k_B$ is the Boltzmann constant. For $\omega_0 \sim 10^{11}$ s$^{-1}$ and $T < 1$ K, we can assume that $W_\infty \approx -0.5$. Then,

$$\alpha_\perp\tau_p^2 \approx \frac{n\hbar\omega_0 F_{44}^2(\omega_0\tau_p)^2}{16g^2\rho a_\perp^2}.$$

For $F_{44} \sim 10$, $\tau_p \sim 10^{-9}$ s, $g = 2$, $\rho = 2$ g/cm$^3$, and $a_\perp = 3\times10^5$ cm/s, we obtain $\alpha_\perp\tau_p^2 \sim 10^{-4} \ll 1$. By the same taken for $F_{11} \sim 10$ we obtain, $\alpha_\| a_\|^{-2}\tau_p \sim 10^{-4}$. According to one-directional propagation approximation, we introduce the "local" time $\tau = t - z/a$ and the "slow" coordinate $\zeta = \epsilon z$, where $\epsilon \sim \alpha_\perp\tau_p^2$ and $\alpha_\| a^{-2}\tau_p$. Then,

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial\tau}, \quad \frac{\partial}{\partial z} = -\frac{1}{a}\frac{\partial}{\partial\tau} + \epsilon\frac{\partial}{\partial\zeta},$$

$$\frac{\partial^2}{\partial z^2} \approx \frac{1}{a^2}\frac{\partial^2}{\partial\tau^2} - \frac{2\epsilon}{a}\frac{\partial^2}{\partial\tau\partial\zeta}.$$

We neglected the term $\sim\epsilon^2$ in the last expression. The right-hand side of Eq. (34) $\sim\alpha_\| \sim \epsilon$; hence we set

$$\partial^2/\partial z^2 \approx a^{-2}\partial^2/\partial\tau^2,$$

there and then integrate this equation with respect to $\tau$. As a result, we obtain, instead of (33) and (34),

$$\frac{\partial\Omega_\perp}{\partial z} = -\mu_\perp\frac{\Omega_\perp}{\Omega}\left(\sin\theta + 2i\frac{\Omega_\|}{\Omega}\sin^2\frac{\theta}{2}\right), \tag{35}$$

$$\frac{\partial\Omega_\|}{\partial z} = \mu_\|\frac{\partial}{\partial\tau}\left(\frac{|\Omega_\perp|^2}{\Omega^2}\sin^2\frac{\theta}{2}\right). \tag{36}$$

Here $\mu_\perp = \alpha_\perp/a$, $\mu_\| = \alpha_\| a/2$, and $\theta$ is determined from (30) accurate to the change $t \longrightarrow \tau$ in the upper limit of the integral, and $\tau_0 \longrightarrow -\infty$ in the lower limit.

The integro-differential system obtained describes the effect of the acoustic self-induced transparency in the mode of long-short-wave resonance under strong spin–phonon nonlinearity. The role of the short-wave component of the elastic field is played by the complex envelope $\Omega_\perp$ of the transverse component of the pulse; the role of the long-wave component is played by the longitudinal component. It must be noted that the change of the two-sided system (33), (34) for the one-sided system (35), (36) is similar to passing from the nonintegrable Zakharov equations [9] to the integrable Yadjima–Oikawa system [10].

## 4. SPECIFIC FEATURES OF SELF-INDUCED TRANSPARENCY FOR TRANSVERSE–LONGITUDINAL PULSES

Let us write the complex envelope $\Omega_\perp$ in the form

$$\Omega_\perp(z, t) = |\Omega_\perp|e^{i\phi(z, t)}. \tag{37}$$

Substituting (37) into (33) and separating the real and imaginary parts, we obtain

$$\frac{\partial|\Omega_\perp|}{\partial t} + a_\perp\frac{\partial|\Omega_\perp|}{\partial z} = -\alpha_\perp\frac{|\Omega_\perp|}{\Omega}\sin\theta, \tag{38}$$

$$\frac{\partial\phi}{\partial t} + a_\perp\frac{\partial\phi}{\partial z} = -2\alpha_\perp\frac{\Omega_\|}{\Omega^2}\sin^2\frac{\theta}{2}. \tag{39}$$

For a purely transverse pulse ($\Omega_\| = G_\| = 0$ and $\Omega = \partial\theta/\partial t = |\Omega_\perp|$), Eq. (39) shows the absence of the phase modulation, and (38) is reduced to the sinus-Gordon equation for $\theta$ obtained in [4]:

$$\frac{\partial^2\theta}{\partial z\partial\tau} = -\mu_\perp\sin\theta.$$

Thus, the phase modulation is caused by the presence of the longitudinal component of the acoustic pulse.

The general analysis of system (34) (or (36)), (38), (39) seems rather difficult. Not pretending to generalize, we will seek a solution to this system in the form of a localized traveling pulse. Below, we call this pulse a soliton; we interpret this notion in the general sense, and do not assume that such a pulse interacts elastically with similar pulses. Thus, we seek a solution in the form $\Omega_\| = \Omega_\|(\xi)$, $\Omega_\perp = \Omega_\perp(\xi)$, $\phi = \phi(\xi)$, where $\xi = t - z/v$ and $v$ is the speed of the pulse propagation. Then, it follows from (34), (38), and (39) with regard for the fact that $|\Omega_\perp|$, $\Omega_\|$, $d\Omega_\|/d\xi \longrightarrow 0$ as $t \longrightarrow -\infty$ that

$$\begin{aligned} \Omega_\| &= -\beta\frac{|\Omega_\perp|^2}{\Omega^2}\sin^2\frac{\theta}{2}, \\ \frac{d|\Omega_\perp|}{d\xi} &= \alpha\frac{|\Omega_\perp|}{\Omega}\sin\theta, \\ \frac{d\phi}{d\xi} &= 2\alpha\frac{\Omega_\|}{\Omega^2}\sin^2\frac{\theta}{2}, \end{aligned} \tag{40}$$

where

$$\beta = \frac{\alpha_\|}{a_\|^2 - v^2}, \quad \alpha = \frac{\alpha_\perp}{a_\perp/v - 1}.$$

Let us derive the equation of the phase trajectory for the dependence of the longitudinal, $\Omega_\|$, and transverse, $|\Omega_\perp|$, pulse components. Considering $\theta$ as a parameter on which $|\Omega_\perp|$ depends, we can write

$$\frac{d|\Omega_\perp|}{d\xi} = \frac{d|\Omega_\perp|}{d\theta} = \Omega\frac{d|\Omega_\perp|}{d\theta}.$$

Substituting this expression into the second equation in (40) and integrating with regard for the conditions at infinity, we find

$$\int_0^{|\Omega_\perp|} \frac{|\Omega_\perp|^2 + \Omega_\parallel^2}{|\Omega_\perp|} d|\Omega_\perp| = 2\alpha \sin^2\frac{\theta}{2}.$$

Now, we find $\sin^2(\theta/2)$ from the first equation in (40) and substitute it into the last equation. Differentiating with respect to $|\Omega_\perp|$, we obtain

$$\frac{d\Phi}{d\sigma} = \frac{\sigma^2 + \sigma\Phi^2 + \Phi^3}{\sigma^2 + 3\sigma\Phi^2}. \qquad (41)$$

Thus, we passed to the dimensionless variables $\sigma = (|\Omega_\perp|/\delta)^2$ and $\Phi = -\Omega_\parallel/\delta$, where $\delta = 4\alpha/\beta$.

Below, Eq. (41) is called the phase trajectory equation.

From the first two equations in (40), we easily find that

$$\frac{d\sigma}{d\xi} = 2\sqrt{\alpha\Phi\left[\sigma - \frac{\delta}{\beta}(\sigma + \Phi^2)\Phi\right]}. \qquad (42)$$

Integrating system (41), (42), we can determine $\Omega_\parallel$ and $|\Omega_\perp|$ as functions of $t - z/v$. In the general case, it seems impossible to find the analytical solution to Eq. (41). However, it can be done for two opposite limiting cases.

1. Let $|\Omega_\perp|^2 \gg \Omega_\parallel^2$. In the variables $\sigma$ and $\Phi$, this corresponds to the condition $\Phi^2/\sigma \ll 1$. Then, only the first terms remain in the numerator and denominator of (41), and $d\Phi/d\sigma = 1$. Hence, taking into account the fact that $\Phi \longrightarrow 0$ as $\sigma \longrightarrow 0$, we obtain $\Phi = \sigma$. Now, neglecting $\Phi^2$ in the parentheses of the radicand in (42) in comparison with $\sigma$ and passing from the dimensionless variables to the original ones $|\Omega_\perp|$ and $\Omega_\parallel$, we find

$$|\Omega_\perp| = \frac{2}{\tau_p}\operatorname{sech}\left(\frac{t - z/v}{\tau_p}\right),$$

$$\Omega_\parallel = -\Omega_{\parallel m}\operatorname{sech}^2\left(\frac{t - z/v}{\tau_p}\right), \qquad (43)$$

$$\frac{d\phi}{d\xi} = -\frac{\Omega_{\parallel m}}{2}\operatorname{sech}^2\left(\frac{t - z/v}{\tau_p}\right),$$

where the speed $v$ is connected with the duration of the pulse $\tau_p$ by the equation

$$\frac{1}{v} = \frac{1}{a_\perp}(1 + \alpha_\perp\tau_p^2), \qquad (44)$$

and the amplitude of the longitudinal component $\Omega_{\parallel m} = \beta = \alpha_\parallel/(a_\parallel^2 - v^2)$.

According to (44) and the estimates made above, the speed of the soliton-like structure is very close to $a_\perp$. If $a_\perp$ is substantially different from $a_\perp$ ($a_\parallel > a_\perp$), then $\Omega_{\parallel m} \approx \alpha_\parallel/(a_\parallel^2 - a_\perp^2)$ (here, we set $v \approx a_\perp$) and it is almost independent of $\tau_p$ and the amplitude of the transverse component. This testifies that the longitudinal and transverse components affect each other insignificantly under the condition of a substantial detuning of the speeds of the longitudinal and transverse acoustic waves. At the same time, as it was mentioned above, the influence of the longitudinal and transverse components on each other is most effective at $a_\perp = a_\parallel = a$ (the long-short-wave resonance). Then,

$$\Omega_{\parallel m} \approx \frac{\alpha_\parallel}{2a^2(1 - v/a)} \approx \frac{\alpha_\parallel}{2\alpha_\perp\tau_p^2} = \frac{8\omega_0(F_{11}/F_{44})^2}{(\omega_0\tau_p)^2}.$$

This approximation is equivalent to the transition from system (33), (34) to (35), (36). Below, all numerical estimates will be made for the case $a_\perp = a_\parallel = a$; general analytic expressions will be given for the case $a_\perp \neq a_\parallel$; i.e., for system (33), (34). It is seen from (43) that the amplitude of the transverse component is $|\Omega_\perp|_m = 2/\tau_p$. Hence, we see that the condition $|\Omega_\perp|_m^2 \gg \Omega_{\parallel m}^2$ is equivalent to the inequality

$$(\omega_0\tau_p)^2 \gg (2F_{11}/F_{44})^2. \qquad (45)$$

For the majority of crystals, $F_{11} > F_{44}$ [11, 14]. The validity condition for the SVAP approximation has the form $(\omega_0\tau_p)^2 \gg 1$, which does not contradict (45). For $\omega_0\tau_p \sim 10^3$, $F_{44} \sim 10$, and $g = 2$, we have for the amplitude of the relative transverse strain the formula

$$|\mathcal{E}_\perp|_m = \frac{g|\Omega_\perp|_m}{\omega_0 F_{44}} = \frac{2g}{\omega_0\tau_p F_{44}} \sim 4 \times 10^{-4},$$

which corresponds to the intensity $I \approx \rho a^3|\mathcal{E}_\perp|^2 \sim 10^3$ W/cm$^2$ ($\rho = 2$ g/cm$^3$ and $a = 3 \times 10^5$ cm/s).

From (43) and (32), we find the dynamics of the inversion of population when the pulse under consideration passes through the medium:

$$W = W_\infty\left[1 - \frac{2|\Omega_\perp|_m^2\operatorname{sech}^2\left(\frac{t - z/v}{\tau_p}\right)}{|\Omega_\perp|_m^2 + \Omega_{\parallel m}^2\operatorname{sech}^2\left(\frac{t - z/v}{\tau_p}\right)}\right]. \qquad (46)$$

This formula shows that the inversion of spins at the maximum of the soliton's amplitude is incomplete ($W \neq -W_\infty$), which is due to the presence of the transverse component ($\Omega_{\parallel m} \neq 0$). This phenomenon can be easily explained in the framework of the model used. Indeed, the increase of $\Omega_\parallel$ (and $|\Omega_\perp|$ along with it) takes the quantum system off the resonance state (see (15) and (21)), which hinders its further excitation.

Solutions (43), (44) are actually corrections to the corresponding solutions obtained in [4] for purely transverse pulses; thus, they can be obtained from (40) by the successive approximation method with respect to $\Omega_\parallel$. It is clear that no radically new solutions can be obtained in this way. From this point of view, the other limiting case seems to be more interesting.

2. Let $|\Omega_\perp|^2 \ll \Omega_\parallel^2$. This condition is equivalent to the inequality $\Phi^2/\sigma \gg 1$; hence, we can neglect the first terms in the numerator and denominator of (41). Then, we have

$$\frac{d\Phi}{d\sigma} = \frac{1}{3}\left(1 + \frac{\Phi}{\sigma}\right).$$

The solution to this equation satisfying the condition $\Phi = 0$ at $\sigma = 0$ has the form $\Phi = \sigma/2$. Neglecting the term $\sigma$ in the parentheses on the right-hand side of (42) in comparison with $\Phi^2$, we obtain after the integration and return from the dimensionless variables to the original ones

$$|\Omega_\perp| = |\Omega_\perp|_m \operatorname{sech}^{1/2}\left(\frac{t - z/v}{\tau_p}\right),$$

$$\Omega_\parallel = (-\Omega_{\parallel m})\operatorname{sech}\left(\frac{t - z/v}{\tau_p}\right), \qquad (47)$$

$$\frac{d\phi}{d\xi} = -\frac{\Omega_{\parallel m}}{4}\operatorname{sech}\left(\frac{t - z/v}{\tau_p}\right),$$

where

$$|\Omega_\perp|_m = \frac{4}{\sqrt{2\beta\tau_p^3}}, \quad \Omega_{\parallel m} = \frac{2}{\tau_p},$$

and the relation between the speed of the soliton and its duration is determined as

$$\frac{1}{v} = \frac{1}{a_\perp}(1 + 2\alpha_\perp\tau_p^2). \qquad (48)$$

Setting $\Omega^2 \approx \Omega_\parallel^2$ in (32) and using (47), we obtain for the inversion of populations the equation

$$W = W_\infty\left[\left(1 - 2\frac{|\Omega_\perp|_m}{\Omega_{\parallel m}}\right)\operatorname{sech}\left(\frac{t - z/v}{\tau_p}\right)\right]. \qquad (49)$$

By virtue of the condition $||\Omega_\perp|_m^2 \ll \Omega_{\parallel m}^2$, the initial inversion, as is seen from (49), undergoes only slight changes as the transverse-longitudinal pulse passes. In the limit under consideration, the transverse component, which causes the quantum transition, is weaker than the longitudinal component, which takes the transition off the resonance with the carrier frequency of the pulse transverse component. This fact explains the weakening of the excitation degree of the quantum system as the ratio $|\Omega_\perp|/\Omega_\parallel$ decreases.

Setting $a_\parallel = a_\perp = a$, we obtain an expression for the amplitude of the transverse component:

$$|\Omega_\perp|_m = \frac{F_{44}}{F_{11}}\sqrt{\frac{2\omega_0}{\tau_p}}.$$

Then, the condition $|\Omega_\perp|_m^2 \ll \Omega_{\parallel m}^2$, together with the applicability condition of the SVAP approximation ($\omega_0\tau_p \gg 1$) is written as

$$1 \ll \omega_0\tau_p \ll (2F_{11}/F_{44})^2. \qquad (50)$$

For paramagnetic inclusions of $Co^{2+}$ in the matrix of the ion crystal, the ratio $2F_{11}/F_{44}$ can reach as much as 10 [11]. Hence, $1 \ll \omega_0\tau_p \ll 100$. Taking into account the fact that the SVAP approximation becomes valid for $\omega_0\tau_p > 10$, we conclude that the soliton-like structure of type (47) can be discovered in experiments with acoustic pulses with the transverse component being quasi-monochromatic with the carrier frequency $\omega$ that is close to $\omega_0$. Notwithstanding the fact that $\Omega_\parallel^2 \gg |\Omega_\perp|^2$, the relative strains $|\mathcal{E}_\perp|$ and $\mathcal{E}_\parallel$ can be comparable in magnitude; in this case, the inequality $\Omega_\parallel^2 \gg |\Omega_\perp|^2$ holds at the expense of the inequality $F_{11}^2 \gg F_{44}^2$. Setting $F_{11} \sim 10^2$, $\omega_0\tau_p \approx 40$, and $g = 2$, we find that

$$\mathcal{E}_\parallel \approx \frac{2g}{F_{11}\omega_0\tau_p} \approx 10^{-3},$$

$$|\mathcal{E}_\perp| \approx \frac{\sqrt{2}g}{F_{11}\sqrt{\omega_0\tau_p}} \approx 5 \times 10^{-3},$$

which corresponds to the intensity $I \approx \rho a_\perp^3 |\mathcal{E}_\perp|^2 \sim 10^5$ W/cm$^2$ (here, we assume that $\rho = 2$ g/cm$^3$ and $a_\perp \approx 3 \times 10^5$ cm/s). Thus, to excite a soliton of form (47), the intensity of hypersonic pulses must be by about two orders of magnitude greater than the intensity of the soliton (43). It was noted above that the soliton-like solution (43) is only a slight perturbation of the soliton of the sinus–Gordon equation obtained in [4]. From this point of view, solution (47) is a fundamentally new one; it necessarily contains a longitudinal acoustic component. In this connection, we note that this solution can be also obtained directly from (34), (38), and (39) if we set $\Omega \approx |\Omega_\parallel|$ in those equations.

It follows from (43), (47), (37), and (17) that the carrier frequency of the transverse component in the neighborhood of the soliton maximum is less than the corresponding frequency on peripheral parts. If the medium was in thermodynamic equilibrium before the pulse action, then $\alpha_\parallel, \alpha_\perp > 0$, and the pulse speed $v < a_\perp, a_\parallel$. Hence, it follows from (40) that $\alpha$, $\beta > 0$ and $\Omega_\parallel$, $d\phi/d\xi < 0$. Thus, for arbitrary ratios $|\Omega_\perp|/\Omega_\parallel$, the frequency of the transverse component of the stationary traveling pulse in the equilibrium medium decreases from its periphery to the center, where the influence of the longitudinal component is the largest (see Fig. 1).

In the case of the equilibrium medium, for $\Omega_\parallel < 0$, according to (15) (or (21)), the longitudinal component of the pulse reduces the frequency of the quantum transition, which takes the spin subsystem off the resonance with the transverse component. Trying to remain in the

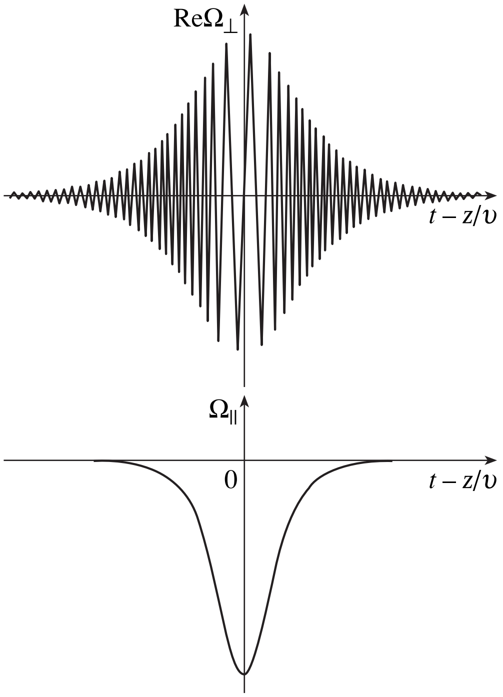

**Fig. 1.** Schematic view of the two-component soliton of the self-induced acoustic transparency propagating along the magnetic field.

resonance state, the spin subsystem modulates the frequency of the transverse component in an appropriate way. This fact can be viewed at as a manifestation of the Le Chatelier–Brown principle [22].

In the intermediate case $|\Omega_\perp| \sim \Omega_\|$, system (41), (42) resists analytical analysis; thus, it was investigated numerically. The results of the numerical analysis for various values of the ratio

$$\frac{\Phi_m^2}{\sigma_m} \equiv \left(\frac{\Omega_{\|m}}{|\Omega_\perp|_m}\right)^2$$

are presented in Fig. 2. For the two limiting cases (Figs. 2a and 2c), the numerical dependences $\Omega_\|(\xi)$, $|\Omega_\perp(\xi)|$, and $W(\xi)$ are in good agreement with the corresponding analytic expressions (43), (46) and (47), (49) presented above. The intermediate case, when $\Omega_{\|m}$ and $||\Omega_\perp|_m$ are close to each other in magnitude, is illustrated in Fig. 2b. As expected, the degree of excitation of the quantum system lies between the two limits (Figs. 2a and 2c) mentioned above. In all the cases, the longitudinal component of the traveling pulse is localized stronger than the transverse one, and the total area of the pulse is

$$\theta_\infty = \int_{-\infty}^{+\infty} \Omega dt = 2\pi.$$

This is best seen from Eq. (32) for the inversion and from the plots presented in Fig. 2. The values of $W$ before sending the pulse ($\theta = 0$) and after its passage ($\theta = 2\pi$) are equal to $W_\infty$. It seems important that the contribution to the area of the pulse is made both by the resonance transverse and the low-frequency longitudinal components of the acoustic pulse. On the other hand, it is seen from (32) that the purely longitudinal pulse ($\Omega_\perp = 0$) does not cause quantum transitions (and, thus, does not change inversion) in the Faraday geometry, which was also mentioned above.

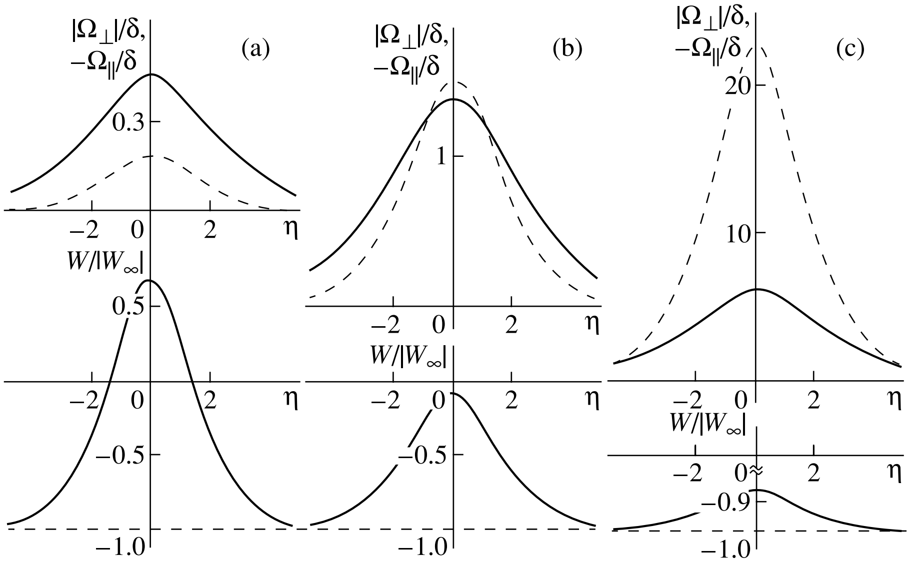

**Fig. 2.** Results of the numerical solution of system (41), (42) upon returning to the original variables $\Omega_\|$ and $|\Omega_\perp|$: $(\Omega_\|/|\Omega_\perp|)^2 = 0.18$ (a), 1.17 (b), and 13 (c); $\eta = 2(t - z/v)/\tau_p$. On the upper plots, the dependences $|\Omega_\perp(\eta)|$ are shown in solid lines and $\Omega_\|(\eta)$ by dashed ones. The inversion $W$ was calculated by formula (32).

## 5. CONCLUSIONS

In this paper, we investigated the structure of transverse–longitudinal soliton-like acoustic pulses interacting with paramagnetic Kramers doublets in the ASIT mode that propagate along an external magnetic field (the Faraday geometry). In the process, the transverse component of the pulse causes resonance quantum transitions inside the doublets, and the role of the longitudinal component is reduced to the dynamic shift in the frequency of those transitions. As a result, the quasi-monochromatic transverse component is taken off the resonance state by the low-frequency longitudinal component. As a result, the degree of excitation of the quantum system (Kramers doublets) decreases. In addition, the existence of the longitudinal component of the elastic field results in a modulation of the carrier frequency of the pulse's transverse component—the frequency is less in the vicinity of the soliton's maximum than it is in peripheral regions. Thus, the long-wave (longitudinal) and short-wave (transverse) components interact nonlinearly through the quantum system. This interaction is most effective when the linear speeds of the longitudinal and transverse sound waves are equal. Then, the system of integro-differential equations (35), (36) holds, which describes the long-short-wave resonance in the strong nonlinearity mode. In this sense, it is different from the Zakharov system [9], Yadjima–Oikawa system [10], and their generalizations [8, 23, 24], where the nonlinearity is weak (expanded in powers of the amplitudes of both components). Thus, in our opinion, further mathematical investigation of system (35), (36) is of interest. The asymptotic soliton-like solution of form (47) ($\Omega_{\parallel}^2 \gg |\Omega_{\perp}|^2$) is fundamentally new and cannot be reduced to known [4] purely transverse solitons of the sinus-Gordon equation.

We considered ions with the effective spin $S = 1/2$ as quantum paramagnetic inclusions. This model is extremely simple and is characterized by ease of interpretation of the results obtained. However, it is well known that paramagnetic ions of the iron group possessing the effective spin $S = 1$ exhibit a much stronger dynamic coupling with the oscillations of the crystal lattice. In this case, the Hamiltonian of the spin–phonon interaction is quadratic in the spin operators. As a result, the mathematical analysis becomes much more difficult. However, the model with $S = 1$ is preferable from the viewpoint of possible experiments. The system of paramagnetic impurities with $S = 1$ is also interesting for another reason. As an acoustic pulse propagates perpendicularly to $\mathbf{B}$ (the Focht geometry), both components of the pulse (the longitudinal and the transverse ones) cause quantum transitions inside the Zeeman triplet. More precisely, the transverse component excites transitions at the frequencies $\omega_0$ and $2\omega_0$, whereas the longitudinal one at the frequency $2\omega_0$ [25]. In this case, both components can be represented in the form of a slowly varying envelope, and the long-short-wave mode is replaced by a two-frequency ASIT. In our opinion, the corresponding investigation is of interest.

The method for obtaining asymptotic solutions to the system of constitutive Eqs. (22) developed in this paper, which is based on formulas (24) and (28), is also valid for microobjects with an arbitrary number of quantum levels. Hence, the application of this approach in the case of paramagnetic impurities with higher effective spins seems promising. At the same time, the method described is inapplicable to weaker pulses, when the inhomogeneous broadening must be taken into account. In this case, new approaches to solving constitutive equations containing both the longitudinal and transverse components of the pulse are required. Taking into account the inhomogeneous broadening is essential for deriving the areas theorem [2]. In this connection, it is important to find a modification of this theorem for two-component pulses compared to the cases of purely longitudinal or transverse solitons.

## ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research, project no. 00-02-17436a.

## REFERENCES

1. S. L. McCall and E. L. Hahn, Phys. Rev. Lett. **18**, 908 (1967).
2. N. S. Shiren, Phys. Rev. B **2**, 2471 (1970).
3. V. V. Samartsev, B. P. Smolyakov, and R. Z. Sharipov, Pis'ma Zh. Éksp. Teor. Fiz. **20**, 644 (1974) [JETP Lett. **20**, 296 (1974)].
4. G. A. Denisenko, Zh. Éksp. Teor. Fiz. **60**, 2269 (1971) [Sov. Phys. JETP **33**, 1220 (1971)].
5. O. Anderson, *Physical Acoustics: Principles and Methods*, Vol. 3, Part B: *Lattice Dynamics*, Ed. by W. P. Mason (Academic, New York, 1965; Mir, Moscow, 1968), p. 62.
6. C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976; Fizmatlit, Moscow, 1963).
7. C. Zener, Phys. Rev. **71**, 323 (1947).
8. S. V. Sazonov, Zh. Éksp. Teor. Fiz. **118**, 20 (2000) [JETP **91**, 16 (2000)].
9. V. E. Zakharov, Zh. Éksp. Teor. Fiz. **62**, 1745 (1972) [Sov. Phys. JETP **35**, 908 (1972)].
10. N. Yadjima and M. Oikawa, Prog. Theor. Phys. **56**, 1719 (1976).
11. J. W. Tucker and V. W. Rampton, *Microwave Ultrasonics in Solid State Physics* (North Holland, Amsterdam, 1972; Mir, Moscow, 1975).
12. V. A. Golenishchev-Kutuzov, V. V. Samartsev, N. K. Solovarov, and B. M. Khabibulin, *Magnetic Quantum Acoustics* (Nauka, Moscow, 1977).
13. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 7: *Theory of Elasticity* (Nauka, Moscow, 1987; Pergamon, New York, 1986).
14. S. A. Altshuler and B. M. Kozyrev, *Electron Paramagnetic Resonance in Compounds of Transition Elements* (Nauka, Moscow, 1972; Halsted, New York, 1975).

15. S. V. Sazonov, J. Phys.: Condens. Matter **6**, 6295 (1994).
16. N. N. Moiseev, *Asymptotic Methods of Nonlinear Mechanics* (Nauka, Moscow, 1981).
17. A. Yu. Parkhomenko and S. V. Sazonov, Kvantovaya Élektron. (Moscow) **27**, 139 (1999).
18. F. R. Gantmacher, *The Theory of Matrices* (Nauka, Moscow, 1966; Chelsea, New York, 1959).
19. U. Kh. Kopvillem and S. V. Prants, *Polarization Echo* (Nauka, Moscow, 1985).
20. S. V. Sazonov and E. V. Trifonov, J. Phys. B **27**, L7 (1994).
21. M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *The Theory of Waves* (Nauka, Moscow, 1990).
22. R. Kubo, *Thermodynamics. An Advanced Course with Problems and Solutions* (North-Holland, Amsterdam, 1968; Mir, Moscow, 1970).
23. E. S. Benilov and S. P. Burtzev, Phys. Lett. A **98A**, 256 (1983).
24. V. G. Makhan’kov, Yu. P. Rybakov, and V. I. Sanyuk, Usp. Fiz. Nauk **164** (2), 121 (1994) [Phys. Usp. **37**, 113 (1994)].
25. V. Yu. Man’kov and S. V. Sazonov, Fiz. Tverd. Tela (St. Petersburg) **41**, 623 (1999) [Phys. Solid State **41**, 560 (1999)].

*Translated by A. Klimontovich*

Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 247–255.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 280–290.
Original Russian Text Copyright © 2001 by Naumov, Sidorov-Biryukov, Giammanco, Fedotov, Marsili, Ruffini, Kolevatova, Zheltikov.

# Four-Wave Mixing of Picosecond Pulses in Hollow Fibers: Phase Matching and the Influence of High-Order Waveguide Modes

**A. N. Naumov[a], D. A. Sidorov-Biryukov[a], F. Giammanco[b], A. B. Fedotov[a], P. Marsili[b], A. Ruffini[b], O. A. Kolevatova[a], and A. M. Zheltikov[a, *]**

[a]*International Laser Center, Physics Faculty, Moscow State University, Moscow, 119899 Russia*
**e-mail: zheltikov@top.phys.msu.su*

[b]*Dipartimento di Fisica e Unitá INFM – Università di Pisa, via F. Buonarroti 2, 56127 Pisa, Italy*

Received January 30, 2001

**Abstract**—The processes of third-harmonic and difference-frequency generation through the four-wave mixing of picosecond pulses in gas-filled hollow fibers are experimentally studied. Due to the improvement of phase-matching conditions with an appropriate choice of the gas pressure and optimal parameters of the hollow fiber, we were able to use hollow fibers with a large length (up to 30 cm) for difference-frequency generation, which resulted in a considerable increase in the power of the difference-frequency signal at the output of the fiber. Our experimental data reveal a considerable influence of high-order waveguide modes on four-wave mixing processes in a hollow fiber. It is shown that the waveguide regime of nonlinear optical interactions implemented in hollow fibers removes the limitations on the efficiency of third-harmonic and sum-frequency generation, which are characteristic of the tight-focusing regime in media with normal dispersion and which are due to the geometric phase shift arising in tightly focused light beams. *© 2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

Gas-filled hollow fibers have recently successfully solved several important problems of nonlinear optics and optics of ultrashort laser pulses. Nonlinear optical interactions in gas-filled hollow fibers are currently widely employed in the generation of extremely short light pulses [1, 2] and optical frequency conversion to the vacuum-ultraviolet and X-ray ranges (including the water-window region) through high-order harmonic generation [3–8]. Fibers of this type also offer much promise for improving the sensitivity and expanding the possibilities of coherent four-wave mixing spectroscopy [9, 10].

An important advantage of hollow fibers is associated with the fact that they permit the pump and the signal generated through a frequency-nondegenerate nonlinear optical process to be phase-matched. With a careful choice of the parameters of a hollow fiber, the gas pressure, and the excitation of appropriate waveguide modes, the phase mismatch related to the gas dispersion can be compensated for by the waveguide component of the phase mismatch [3, 8, 9, 11]. When these conditions are satisfied, the energy of the nonlinear signal can be considerably increased by using longer hollow fibers. The parameters of short pulses of short-wavelength radiation generated through nonlinear optical interactions in gas-filled hollow fibers can be controlled due to cross-phase modulation [12–14].

This paper is devoted to the investigation of the properties of nonlinear optical interactions in gas-filled hollow fibers that are, in our opinion, of considerable methodological interest and that open new avenues for numerous practical applications of hollow fibers in nonlinear optics, optics of ultrashort pulses, and nonlinear spectroscopy. In particular, one of the most general properties of the waveguide regime of nonlinear optical interactions in hollow fibers is associated with a fact that the use of a hollow fiber allows high power densities of laser radiation typical of a tight-focusing regime to be achieved with an appropriate focusing of pump beams. The waveguide regime of radiation propagation under these conditions ensures the geometry of nonlinear optical processes that is characteristic of plane-wave interaction, thus allowing the efficiency of sum-frequency generation to be considerably improved relative to nonlinear optical interactions of tightly focused light beans in a medium with a normal dispersion.

Most of the nonlinear optical hollow-fiber experiments performed to date were performed with the use of high-intensity femtosecond pulses (the pioneering work by Miles *et al.* [9] is an exception, but their work does not deal with sum-frequency generation processes). The prohibition on third-harmonic generation (THG) and sum-frequency generation in a gas with an initially positive dispersion in these experiments may be removed due to the ionization of the gas (such effects were observed in numerous experiments) and due to the self-action of laser pulses (which was also observed experimentally, see [15, 16]).

The experiments presented in this paper provide direct evidence of the possibility of sum-frequency generation and optical frequency multiplication in media with a normal dispersion due to the use of the waveguide regime of nonlinear optical interactions. For this purpose, we employed picosecond pulses of moderate intensities (the maximum intensities of laser pulses in our experiments were on the order of $10^{11}$ W/cm$^2$). No third harmonic was generated when such laser beams were tightly focused in the gas in the absence of a fiber, which indicates that the perturbation of the gas medium and the pump beams themselves does not have a considerable influence on nonlinear optical processes. Such an approach allowed us to study the main properties of four-wave mixing (FWM) processes in the waveguide regime and to examine the ways to phase match FWM processes under these conditions. One of the important results of our study is the experimental demonstration of a considerable influence of high-order waveguide modes on FWM processes in hollow fibers. The investigation of FWM of picosecond pulses is also of considerable interest in the context of the possibility of using hollow fibers to improve the sensitivity of nonlinear optical techniques for gas-phase analysis. Picosecond pulses are often a reasonable choice for stationary spectroscopic techniques, which are widely employed for various practical applications and which may often impose certain limitations on the duration of laser pulses.

The plan of this paper is the following. In Section 2, we employ the slowly varying envelope approximation to derive expressions describing FWM processes in gas-filled hollow fibers including the phase mismatch and the influence of high-order waveguide modes. Section 3 describes the experimental technique and the procedure of measurements. The results of experiments are discussed in Section 4. Finally, the main conclusions will be briefly summarized in the last section of this paper.

## 2. THE THEORY OF FOUR-WAVE MIXING IN HOLLOW FIBERS

### *2.1. The Amplitude of the FWM Signal*

In this section, we will study the specific features of FWM processes in gas-filled hollow fibers taking into consideration the influence of phase-matching effects, optical losses of hollow-fiber modes, and high-order waveguide modes. We will consider FWM processes of third-harmonic and difference-frequency generation giving rise to a signal at the frequency of the third harmonic in accordance with the following FWM schemes:

$$3\omega = \omega + \omega + \omega \quad \text{and} \quad 3\omega = 2\omega + 2\omega - \omega,$$

where $\omega$ and $2\omega$ are the frequencies of pump pulses (fundamental radiation of the pump laser and its second harmonic). Processes of this type, as demonstrated by experiments [3], allow high efficiencies of nonlinear optical frequency conversion to be achieved by phase-matching the light pulses involved in FWM in a hollow fiber. The results of our experimental studies for such processes will be presented in Section 4 of this paper.

Suppose that fundamental radiation and its second harmonic (pump pulses) propagate along the $z$-axis of a hollow fiber with an inner radius $a$. We assume that the hollow fiber is filled with a gas with a cubic nonlinearity and a refractive index $n$. The dielectric constant of the cladding of the hollow fiber is assumed to be a real quantity meeting the condition $\varepsilon > n^2$. The fields of the pump and FWM pulses can be then represented as

$$E_1 = \frac{1}{2} f_1^q(\rho) A_0^q \times \exp[-i\omega t + (iK_1^q - \alpha_1^q/2)z] + \text{c.c.}, \tag{1}$$

$$E_2 = \frac{1}{2} \sum_l f_2^l(\rho) B_0^l \times \exp[-2i\omega t + (iK_2^l - \alpha_2^l/2)z] + \text{c.c.}, \tag{2}$$

$$E_3 = \frac{1}{2} f_3^m(\rho) C^m(z) \exp[-3i\omega t + iK_3^m z] + \text{c.c.}, \tag{3}$$

where $f_1^q(\rho)$, $f_2^l(\rho)$, and $f_3^m(\rho)$ are the transverse field distributions corresponding to the $EH_{1q}$, $EH_{1l}$, and $EH_{1m}$ hollow-fiber modes of fundamental radiation, the second harmonic, and the FWM pulse, respectively; $\rho$ is the distance from the axis of the hollow fiber; $A_0^q$ and $B_0^l$ are the amplitudes of the pulses of fundamental radiation and the second harmonic at the input of the fiber; $C^m(z)$ is the slowly varying amplitude of the FWM signal; $K_1^q$, $K_2^l$, and $K_3^m$ are the propagation constants of fundamental radiation, the second harmonic, and the FWM signal in the hollow fiber, respectively; and $\alpha_1^q$ and $\alpha_2^l$ are the attenuation coefficients for the $EH_{1q}$ waveguide mode at the fundamental frequency and the $EH_{1l}$ waveguide mode at the frequency of the second harmonic. Representing the field of the second harmonic in Eq. (2) as a sum of hollow-fiber modes, we extend our analysis to FWM processes where two of the four waves have equal frequencies $2\omega$, but different transverse field distributions corresponding to different waveguide modes $EH_{1l'}$ and $EH_{1l''}$ (i.e., $l = l', l''$).

We assume that each of the waves involved in the FWM process has a small attenuation coefficient and a wavelength much less than the fiber core radius $a$:

$$\frac{\omega a}{c} \gg 1, \tag{4}$$

$$\left| \frac{K_s^m c}{s\omega n(s\omega)} - 1 \right| \ll 1, \tag{5}$$

where $s = 1, 2, 3$ and $n(s\omega)$ is the refractive index of the gas filling the hollow fiber at the frequency $s\omega$. Once conditions of Eqs. (4) and (5) are met, we can employ the following approximate expression for the transverse field distribution in $EH_{1m}$ modes of a hollow fiber [17, 18]:

$$f_s^m(\rho) = J_0\left(\frac{u^m\rho}{a}\right), \tag{6}$$

where $J_0(x)$ is the zeroth-order Bessel function and $u^m$ is the eigenvalue for the $EH_{1m}$ mode. Propagation constants and attenuation coefficients in this case are written as [18]

$$K_s^m \approx \frac{s\omega n(s\omega)}{c}\left[1 - \frac{1}{2}\left(\frac{u^m c}{as\omega n(s\omega)}\right)^2\right], \tag{7}$$

$$\alpha_s^m \approx \frac{2}{an(s\omega)}\left(\frac{u^m c}{as\omega}\right)^2 \times \frac{\varepsilon(s\omega) + n^2(s\omega)}{2n^2(s\omega)(\varepsilon(s\omega) - n^2(s\omega))^{1/2}}. \tag{8}$$

Using a procedure similar to that described in [19], we arrive at the following equation for slowly varying envelope of the third harmonic in a lossy hollow fiber:

$$\frac{d}{dz}C_{THG}^m + \frac{\alpha_3^m}{2}C_{THG}^m = i\beta_{THG}^{mq}(A_0^q)^3\exp\left(-i\Delta k_{THG}^{mq}z - \frac{3\alpha_1^q z}{2}\right), \tag{9}$$

where $\alpha_3^m$ is the attenuation coefficient for the $EH_{1m}$ mode of the THG signal. The phase mismatch is then written as

$$\Delta k_{THG}^{mn} = K_3^m - 3K_1^q \approx \Delta k_{THG}^g + \Delta k'^{mq}_{THG}, \tag{10}$$

where

$$\Delta k^g = \frac{3\omega}{c}[n(3\omega) - n(\omega)], \tag{11}$$

$$\Delta k'^{mq}_{THG} = \frac{c}{2\omega a^2}\left[3(u^q)^2 - \frac{(u^m)^2}{3}\right] \tag{12}$$

are the components of the phase mismatch due to the dispersion of the gas and waveguide dispersion, respectively [the total phase mismatch can be represented as a sum of these two components in the case when $n(\omega)$, $n(3\omega) \approx 1$]. The nonlinear coefficient $\beta_{THG}^{mq}$ can be expressed in terms of the relevant nonlinear optical cubic susceptibility [14]:

$$\beta_{THG}^{mq} = \frac{9\pi\omega^2}{2K_3^m c^2}\chi_{THG}^{(3)} \times \frac{\iint f_3^m(\rho)[f_1^q(\rho)]^3\rho\, d\rho\, d\theta}{\iint [f_3^m(\rho)]^2\rho\, d\rho\, d\theta}, \tag{13}$$

where $\chi_{THG}^{(3)}$ is the third-order nonlinear optical susceptibility responsible for third-harmonic generation.

In contrast to the standard plane-wave approximation (see, e.g., [19]), Eq. (9) includes the influence of a waveguide through the propagation constants (7) and the nonlinear coefficient (13), which is normalized to include the transverse distributions of the pump and third-harmonic fields in the relevant waveguide modes. In particular, the phase mismatch, which appears in Eq. (9) and which determines the efficiency of THG, depends not only on the gas dispersion, but also on the dispersion of waveguide modes. This circumstance, as highlighted in a number of earlier papers [3, 8, 14], provides an opportunity to improve phase matching for a given combination of waveguide modes of pump and third-harmonic radiation.

Integrating Eq. (9), we derive the following expression for the amplitude of the $EH_{1m}$ mode of the third harmonic:

$$C_{THG}^m = i\beta_{THG}^{mq}(A_0^q)^3 \times \frac{\exp\left[-\frac{\alpha_3^m}{2}L\right] - \exp\left[\left(-\frac{3\alpha_1^q}{2} - i\Delta k_{THG}^{mq}\right)L\right]}{\frac{\alpha_3^m - 3\alpha_1^q}{2} - i\Delta k_{THG}^{mq}}, \tag{14}$$

where $L$ is the length of the gas-filled hollow fiber. In the limiting case of low losses and zero phase mismatch, Eq. (14) is reduced to

$$C_{THG}^m = i\beta_{THG}^{mq}(A_0^q)^3 L. \tag{15}$$

Using Eq. (15), we can obtain the following estimate for the power of the third-harmonic signal:

$$P_{THG} \sim P_1^3\frac{L^2}{a^4}, \tag{16}$$

where $P_1$ is the power of fundamental radiation.

In the case of a difference-frequency generation (DFG) process $3\omega = 2\omega + 2\omega - \omega$, involving the $EH_{1q}$ hollow-fiber mode of fundamental radiation and $EH_{1l'}$ and $EH_{1l''}$ modes of the second harmonic, generating the $EH_{1m}$ mode of the DFG signal at the frequency of the third harmonic in a lossy hollow fiber, the slowly varying envelope of the DFG signal is governed by the following equation:

$$\frac{d}{dz}C_{DFG}^{m} + \frac{\alpha_3^m}{2}C_{DFG}^{m} = i\beta_{DFG}^{ml'l''q}A_0^{q*}B_0^{l'}B_0^{l''} \times \exp\left(-i\Delta k_{DFG}^{ml'l''q}z - \frac{\alpha_1^q + \alpha_2^{l'} + \alpha_2^{l''}}{2}z\right). \tag{17}$$

Here, the phase mismatch includes the dispersion of waveguide modes and can be represented as

$$\begin{aligned}\Delta k_{DFG}^{ml'l''q} &= K_3^m - K_2^{l'} - K_2^{l''} + K_1^q \\ &\approx \Delta k_{DFG}^g + \Delta k'^{ml'l''q}_{DFG},\end{aligned} \tag{18}$$

where

$$\Delta k_{DFG}^g = \frac{\omega}{c}[3n(3\omega) - 4n(2\omega) + n(\omega)], \tag{19}$$

$$\begin{aligned}&\Delta k'^{ml'l''q}_{DFG} \\ &= \frac{c}{2\omega a^2}\left[\frac{(u^{l'})^2}{2} + \frac{(u^{l''})^2}{2} - \frac{(u^m)^2}{3} - (u^q)^2\right]\end{aligned} \tag{20}$$

are the components of the phase mismatch due to the dispersion of the gas and waveguide modes, respectively. The nonlinear coefficient can be expressed in terms of the relevant nonlinear optical cubic susceptibility [14]:

$$\beta_{DFG}^{ml'l''q} = \frac{27\pi\omega^2}{K_3^m c^2}\chi_{DFG}^{(3)} \times \frac{\iint f_3^m(\rho) f_2^{l'}(\rho) f_2^{l''}(\rho)\rho d\rho d\theta}{\iint [f_3^m(\rho)]^2 \rho d\rho d\theta}, \tag{21}$$

where $\chi_{DFG}^{(3)}$ is the third-order nonlinear optical susceptibility responsible for difference-frequency generation.

Integrating Eq. (21), we derive the following expression for the amplitude of the DFG signal excited in the $EH_{1m}$ mode:

$$C_{DFG}^{m} = i\beta_{DFG}^{ml'l''q}A_0^{q*}B_0^{l'}B_0^{l''} \times \frac{\exp\left[-\frac{\alpha_3^m}{2}L\right] - \exp\left[\left(-\frac{\alpha_1^q + \alpha_2^{l'} + \alpha_2^{l''}}{2} - i\Delta k_{DFG}^{ml'l''q}\right)L\right]}{\frac{\alpha_3^m - \alpha_1^q - \alpha_2^{l'} - \alpha_2^{l''}}{2} - i\Delta k_{DFG}^{ml'l''q}}. \tag{22}$$

In the limiting case of low losses and zero phase mismatch, Eq. (22) is reduced to

$$C_{DFG}^{m} = i\beta_{DFG}^{ml'l''q}A_0^{q*}B_0^{l'}B_0^{l''}L. \tag{23}$$

Using Eq. (23), we derive the following estimate for the power of the DFG signal:

$$P_{DFG} \sim P_1 P_2^2 \frac{L^2}{a^4}, \tag{24}$$

where $P_2$ is the power of pump radiation at the frequency of the second harmonic.

Expressions (16) and (24) describe the dependence of the FWM signal power on geometrical sizes of a hollow fiber. In the following section, we will use these formulas to analyze the physical factors allowing the efficiency of four-wave mixing of short laser pulses to be increased in gas-filled hollow fibers.

### *2.2. Improving the Efficiency of Four-Wave Mixing in the Waveguide Regime*

To illustrate how the efficiency of FWM processes can be improved by using hollow fibers, it would be instructive to consider the basic formulas of the elementary theory of four-wave mixing. Expressions for the powers of signals produced through FWM processes of third-harmonic and difference-frequency generation (generally, $\omega_{DF} = 2\omega_2 - \omega_1$) can be found in many textbooks on nonlinear optics [19, 20]. In particular, in the regime of loose focusing, when the condition

$$b \gg L \tag{25}$$

is satisfied, where $b$ is the confocal parameter, these expressions can be written as

$$P_{THG} \sim P_1^3 \frac{L^2}{b^2} \frac{\sin^2[(\Delta k + 4/b)L/2]}{[(\Delta k + 4/b)L/2]^2} \tag{26}$$

in the case of third-harmonic generation,

$$P_{DFG} \sim P_1 P_2^2 \frac{L^2}{b^2} \frac{\sin^2[\Delta k L/2]}{[\Delta k L/2]^2} \tag{27}$$

in the case of difference-frequency generation. Here, we used the following notations: $P_1$ and $P_2$ are the powers of the pump waves and $\Delta k$ is the phase mismatch for the corresponding FWM process.

Let us consider in greater detail the enhancement of the efficiency of FWM processes in hollow fibers with respect to the geometry of tight focusing due to the increase in the interaction length of light beams attainable with hollow fibers. Physically, a hollow fiber enhances FWM processes since it allows light intensities typical of the tight-focusing regime to be achieved, simultaneously letting these beams interact in a nearly plane-wave regime. Since the intensity of the FWM signal is proportional to the intensities of the pump beams, the power of the FWM signal can be increased by decreasing the diameter of a hollow fiber and keeping the powers of pump beams constant as long as the phase-matching conditions are satisfied and the losses of the waves interacting in the fiber are low. The role of a hollow fiber is thus to ensure the regime of interaction of collimated beams [see Eqs. (26), (27) and (16), (24)] for light beams having intensities typical of the regime of tight focusing, simultaneously providing large interaction lengths for these beams and improving the phase matching.

Figure 1 displays the phase mismatch calculated with the use of Eqs. (10)–(12) and (18)–(20) for THG and DFG processes involving different waveguide modes as a function of the inner radius of a hollow fiber filled with air at standard atmospheric pressure under normal conditions. As can be seen from these plots, the phase mismatch for the DFG process involving fundamental waveguide modes of pump and signal radiation can be completely compensated with an appropriate choice of the hollow-fiber inner radius. In the case of third-harmonic generation in the field of the fundamental mode of pump radiation, phase matching can be achieved only for high-order modes of the hollow fiber. Generally, the phase-matching problems under these conditions can be solved by adjusting the gas pressure, choosing optimal parameters of the hollow fiber, and excitation of appropriate waveguide modes [3, 8, 14].

Thus, there are two natural ways of increasing the efficiency of nonlinear optical processes in hollow fibers: (i) increasing the fiber length and (ii) reducing the inner radius of the fiber. The enhancement factor cannot be increased infinitely, of course. The increase of the fiber length is limited by optical losses of hollow-fiber modes, while the decrease of the inner radius requires a tighter focusing of the pump beam, eventually leading to the breakdown of the gas filling the fiber. The use of shorter pulses under these conditions allows further improvement in the efficiency of nonlinear optical processes due to the increase in the breakdown threshold of the gas.

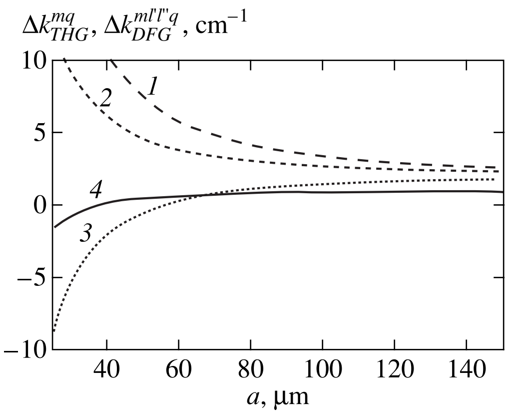

**Fig. 1.** Phase mismatches for (*1*–*3*) third-harmonic generation and (*4*) difference-frequency generation in a hollow fiber filled with air at standard atmospheric pressure under normal conditions as functions of the inner radius of the hollow fiber $a$. The transverse distribution of the pump field corresponds to the $EH_{11}$ waveguide mode. The third harmonic is produced in the $EH_{11}$ (*1*), $EH_{12}$ (*2*), and $EH_{13}$ (*3*) waveguide modes. The DFG signal is generated in the $EH_{11}$ waveguide mode.

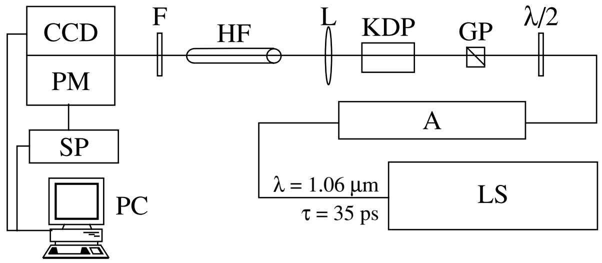

**Fig. 2.** Diagram of the experimental setup for studying FWM processes in gas-filled hollow fibers based on a passively mode-locked picosecond laser system: LS, picosecond laser system; A, amplification stages; GP, Glan prism; L, achromatic lens; HF, hollow fiber; F, bandpass filter; PM, photomultiplier; CCD, CCD camera; SP, signal-processing unit; and PC, personal computer.

### *2.3. Removing the Prohibition on Third-Harmonic Generation*

Hollow fibers may play an even more important role in the case of sum-frequency and third-harmonic generation. In media with a normal dispersion, such FWM processes are characterized by a low efficiency in the tight-focusing regime due to an additional phase shift of a focused beam with respect to a plane wave. This geometric phase shift around the axis of a Gaussian beam can be written as [20]

$$\Delta\varphi = -\arctan\frac{2(z-z_0)}{b}, \qquad (28)$$

where $z_0$ is the coordinate of the beam waist.

In accordance with Eq. (28), the phase shift between the field of the third harmonic and the nonlinear polarization responsible for THG tends to $\pm\pi$ as $z \longrightarrow \pm\infty$. Therefore, no third harmonic can be observed at the output of the medium in this regime because of the destructive interference of the pump and third-harmonic fields generated before and after the focus.

The situation radically changes in the case of hollow fibers, where nonlinear optical interactions occur in the regime of collimated light beams, giving rise to no $\pm\pi$ phase shift between the signal field and the relevant polarization of the medium. This allows efficient third-harmonic generation. In Section 4, we will present the experimental data confirming this conclusion.

## 3. EXPERIMENTAL SETUP

The experimental setup for studying FWM processes in gas-filled hollow fibers (Fig. 2) consisted of a picosecond laser system, which generated pumping radiation at the wavelengths of 1.06 and 0.53 μm, a hollow fiber, and a detection system based on a photodetector, photomultiplier, and a CCD camera. The pico-

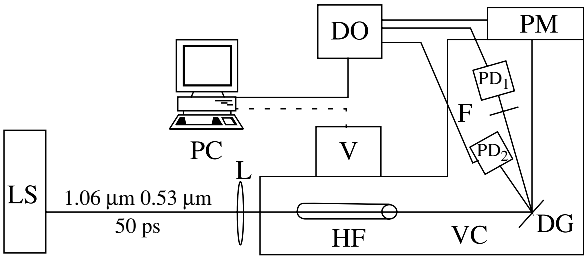

**Fig. 3.** Diagram of the experimental setup for studying the influence of the gas pressure on FWM processes in a gas-filled hollow fiber: LS, picosecond laser system; L, achromatic lens; VC, vacuum chamber; HF, hollow fiber; DG, diffraction grating; PM, photomultiplier; F, filter blocking pump beams; PD$_1$ and PD$_2$, photodiodes measuring the energy of fundamental radiation and the second harmonic, respectively; DO, digital oscilloscope; V, computer-controlled valve for gas delivery.

second system included a passively mode-locked Nd:YAG master oscillator with a negative-feedback-controlled cavity *Q*-factor [21], a single-pulse selection unit, and amplifying stages. Passive mode locking in the master oscillator was implemented with the use of a saturable absorber film, which was placed in front of the rear cavity mirror and which made it possible to generate laser pulses with a duration of 35 ps. Negative feedback was introduced by inserting an electro-optical switch controlled with a fast-response photomultiplier inside the cavity. An optical signal served as an input for the fast-response photomultiplier. This feedback loop considerably improved the stability of parameters of laser pulses [22], providing an opportunity to generate trains of picosecond light pulses with a duration of the envelope on the order of 30–40 μs. As the regime of stationary lasing was established in the master oscillator, the negative feedback loop was switched off, and a short train of highly stable picosecond pulses with a duration of the envelope of about 100 ns and an energy of approximately 1.5 mJ was generated.

An electro-optical switch was used to separate a single pulse from this train. The energy of a single 35-ps laser pulse thus selected ranged from 30 to 40 μJ. The single-pulse selection unit also served as an optical decoupler, suppressing the parasitic feedback between amplifying stages and the master oscillator and preventing radiation reflected from optical elements of the amplification system from influencing the formation of trains of pulses in the master oscillator. The further details of our picosecond laser system can be found elsewhere [21, 22].

A single pulse of 1.06-μm radiation passes through three amplifying stages. The energy of the laser pulse at the output of the third stage may reach 50 mJ. The spatial distribution of intensity in such a laser beam is close to that characteristic of the Gaussian mode. This radiation was used as a pump beam in the THG scheme and one of the pump beams in sum- and difference-frequency generation. A KDP crystal was used to produce the second harmonic of Nd:YAG laser radiation for two-color experiments. A spherical lens was employed to couple pump beams into the fiber (Fig. 2). We used commercially available hollow fibers with inner diameters of 70, 100, 127, 152, and 203 μm in our experiments. The lengths of the fibers were varied from 1 to 30 cm. The attenuation coefficients of 1.06-μm radiation in these fibers were estimated as 0.6, 0.2, 0.1, 0.06, and 0.04 dB/cm, respectively. The signals produced through third-harmonic, sum-frequency, and difference-frequency generation in these fibers were selected with a monochromator and bandpass filters and were detected with a photomultiplier. A CCD camera was used to investigate the spatial profiles of light beams coming out of the fiber.

To measure the gas-pressure dependences of nonlinear optical signals, we used an experimental setup (Fig. 3) consisting of a picosecond laser, a vacuum chamber with a hollow fiber inside, and a detection system based on a photomultiplier. The Nd:YAG picosecond laser generated 50-ps pump pulses at 1.06 μm and 0.53 μm. The maximum energy of 1.06-μm radiation reached 100 mJ. A KDP crystal was used to produce the second harmonic of fundamental radiation. An achromatic lens was used to couple the pump laser beams into a hollow fiber. Two photodiodes were used to monitor the energies of both of these laser beams transmitted through the fiber. The energies of fundamental radiation and the second harmonic in these experiments were equal to 1 and 0.1 mJ, respectively. The signal produced through an FWM process in a hollow fiber was detected with a photomultiplier and was then processed and displayed with a digital oscilloscope. The result of averaging over 30 FWM pulses was stored in a personal computer.

## 4. RESULTS AND DISCUSSION

The results of our experiments fully justify our expectations that the use of hollow fibers allows the efficiency of FWM processes to be improved, the prohibition on THG to be removed, and phase-matching conditions in FWM processes to be improved. Our experiments also revealed a noticeable influence of high-order waveguide modes on FWM processes in gas-filled hollow fibers. The use of picosecond pulses allowed us to couple pump beams with a sufficiently high intensity into a hollow fiber without any noticeable perturbation of the gas medium in the fiber or the self-action of laser pulses.

Figure 4 presents the powers of third-harmonic and the difference-frequency signals as functions of the pump energy. As can be seen from the data presented in Fig. 4a, the power of the third harmonic can be approximated with high accuracy with a cubic function of the power of fundamental radiation. The power of the DFG signal at the frequency $\omega_{DF} = 2\omega_2 - \omega_1$ (where $\omega_1$ is the frequency of fundamental radiation of the Nd:YAG laser and $\omega_2 = 2\omega_1$ is the frequency of the second harmonic of this laser) is linear in the fundamental radia-

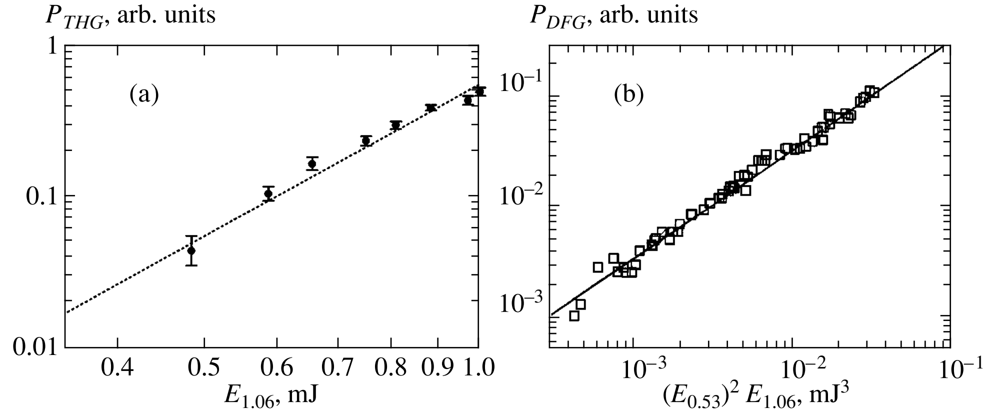

**Fig. 4.** The powers of (a) the third-harmonic and (b) difference-frequency signals produced in a hollow fiber filled with atmospheric-pressure air under normal conditions as functions of the energy of fundamental radiation $E_{1.06}$ (a) and the parameter $(E_{0.53})^2E_{1.06}$ (b) ($E_{0.53}$ is the energy of second-harmonic radiation). The inner diameter of the fiber is (a) 127 and (b) 152 μm. The fiber length is equal to (a) 4 and (b) 20 cm. Cubic (a) and linear (b) dependences are given for the convenience of comparison.

tion power and quadratic in the power of the second harmonic of the Nd:YAG laser (see Fig. 4b). These results indicate that nonlinear optical interactions in our experiments occur in a weak-field regime, and effects related to the ionization of the medium and the self-action of pump pulses do not exert a considerable influence on nonlinear optical processes. This conclusion is also supported by the fact that no THG signal was observed in our experiments in the tight-focusing regime with the same focusing parameters as in hollow-fiber experiments, but in the absence of a hollow fiber, until a plasma was produced due to the gas breakdown. The THG signal was easily detectable under conditions of ionization of the medium and ionization-induced self-action of pump pulses [15, 16].

The influence of phase-matching effects on FWM processes in hollow fibers is illustrated by the experimental data presented in Fig. 5. In particular, these experimental data indicate that, for gases whose dispersion properties are similar within the studied frequency range, the pressure dependences of the DFG signal power have much in common. Specifically, the pressure dependences of the DFG signal for argon and nitrogen display qualitatively similar tendencies (Fig. 5). At the same time, the pressure dependence of the DFG signal for carbon dioxide qualitatively differs from similar dependences for argon and nitrogen. This is due to considerable differences in dispersion properties of carbon dioxide and dispersion properties of argon and nitrogen. In particular, the phase mismatch for the DFG process at the atmospheric pressure of carbon dioxide is estimated as $\Delta k^g_{DFG} = 1.8$ cm$^{-1}$, which noticeably differs from the phase mismatch corresponding to the atmospheric pressure of nitrogen or argon ($\Delta k^g_{DFG} = 1$ cm$^{-1}$).

Comparison of the experimental data presented in Figs. 6 and 7 with the results of calculations performed with the use of Eqs. (18)–(20) and (22) reveals a noticeable role of high-order waveguide modes in nonlinear optical processes in hollow fibers. It is instructive in this context to consider in greater detail the results obtained for the DFG process $\omega_{DF} = 2\omega_2 - \omega_1$ in an argon-filled hollow fiber with a length of 17.4 cm and an inner diameter $a = 100$ μm (Fig. 6). A satisfactory agreement between the experimental data (squares) and theoretical predictions (the solid line) is achieved when effects related to high-order waveguide modes are included in the analysis. In particular, a satisfactory agreement between the experimental data in Fig. 6 and the results of calculations performed with the use of Eqs. (18)–(20) and (22) was achieved when not only the DFG process occurring in the fundamental waveguide mode (i.e., the DFG process involving the $EH_{11}$ hollow-fiber modes of fundamental radiation, second harmonic, and the DFG signal) was included in the calcu-

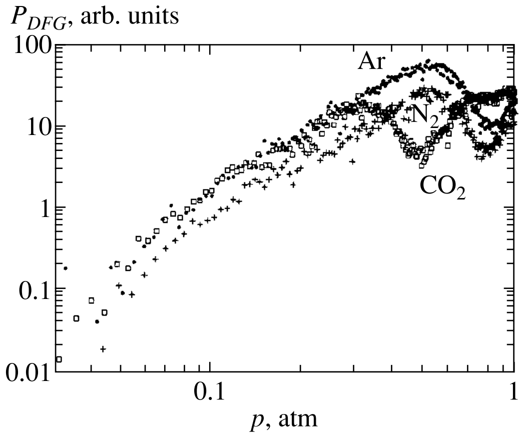

**Fig. 5.** The power of the difference-frequency signal produced in a hollow fiber filled with different gases [(●) Ar, (+) $N_2$, and (□) $CO_2$] as a function of the gas pressure $p$. The fiber length is 19.3 cm. The inner diameter of the fiber is 127 μm.

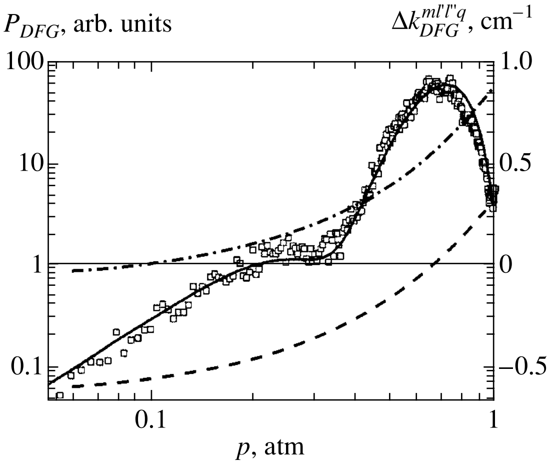

**Fig. 6.** The power of the DFG signal (the dots show the experimental data, and the solid line represents the results of calculations) and the phase mismatch for the DFG process in an argon-filled hollow fiber as functions of the argon pressure $p$. The dashed line shows the phase mismatch for the DFG process occurring in the fundamental waveguide mode $EH_{11}$. The dash–dotted line represents the phase mismatch for the DFG process involving the $EH_{12}$ mode of fundamental radiation, $EH_{11}$ and $EH_{13}$ modes of the second harmonic, and the $EH_{12}$ mode of the DFG signal. The length of the hollow fiber is 17.4 cm, and the inner diameter is 100 μm.

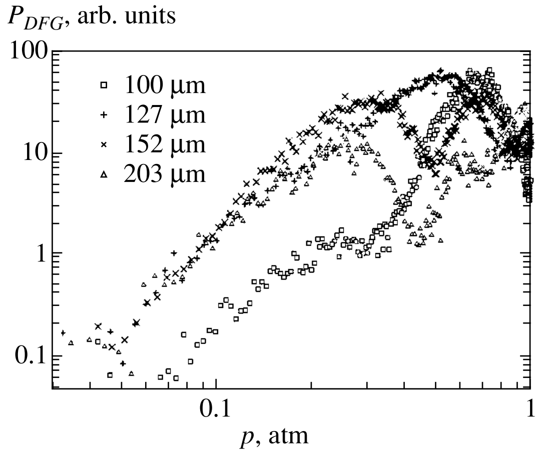

**Fig. 7.** The power of the DFG signal in argon-filled hollow fibers with different inner diameters as a function of the argon pressure $p$. The length of hollow fibers in all the experiments was approximately equal to 20 cm. The inner diameter of the hollow fiber was (□) 100, (+) 127, (×) 152, and (△) 203 μm.

lations; but also the DFG process involving the $EH_{12}$ mode of fundamental radiation, $EH_{11}$ and $EH_{13}$ modes of the second harmonic, and the $EH_{12}$ mode of the DFG signal was included. The maximum of the DFG signal around an argon pressure of about 0.7 atm is observed within the pressure range where the DFG process in the fundamental waveguide mode is phase matched (the dashed line in Fig. 6 shows the phase mismatch for this process). At lower pressures, difference-frequency generation through the FWM interaction of the $EH_{12}$ mode of fundamental radiation, $EH_{11}$ and $EH_{13}$ modes of the second harmonic, and the $EH_{12}$ mode of the DFG signal begins to play a more important role (the phase mismatch for this process is shown by the dash–dot line in Fig. 6). Thus, high-order waveguide modes of a hollow fiber may have a noticeable influence on FWM processes.

Figure 7 presents the experimental dependences of the DFG signal produced in argon-filled hollow fibers with different inner diameters and a length of ≈20 cm on the gas pressure $p$. As can be seen from Fig. 7, with the increase in the inner diameter of the fiber, the maximum of the DFG signal related to the FWM interaction of the pump and DFG beams in the fundamental waveguide mode is shifted toward lower pressures, tending in the case of large inner diameters to the limiting value corresponding to the FWM process in collimated beams. In this limiting case, the maximum power of the DFG signal is achieved, in accordance with Eq. (27), at the gas pressure in the fiber equal to 0.16 atm (at this pressure, the coherence length $L_{ph} = \pi/\Delta k_{DFG}^{g}$ becomes equal to the fiber length). These results are in perfect agreement with our expectations based on the analysis of Eqs. (18)–(20). The maxima observed in the DFG signal at an argon pressure of about 0.7 atm for a hollow fiber with an inner diameter of 152 μm and a pressure of 0.9 atm for a fiber with an inner diameter of 203 μm can be attributed, by analogy with the case considered above, to FWM processes involving high-order waveguide modes of the pump and FWM beams.

A methodologically important aspect of nonlinear optical experiments reported in this paper is that the use of sufficiently long picosecond pulses allowed us to study phase-matching effects and effects related to high-order waveguide modes in nonlinear optical interactions in gas-filled hollow fibers under conditions when group-delay effects remain negligible (see the estimates in [12, 14]). Due to the optimization of the phase matching for FWM processes in hollow fibers under these conditions, we were able to considerably increase the fiber length (DFG experiments were performed with the use of hollow fibers with a length up to 30 cm). This allowed us to achieve efficiencies of FWM processes comparable in their order of magnitude with efficiencies attainable with femtosecond pulses of much higher intensities. This result seems to be very important for numerous spectroscopic applications where the bandwidths of laser pulses should be narrower than those typical of the femtosecond range.

## 5. CONCLUSION

Thus, the results of experimental studies devoted to FWM processes of third-harmonic and difference-frequency generation in the field of picosecond laser pulses presented in this paper reveal several important features of nonlinear optical processes in gas-filled hollow fibers, giving a deeper insight into methodological aspects of the problem and opening new possibilities for practical applications of hollow fibers in nonlinear optics, optics of ultrashort pulses, and nonlinear spectroscopy. Unlike experiments carried out with the use of high-intensity laser pulses, our approach allowed us to implement conditions when effects related to the ionization of a gas medium and the self-action of pump pulses do not exert a noticeable influence on nonlinear optical processes in hollow fibers, as indicated by the absence of optical-harmonic generation in tightly focused pump beams in the absence of a hollow fiber. The results of our experiments also demonstrate the possibility of coupling high-intensity picosecond pulses into a hollow fiber without any noticeable perturbation of the gas medium or self-action of laser pulses. Our experimental approach, based on the use of picosecond pulses, allowed us to analyze the main properties of FWM processes in the waveguide regime and to explore the ways of phase-matching such processes under these conditions.

Due to the improvement of phase-matching conditions with an appropriate choice of the gas pressure and optimal parameters of the hollow fiber, we were able to use hollow fibers with a large length (up to 30 cm) for difference-frequency generation, which resulted in a considerable increase in the power of the difference-frequency signal at the output of the fiber. Our experimental results indicate that high-order waveguide modes may have a considerable influence on four-wave mixing processes in gas-filled hollow fibers. This effect can be employed to increase the total energy of short-wavelength radiation produced through nonlinear optical processes in hollow fibers. On the other hand, effects related to high-order waveguide modes should be taken into consideration in the optimization of hollow-fiber frequency converters and pulse compressors, where the excitation of high-order waveguide modes may lead to unwanted energy losses.

Finally, we have shown that the waveguide regime of nonlinear optical interactions implemented in hollow fibers removes the limitations on the efficiency of third-harmonic and sum-frequency generation, which are characteristic of the tight-focusing regime in media with normal dispersion and which are due to the geometric phase shift arising in tightly focused light beams. This finding considerably expands the possibilities of using hollow fibers for frequency conversion and nonlinear optical gas-phase analysis.

## ACKNOWLEDGMENTS

This research was supported in part by the Russian Federation Presidential Grant no. 00-15-99304, the Russian Foundation for Basic Research (project no. 00-02-17567), the federal program “Integration” (State Contract no. AO155), and the US Civilian Research and Development Foundation for the Independent States of the former Soviet Union (CRDF, Award no. RP2-2266).

## REFERENCES

1. M. Nisoli, S. De Silvestri, and O. Svelto, Appl. Phys. Lett. **68**, 2793 (1996).
2. M. Nisoli, S. De Silvestri, O. Svelto, *et al.*, Opt. Lett. **22**, 522 (1997).
3. C. G. Durfee III, S. Backus, M. M. Murnane, and H. C. Kapteyn, Opt. Lett. **22**, 1565 (1997).
4. A. Rundquist, C. G. Durfee III, Z. Chang, *et al.*, Science **5368**, 1412 (1998).
5. Y. Tamaki, K. Midorikawa, and M. Obara, Appl. Phys. B: Lasers Opt. **B67**, 59 (1998).
6. C. G. Durfee III, A. Rundquist, Z. Chang, *et al.*, in Proc. Int. Quant Electronics Conf. (IQEC’98), San Francisco, CA (1998), QPD5.
7. E. Constant, D. Garzella, P. Breger, *et al.*, Phys. Rev. Lett. **82**, 1668 (1999).
8. C. G. Durfee III, A. R. Rundquist, S. Backus, *et al.*, Phys. Rev. Lett. **83**, 2187 (1999).
9. R. B. Miles, G. Laufer, and G. C. Bjorklund, Appl. Phys. Lett. **30**, 417 (1977).
10. A. B. Fedotov, F. Giammanco, A. N. Naumov, *et al.*, Laser Phys. **11** (2001) (in press).
11. A. N. Naumov, A. M. Zheltikov, A. B. Fedotov, *et al.*, submitted to J. Opt. Soc. Am. B.
12. N. I. Koroteev and A. M. Zheltikov, Appl. Phys. B: Lasers Opt. **B67**, 53 (1998).
13. C. G. Durfee III, S. Backus, H. C. Kapteyn, and M. M. Murnane, Opt. Lett. **24**, 697 (1999).
14. A. M. Zheltikov, N. I. Koroteev, and A. N. Naumov, Zh. Éksp. Teor. Fiz. **115**, 1561 (1999) [JETP **88**, 857 (1999)].
15. S. Backus, J. Peatross, Z. Zeek, *et al.*, Opt. Lett. **21**, 665 (1996).
16. A. B. Fedotov, N. I. Koroteev, M. M. T. Loy, *et al.*, Opt. Commun. **133**, 587 (1997).
17. E. A. J. Marcatili and R. A. Schmeltzer, Bell Syst. Tech. J. **43**, 1783 (1964).
18. M. J. Adams, *An Introduction to Optical Waveguides* (Wiley, Chichester, 1981; Mir, Moscow, 1984).
19. Y. R. Shen, *The Principles of Nonlinear Optics* (Wiley, New York, 1984; Nauka, Moscow, 1989).
20. J. F. Reintjes, *Nonlinear Optical Parametric Processes in Liquids and Gases* (Academic, Orlando, 1984; Mir, Moscow, 1987).
21. A. M. Zheltikov, N. I. Koroteev, and A. B. Fedotov, Laser Phys. **4**, 569 (1994).
22. A. B. Fedotov, N. I. Koroteev, A. N. Naumov, *et al.*, J. Nonlinear Opt. Phys. Mater. **6**, 387 (1997).

*Translated by A. Zheltikov*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 256–277.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 291–314.*
*Original Russian Text Copyright © 2001 by Chibisov.*

## ATOMS, SPECTRA, RADIATION

# Electron Capture in Collisions between Negative and Positive Ions

**M. I. Chibisov***

*Institute of Nuclear Fusion, Kurchatov Institute Russian Research Centre, pl. Kurchatova 1, Moscow, 123182 Russia*
**e-mail: chib@nfi.kiae.ru.*

Received February 16, 2001

**Abstract**—Collisions between negative and positive atomic ions are investigated. The ionic wave function is expressed in terms of the Coulomb Green's function. Normalizing this function allows the system of two ions to be described completely. The exchange matrix elements turn out to be the sums of products of the Coulomb wave functions over degenerate states. These sums are expressed in terms of the quadratic form of the wave function for a state with zero angular quantum numbers, $l = m = 0$. The nonadiabatic coupling of quasi-crossing terms with other terms of the system is analyzed; this effect significantly increases the cross section for single-electron capture. *© 2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

Here, we investigate collisions between negative and positive atomic ions. We study the single-electron capture

$$A^- + B^{Z+} \longrightarrow A + B^{(Z-1)+}(n) \qquad (1)$$

at low collision velocities, $v \ll v_0 = 2.19 \times 10^8$ cm s$^{-1}$. The cross section for single-electron capture in a collision of H$^-$ with a proton, $H^- + H^+ = H + H(n)$, and with other singly charged positive ions has previously been calculated in [1–12] and measured in [13–16]. The recombination cross section for H$^-$ in its collision with an α particle, $H^- + He^{++} = H + He^+(n)$, was measured experimentally [17, 18] and calculated theoretically [17, 19] (see also [20, 21]); $H^- + \{Ne^{3+}, Ar^{3+}\}$ collisions were studied in experimental measurements [22, 23] and in [24]. The Ca$^-$ production cross section in the reverse reaction

$$Ca + B^*(n) \longrightarrow Ca^- + B^+ \qquad (2)$$

was calculated in [25–27] and measured in [28–30].

The properties of a collisional system that can produce a pair of oppositely charged ions are determined by the crossing of energy terms. The probability of single-electron exchange reactions (1) and (2) at low collision velocities significantly increases in the presence of such crossings. Since the binding energies of negative ions are low, the ionic term crosses the terms of highly excited states with large principal quantum numbers $n$. Such states of any atom are similar in properties to purely Coulomb states, whose energies do not depend (or depend weakly) on orbital quantum numbers $l$ and $m$. Consequently, in collisions (1) and (2), the ionic term crosses a complex variety of degenerate states whose number for each $n$ is $n^2$.

Previously, the authors of [31–36] used the approximation of zero-range potentials, the approximation of δ potentials, to investigate reaction (1). The latter approximation is applicable to negative ions whose weakly bound electrons have zero orbital angular momenta, as, for example, for H$^-$. The energy levels in the above papers were determined by solving a transcendental equation with the logarithmic derivative of Coulomb Green's function. It was found that for each crossing and for each $n$, only one of the total number $n^2$ of states interacted with the ionic state, while the energies of the remaining $n^2 - 1$ states, the so-called passive states, were not perturbed by the ionic term, and these states were disregarded in specific calculations [33].

In reality, however, the passive states are involved in processes (1) and (2). Although the energies of the passive states are not perturbed, their wave functions are time dependent, because they result when reconstructing the basis of Coulomb functions and prove to be dependent on the separation between the colliding ions. The adiabatic matrix element of the time derivative between active and passive states is nonzero. Here, we construct a complete adiabatic basis of wave functions, which is used to solve an adiabatic system of close-coupling equations.

In our approximation, the wave function of the outer weakly bound electron is expressed in terms of Coulomb Green's function. This function, normalized to unity as the wave function of a bound state with a negative energy [7, 8], was analyzed in detail on the $E$, $R$ (energy, internuclear separation; $-\infty \le E \le +\infty$, $0 \le R \le \infty$) half-plane. The analysis yielded energy separations between the adiabatic terms for quasi-crossings and all the matrix elements between the adiabatic states; i.e., it allowed the $A^- + B^{Z+}$ system to be described completely.

This approach allows negative ions with any orbital angular momentum $L$ to be considered.

In this paper, we consider $L = 0$ ($H^-$) and $L = 1$ ($Ca^-$). The term separation for quasi-crossings was found to be expressed in terms of the sums of products of the Coulomb wave functions over degenerate states, i.e., over angular quantum numbers $l$ and $m$. An analysis of the closed expression for Coulomb Green's function derived by Hostler and Pratt [37, 38] allows these sums to be expressed in terms of the quadratic form of the wave function only for one state with zero angular quantum numbers, $l = m = 0$. Such sums, which describe the behavior of a negative ion with a zero orbital angular momentum, $L = 0$, in the field of a positive ion, were calculated previously [39–41], while similar sums for ions with $L = 1$ are calculated here.

We also study the following nonadiabaticity effect discussed previously [24] in the atomic-basis approximation. If the ionic term quasi-crosses the $n$th covalent term, then the matrix element between the ionic state and another covalent state $n'$ experiences a sharp jump, causing the population of state $n'$ to increase. This effect proves to be strong. It appreciably increases the total cross section for single-electron capture (1).

Here, we use the system of atomic units, $e^2 = m = \hbar = 1$.

## 2. THE WAVE FUNCTION OF THE OUTER WEAKLY BOUND ELECTRON

The wave function $\Phi(\mathbf{R}, \mathbf{r})$ of the weakly bound electron in the $A^- + B^{Z+}$ system is the solution of the wave equation

$$\left(-\frac{\Delta}{2} - \frac{Z}{r} + V(|\mathbf{R} - \mathbf{r}|) - E\right)\Phi(\mathbf{R}, \mathbf{r}) = 0, \tag{3}$$

where $V(|\mathbf{R} - \mathbf{r}|)$ is the potential energy of interaction between the weakly bound electron and the core of atom A. The wave function $\Phi_0(\mathbf{r})$ of the unperturbed negative ion $A^-$ satisfies the equation

$$\left(-\frac{\Delta}{2} + V(|\mathbf{R} - \mathbf{r}|) - E_0\right)\Phi_0(\mathbf{r}) = 0. \tag{4}$$

Let us consider Green's function $G(\mathbf{r}, \mathbf{r}', E)$ that is the solution of the equation

$$\left(-\frac{\Delta}{2} - \frac{Z}{r} - E\right)G(\mathbf{r}, \mathbf{R}, E) = \delta(\mathbf{R} - \mathbf{r}), \tag{5}$$

and whose spectral expansion over Coulomb eigenfunctions is [42]

$$G(\mathbf{r}, \mathbf{R}, E) = \sum_{nlm} \frac{\psi_{nlm}(\mathbf{r})\psi^*_{nlm}(\mathbf{R})}{E_n - E}, \tag{6}$$

where the sum denotes summation over the discrete states with a negative energy and integration over the continuum. In spherical coordinates, the Coulomb eigenfunctions of bound states with a negative energy are [43]

$$\begin{aligned} \psi_{nlm}(\mathbf{r}) &= f_{nl}(r)Y_{lm}(\theta, \phi), \\ Y_{lm}(\theta, \phi) &= N_{lm}P_l^{|m|}(\cos\theta)\frac{\exp(im\phi)}{\sqrt{2\pi}}, \end{aligned} \tag{7}$$

$$\begin{aligned} &\frac{d^2 f_{nl}}{dr^2} + \frac{2}{r}\frac{d f_{nl}}{dr} \\ &+ 2\left(E_n + \frac{Z}{r} - \frac{l(l+1)}{2r^2}\right)f_{nl} = 0, \end{aligned} \tag{8}$$

$$\begin{aligned} f_{nl}(r) &= N_{nl}\left(\frac{2Zr}{n}\right)^l \exp\left(\frac{-Zr}{n}\right) \\ &\times F\left(-n + l + 1; 2l + 2; \frac{2Zr}{n}\right), \end{aligned} \tag{9}$$

$$E_n = -Z^2/2n^2, \quad N_{nl} = \frac{2Z^{3/2}}{n^2(2l+1)!}\sqrt{\frac{(n+l)!}{(n-l-1)!}},$$

$$N_{lm} = \sqrt{\frac{2l+1}{2}\frac{(l-|m|)!}{(l+|m|)!}}.$$

Using Green's function, we may write the following integral equation instead of the differential equation (3):

$$\Phi(\mathbf{R}, \mathbf{r}) = -\int G(\mathbf{r}, \mathbf{r}', E)V(|\mathbf{r}' - \mathbf{R}|)\Phi(\mathbf{R}, \mathbf{r}')d\mathbf{r}'. \tag{10}$$

Since the effective potential $V$ is relatively small, we may expand Green's function into a Taylor series about variable $\mathbf{r}'$ at point $\mathbf{r}' = \mathbf{R}$:

$$G(\mathbf{r}, \mathbf{r}', E) \approx G(\mathbf{r}, \mathbf{R}, E) + \left.\frac{dG}{d\mathbf{r}'}\right|_{\mathbf{r}' = \mathbf{R}}(\mathbf{r}' - \mathbf{R}) + \dots. \tag{11}$$

For large internuclear separations satisfying the condition

$$\gamma R \gg 1, \quad \gamma = \sqrt{-2\varepsilon_0}, \tag{12}$$

the perturbation of a negative ion by a positive ion is small; it reduces to a comparatively small polarization displacement of its energy level $\varepsilon_0$. For $H^-$, $\varepsilon_0 = -0.75421$ eV $= -0.027716$ a.u., $\gamma = 0.23544$, and it follows from condition (12) that the internuclear separations are $R \gg 4.25a_0$, where $a_0$ is the Bohr radius. For $Ca^-(4s^24p)$ with two bound levels, two fine-structure components with total angular momenta $j = 1/2$ and $3/2$ are $\varepsilon_0^{1/2} = -0.02455$ eV, $\gamma_{1/2} = 0.04248$, $R \gg 24a_0$ and $\varepsilon_0^{3/2} = -0.01973$ eV, $\gamma_{3/2} = 0.03808$, $R \gg 26a_0$. The distances that mainly contribute to the cross sections for the processes under study are larger by more than a factor of 10 than those limiting distances (see below). Therefore, condition (12) is satisfied with a large margin, and the unperturbed ionic wave function may be substituted in the integral of motion (10).

### *2.1. A Negative Ion with a Zero Orbital Angular Momentum, L = 0*

Let us first consider a negative ion whose weakly bound electron has a zero orbital angular momentum, $L = 0$. In this case, the wave function of the unperturbed negative ion may be chosen to be

$$\Phi_0(|\mathbf{r}-\mathbf{R}|) = N_0\frac{\exp(-\gamma|\mathbf{r}-\mathbf{R}|)}{|\mathbf{r}-\mathbf{R}|},$$
$$N_0 = \sqrt{\frac{\gamma}{2\pi}}, \quad \varepsilon_0 = -\frac{\gamma^2}{2}, \quad L = 0, \tag{13}$$

which is the limit of the ionic wave function $\Phi$ for an infinite internuclear separation.

Substituting function (13) and expanding Green's function (11) in integral (10), we find that the wave function of the outer electron for $L = 0$ is Green's function at $\mathbf{r}' = \mathbf{R}$,

$$\Phi(\mathbf{R}, \mathbf{r}) \approx C_0 G(\mathbf{r}, \mathbf{R}, E), \tag{14}$$

because $\Phi_0$ for $L = 0$ does not depend on the angles and the contribution from the second term of expansion (11) vanishes. The contributions from the higher order terms of expansion (11) are proportional to positive powers of the low binding energy for the negative ion, $|\varepsilon_0|^p \ll 1$, $p \geq 1$, and they may be ignored. The constant $C_0$ in (14) can be determined by normalization.

A closed expression for Coulomb Green's function was derived by Hostler and Pratt [37, 38]:[1]

$$G(\mathbf{r}, \mathbf{r}', E) = \frac{\Gamma(1-Z\nu)}{2\pi|\mathbf{r}-\mathbf{r}'|}$$
$$\times [W(\tau_x)M'(\tau_y) - W'(\tau_x)M(\tau_y)], \tag{15}$$

$$\tau_{\{x,y\}} = \frac{\{x, y\}}{\nu}, \quad x = r + r' + |\mathbf{r}-\mathbf{r}'|,$$
$$y = r + r' - |\mathbf{r}-\mathbf{r}'|, \quad \nu \equiv \frac{1}{\sqrt{-2E}}. \tag{16}$$

The Whittaker functions $W(\tau)$ and $M(\tau)$ have the subscripts $Z\nu$ and 1/2 in (15) and are the solutions of the equation [44]

$$W''_{Z\nu,\,1/2}(\tau) + \left(-\frac{1}{4} + \frac{Z\nu}{\tau}\right)W_{Z\nu,\,1/2}(\tau) = 0 \tag{17}$$

and a similar equation for $M_{Z\nu,\,1/2}(\tau)$. In the limit $E \longrightarrow E_n$, where $Z\nu \longrightarrow n$, Eq. (17) when changing variables $\tau = 2Zr/n$ transforms to the equation for the Coulomb function $\phi_{n0} = r\psi_{n0}(r)$ for the state with a zero orbital angular momentum, $l = 0$,

$$\frac{d^2\phi_{n0}}{dr^2} + 2\left(E_n + \frac{Z}{r}\right)\phi_{n0} = 0. \tag{18}$$

The Whittaker function $M$ is expressed via a linear combination of $W_{Z\nu}(\tau)$ and $W_{-Z\nu}(-\tau)$ [44]:

$$\Gamma(1-Z\nu)M_{Z\nu,\,1/2}(\tau) = (-1)^{1+Z\nu}\frac{\Gamma(1-Z\nu)}{\Gamma(1+Z\nu)}W_{Z\nu,\,1/2}(\tau)$$
$$+ (-1)^{Z\nu}W_{-Z\nu,\,1/2}(-\tau), \tag{19}$$

$W_{Z\nu}(\tau)$ exponentially decreases, while $W_{-Z\nu}(-\tau)$ exponentially increases as $\tau \longrightarrow \infty$.

Using the linear combination (19), we may write Green's function (15) at $\mathbf{r}' = \mathbf{R}$ as the sum of two terms:

$$G(\mathbf{r}, \mathbf{R}, E) = \frac{(-1)^{1-Z\nu}}{2\pi}\frac{\Gamma(1-Z\nu)}{\Gamma(1+Z\nu)}G_1(\mathbf{r}, \mathbf{R}, E)$$
$$+ \frac{(-1)^{Z\nu}}{2\pi}G_2(\mathbf{r}, \mathbf{R}, E), \tag{20}$$

$$G_1(\mathbf{r}, \mathbf{R}, E) = \frac{W_+(\tau_x)W'_+(\tau_y) - W'_+(\tau_x)W_+(\tau_y)}{|\mathbf{r}-\mathbf{R}|}, \tag{21}$$

$$G_2(\mathbf{r}, \mathbf{R}, E) = \frac{W_+(\tau_x)W'_-(\tau_y) - W'_+(\tau_x)W_-(\tau_y)}{|\mathbf{r}-\mathbf{R}|}, \tag{22}$$

$$\mathbf{r}' = \mathbf{R},$$

where $W_+ \equiv W_{Z\nu,\,1/2}(\tau)$ and $W_- \equiv W_{-Z\nu,\,1/2}(-\tau)$. The functions $W_+$ and $W_-$ are not regular at zero, but their linear combination (19), i.e., the function $M$, is regular at zero. Accordingly, the functions $G_1$ and $G_2$ are not regular in the limit $\mathbf{r} \longrightarrow 0$ or $\mathbf{r}' \longrightarrow 0$, but their sum, i.e., the full Green's function (15), is regular in these limits. The functions $G_2$ and $G_1$ may be called the ionic and covalent parts of Green's function, respectively.

The binding energies $\varepsilon_0$ of negative ions are low, and we are interested in the low energies $E$ of the $H^- + A^{Z+}$ system. For such energies, we can use either the asymptotics of the Whittaker functions in subscript $Z\nu \longrightarrow \infty$ [45] or calculate them in the semiclassical approximation to solve Eq. (17). We use the semiclassical approximation [7], in which the solutions of Eq. (17) in the subbarrier region are

$$W_{Z\nu,\,1/2}(\tau) = C_1\frac{\exp\left(-\int_{\tau_0}^{\tau}|p(\tau')|d\tau'\right)}{(1-\tau_0/\tau)^{1/4}},$$

---

[1] Our definition of Green's function in formulas (6) and (15) differ from those in [37, 38] by the factor (–2).

$$W_{-Z\nu,\,1/2}(-\tau) = C_2\frac{\exp\left(+\int\limits_{\tau_0}^{\tau}|p(\tau')|\,d\tau'\right)}{(1-\tau_0/\tau)^{1/4}}, \tag{23}$$

$$|p(\tau)| = \sqrt{\frac{1}{4}-\frac{Z\nu}{\tau}}, \quad \tau > \tau_0 \equiv 4Z\nu.$$

Integration yields the function $W$ [7]:

$$W(\tau) = \frac{\exp\left(\pm\frac{1}{2}\sqrt{\tau(\tau-\tau_0)} \pm Z\nu \pm 2Z\nu\ln\frac{\sqrt{\tau}-\sqrt{\tau-\tau_0}}{2Z\nu}\right)}{(1-\tau_0/\tau)^{1/4}}, \tag{24}$$

where the plus and minus signs correspond to $W_{-Z\nu,\,1/2}(-\tau)$ and $W_{Z\nu,\,1/2}(\tau)$, respectively. In view of relations (24), $M$ is the linear combination (19).

The asymptotics of the Whittaker functions $W_\pm$ for $\tau \longrightarrow \infty$ are [44]

$$W_{Z\nu,\,1/2}(\tau) \approx \tau^{Z\nu}\exp\left(-\frac{\tau}{2}\right)$$
$$\times\left(1-\frac{Z\nu(Z\nu-1)}{\tau}+\frac{Z\nu(Z\nu-1)^2(Z\nu-2)}{2\tau^2}+\ldots\right), \tag{24a}$$
$$W_{-Z\nu,\,1/2}(-\tau) \approx \tau^{-Z\nu}\exp\left(\frac{\tau}{2}\right)$$
$$\times\left(1+\frac{Z\nu(Z\nu+1)}{\tau}+\frac{Z\nu(Z\nu+1)^2(Z\nu+2)}{2\tau^2}+\ldots\right).$$

Note that for highly excited states, when $Z\nu \geq 3$–$5$, the second and third terms of these asymptotic expansions become small at comparatively large distances exceeding the size of the region of classically permitted motion, $r_\nu = Z|E|^{-1}$, for a given energy $E$ by a factor of 3 or 4. At the same time, calculating $W$ from formulas (24a) with three terms of the asymptotic expansion provides high accuracy (~1%) up to distances equal to or smaller than $r_\nu$.

It is convenient to use the semiclassical representation (24) of $W_\pm$ to analyze the ionic part of Green's function $G_2$ in the vicinity of the negative ion, i.e., at $|\mathbf{r}-\mathbf{R}| \ll R$. The corresponding expansion yields

$$G_2(\mathbf{r},\mathbf{R},E) \approx \frac{\exp(-\gamma_{scl}|\mathbf{r}-\mathbf{R}|)}{|\mathbf{r}-\mathbf{R}|},$$
$$|\mathbf{r}-\mathbf{R}| \ll R, \quad \gamma_{scl}(R) \equiv \sqrt{-2(E+Z/R)}. \tag{25}$$

The form of this function is the same as that of the unperturbed function (13). If, however, the energy $E$ is equal to the negative-ion energy in the zero-order approximation,

$$E \approx E_0(R) = \varepsilon_0 - \frac{Z}{R}, \tag{26}$$

then the exponents in functions (25) and (13) are also equal, $\gamma_{scl} = \gamma$. For several energies $E$, the product $|\mathbf{r}-\mathbf{R}|G_2$ was calculated exactly using series in positive powers of the argument for the Whittaker functions [44]. The result obtained was compared with formula (25). The discrepancies did not exceed a few percent; i.e., the semiclassical approximation proves to be highly accurate.

The function $G_1$ has no singularity for $\mathbf{r} \longrightarrow \mathbf{R}$, and its behavior in the vicinity of the negative ion can be determined by expanding expression (21) in a Taylor power series of $\tau_x - \tau_0$ and $\tau_y - \tau_0$. The result of this expansion is

$$G_1(\mathbf{r},\mathbf{R},E) = \frac{W_+(\tau_x)W'_+(\tau_y) - W'_+(\tau_x)W_+(\tau_y)}{|\mathbf{r}-\mathbf{R}|}$$
$$\approx \frac{\nu}{2}\left[\left(\frac{dW_+}{dR}\right)^2 + 2\left(E+\frac{Z}{R}\right)W_+^2\right]_{\tau = 2R/\nu} \tag{27}$$
$$-\frac{Z\nu}{2}(r-R)\frac{W_+^2(2R/\nu)}{R^2}+\ldots$$

The first expansion term describes the separation of energy terms for quasi-crossings.

Let us now turn to normalizing Green's function considered as the wave function [see relation (14)]. The normalized wave function $\Phi(\mathbf{R},\mathbf{r})$ can be written as

$$\Phi(\mathbf{R},\mathbf{r}) = \frac{N(R)}{|\mathbf{r}-\mathbf{R}|}\left[W\left(\frac{x}{\nu}\right)M'\left(\frac{y}{\nu}\right) - W'\left(\frac{x}{\nu}\right)M\left(\frac{y}{\nu}\right)\right],$$
$$N(R) = \left[\int\left(W\left(\frac{x}{\nu}\right)M'\left(\frac{y}{\nu}\right)\right.\right. \tag{28}$$
$$\left.\left.-W'\left(\frac{x}{\nu}\right)M\left(\frac{y}{\nu}\right)\right)^2\frac{d\tau}{(\mathbf{r}-\mathbf{R})^2}\right]^{-1/2},$$

or as

$$\Phi(\mathbf{R},\mathbf{r}) = 2\pi N_0 B(R)G(\mathbf{r},\mathbf{R},E)$$
$$= \frac{N(R)}{\Gamma(1+Z\nu)}G_1(\mathbf{r},\mathbf{R},E) + B(R)N_0G_2(\mathbf{r},\mathbf{R},E), \tag{29}$$

where

$$B(R) \equiv \frac{N(R)}{N_0\Gamma(1-Z\nu)}. \tag{30}$$

When calculating the normalization factor $N(R)$, we performed numerical integration over positive powers of the argument using series for the Whittaker functions $W$ and $M$ [44]. The energy was assumed to depend on the internuclear separation via relation (26) of the zero-

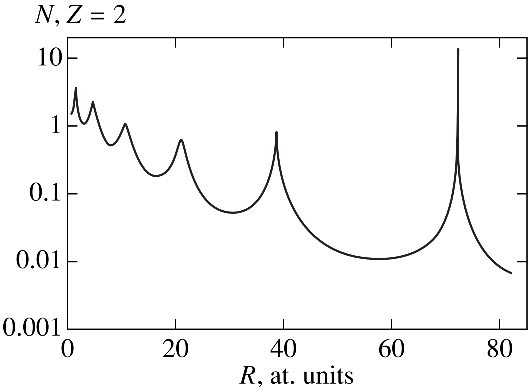

**Fig. 1.** Normalization factor $N$ as a function of the internuclear separation $R$ for the $H^-$ + $He^{++}$ system.

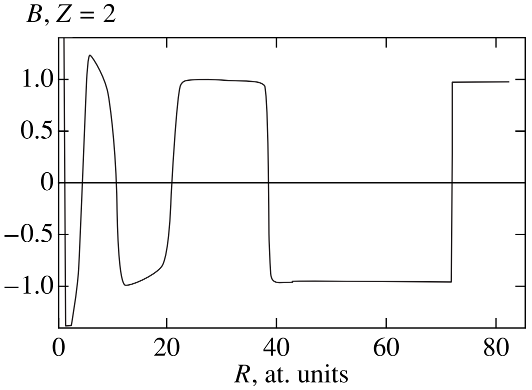

**Fig. 2.** Function $B$ as a function of internuclear separation $R$ for the $H^-$ + $He^{++}$ system.

order approximation. The question of energy is considered in detail in Sect. 4. The derived functions $N(R)$ and $B(R)$ are shown in Figs. 1 and 2 for $Z = 2$. The normalization factor $N(R)$ is seen from Fig. 1 to have sharp peaks for term crossings, when $E = E_n = -Z^2/2n^2$.

Far from term crossings, when the integrand in (28) is large and the normalization factor $N(R)$ is small, the exponentially increasing term is dominant in the linear combination (19) (the argument of $M$ varies over the range $0 \le \tau_y \le 2R/\nu$). Alternatively, near crossings, the exponentially decreasing term is dominant in (19); the integrand in (28) is small and $N(R)$ is large.

Figure 2 shows the function $B(R)$. Far from term crossings, when the exponentially increasing term is dominant in (19), $B(R)$ is close to 1. In this case, in the normalized Green's function (29), the first term with $G_1$ is small, while the second term with $G_2$ matches the unperturbed function (13). The second term in (29) is given by expression (25) at any distances $R$, while at energy of the zero-order approximation (26), when $\gamma_{scl} = \gamma$, the second term in (29) and, hence, the entire ionic function $\Phi(\mathbf{R}, \mathbf{r})$ are close to the unperturbed function (13). In this case, the electron stays mainly near the negative ion. For term crossings, when the exponentially decreasing term in (19) becomes dominant and $B(R)$ decreases to zero, the electron moves to the positively charged center. The first term with $G_1$ becomes dominant in (29).

Using the spectral representation (6) for Green's function, we can write a similar representation for the wave function of the outer electron,

$$\Phi(\mathbf{R}, \mathbf{r}) = \sum_n C_n(E)\Psi_n(\mathbf{R}, \mathbf{r}), \tag{31}$$

where the normalized adiabatic wave functions of covalent states $\Psi_n(\mathbf{R}, \mathbf{r})$ are determined by the sums over degenerate Coulomb states:

$$\begin{aligned}\Psi_n(\mathbf{R}, \mathbf{r}) &= \frac{1}{\sqrt{Q_n(R)}}\sum_{l=0}^{n-1}\sum_{m=-l}^{l}\psi^*_{nlm}(\mathbf{r})\psi_{nlm}(\mathbf{R}) \\ &= \sum_{l=0}^{n-1}\sum_{m=-l}^{l}J_{nlm}(\mathbf{R})\psi^*_{nlm}(\mathbf{r}),\end{aligned} \tag{32}$$

$$J_{nlm}(\mathbf{R}) \equiv \frac{\psi_{nlm}(\mathbf{R})}{\sqrt{Q_n(R)}} = \int\psi_{nlm}(\mathbf{r})\Psi_n(\mathbf{R}, \mathbf{r})d\mathbf{r}, \tag{32a}$$

while the expansion coefficients $C_n(E)$ and the sum $Q_n(R)$ are

$$C_n(E) = -\frac{2\pi N_0 B(E)}{E - E_n}Q_n^{1/2}(R), \tag{33}$$

$$Q_n(R) \equiv \sum_{l=0}^{n-1}\sum_{m=-l}^{l}|\psi_{nlm}(\mathbf{R})|^2. \tag{34}$$

Note that using the summation rule for spherical functions $Y_{lm}$ [43]:

$$\sum_{m=-l}^{+l}Y^*_{lm}(\boldsymbol{\nu}_1)Y_{lm}(\boldsymbol{\nu}_2) = \frac{2l+1}{4\pi}P_l(\cos\alpha), \tag{35}$$

where $\alpha$ is the angle between unit vectors $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$, the sum in (32), in terms of which the function $\Psi_n(\mathbf{R}, \mathbf{r})$ is expressed, can be written as a single sum only over quantum number $l$:

$$\begin{aligned}\hat{Q}_n(\mathbf{r}, \mathbf{R}) &\equiv \sum_{lm}\psi^*_{nlm}(\mathbf{r})\psi_{nlm}(\mathbf{R}) \\ &= \frac{1}{4\pi}\sum_{l=0}^{n-1}(2l+1)P_l(\cos\theta_{\mathbf{r},\mathbf{R}})f_{nl}(r)f_{nl}(R).\end{aligned} \tag{36}$$

It follows from (36) that $\hat{Q}_n(\mathbf{r}, \mathbf{R})$ does not depend on the $z$-axis orientation in the coordinate system in which the functions $\psi_{nlm}(\mathbf{r})$ were defined.

Analyzing the limit of Coulomb Green's function for $E \longrightarrow E_n$ allowed the sums (32) and (34) to be expressed in terms of the quadratic form of the wave function only for one state with zero orbital quantum numbers, $l = m = 0$ [39, 40]:

$$\Psi_n(\mathbf{R}, \mathbf{r}) = \frac{\hat{Q}_n(\mathbf{R}, \mathbf{r})}{\sqrt{Q_n(R)}},$$

$$\hat{Q}_n(\mathbf{R}, \mathbf{r}) = \frac{4Z^2}{n^2}\frac{\phi'_{n0}(\tau_y)\phi_{n0}(\tau_x) - \phi_{n0}(\tau_y)\phi'_{n0}(\tau_x)}{\tau_x - \tau_y}, \qquad (37)$$

$$Q_n(R) = \sum|\psi_{nlm}(\mathbf{R})|^2 = \left(\frac{d\phi_{n0}(R)}{dR}\right)^2$$

$$+ 2\left(E_n + \frac{Z}{R}\right)\phi^2_{n0}(R) = 2Z\int_R^\infty \frac{\phi^2_{n0}(r)}{r^2}dr, \qquad (38)$$

$$\phi_{n0}(r) = r\psi_{n0}(\mathbf{r}), \quad \tau_{x,y} = \frac{Z}{n}[r + R \pm |\mathbf{r} - \mathbf{R}|].$$

It follows from (36) and (37) that $\Psi_n(\mathbf{R}, \mathbf{r})$ are real functions. From (38), it follows that $Q_n(R)$ is nonzero at finite $R$ [but $Q_n(R) \longrightarrow 0$ as $R \longrightarrow \infty$], because the integrand on the right-hand side of (38) is positive. We also see from (38) that zeros of the derivative $Q'_n(R)$ coincide with zeros of the function $\phi_{n0}(r)$; hence, $Q_n(R)$ decreases to zero in steps as $R \longrightarrow \infty$ (see Fig. 3).

### *2.2. A Negative Ion with a Nonzero Orbital Angular Momentum, L = 1*

To investigate the energy levels of the $A^- + B^{Z+}$ system with a nonzero orbital angular momentum of the $A^-$ ion and the parameters on which these energies depend, it is convenient first to introduce a coordinate system with the $z$ axis directed along the vector of internuclear separation $\mathbf{R}$. The energy of the system under study depends on the internuclear separation of the colliding particles and on the absolute value of $R$, but does not depend on the direction of vector $\mathbf{R}$.

The three angular components $\{x, y, z\}$ of the unperturbed wave function for a weakly bound electron with orbital angular momentum $L = 1$ are

$$\Phi_{0, \{x, y, z\}}(\mathbf{r}_b) = \frac{\chi_0(r_b)}{r_b}$$

$$\times\sqrt{\frac{3}{4\pi}}\{\cos\theta_b, \sin\theta_b\cos\phi_b, \sin\theta_b\sin\phi_b\}, \qquad (39)$$

$$\mathbf{r}_b = \mathbf{r} - \mathbf{R},$$

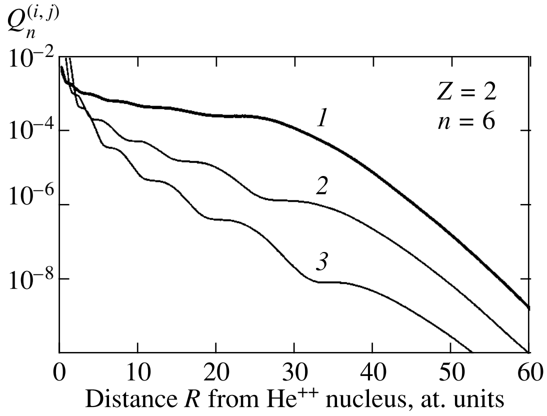

**Fig. 3.** The sums of products of the Coulomb functions over angular quantum numbers $l$ and $m$ for $Z = 2$ and principal quantum number $n = 6$: *1*—$Q_n(R)$ [$Q_n^{(0,0)}(R)$], formulas (38), (100); *2*—$Q_n^{(1,1)}(R)$, formulas (96), (100); and *3*—$Q_n^{(2,2)}(R)$, formulas (97), (100).

where the radial function $\chi_0(r_b)$ is the solution of the equation

$$\frac{d^2\chi_0(r_b)}{dr_b^2} + \left[2(\varepsilon_0 - V(r_b)) - \frac{L(L+1)}{r_b^2}\right]\chi_0(r_b) = 0, \qquad (40)$$

$$L = 1.$$

The azimuthal angle $\phi$ is measured from the collision plane.

When substituting the wave function (39) and the expansion of Green's function (11) in the right-hand part of integral equation (10), we find the integral (10) of the first term of expansion (11) to be zero. For a negative ion with orbital angular momentum $L = 1$, the two-center ionic wave function is given by the second term of expansion (11), and, hence, it is proportional to the derivatives of Green's function

$$\Phi_x \propto \left.\frac{\partial G}{\partial x'}\right|_{\mathbf{r}' \to \mathbf{R}}, \quad \Phi_y \propto \left.\frac{\partial G}{\partial y'}\right|_{\mathbf{r}' \to \mathbf{R}}, \quad \Phi_z \propto \left.\frac{\partial G}{\partial z'}\right|_{\mathbf{r}' \to \mathbf{R}}, \qquad (41)$$

with

$$\left.\frac{\partial G}{\partial z'}\right|_{\mathbf{r}' \to \mathbf{R}} = -\cos\theta_b\frac{\Gamma(1 - Z\nu)}{2\pi}\frac{F_z\{M, W\}}{|\mathbf{r} - \mathbf{R}|}, \qquad (42)$$

$$\left.\frac{\partial G}{\partial x'}\right|_{\mathbf{r}' \to \mathbf{R}} = -\sin\theta_b\cos\varphi_b\frac{\Gamma(1 - Z\nu)}{2\pi}\frac{F_x\{M, W\}}{|\mathbf{r} - \mathbf{R}|}, \qquad (43)$$

$$\left.\frac{\partial G}{\partial y'}\right|_{\mathbf{r}' \to \mathbf{R}} = -\sin\theta_b\sin\varphi_b\frac{\Gamma(1 - Z\nu)}{2\pi}\frac{F_y\{M, W\}}{|\mathbf{r} - \mathbf{R}|}. \qquad (44)$$

Here, $\theta_b$ and $\varphi_b$ are the spherical angles of vector $\mathbf{r}_b \equiv \mathbf{r} - \mathbf{R}$ in the coordinate system centered on the nucleus of the negative ion $A^-$:

$$\cos\theta_b = \frac{z - z_R}{|\mathbf{R} - \mathbf{r}|}, \quad \sin\theta_b\cos\varphi_b = \frac{x - x_R}{|\mathbf{R} - \mathbf{r}|}, \quad \sin\theta_b\sin\varphi_b = \frac{y - y_R}{|\mathbf{R} - \mathbf{r}|}; \tag{45}$$

$x_R$, $y_R$, and $z_R$ are the components of vector $\mathbf{R}$ in the coordinate system centered on the nucleus of the positive ion. If the $z$ axis is directed along vector $\mathbf{R}$, then $z_R = R$, $x_R = y_R = 0$, and $F_{xyz}\{M, W\}$ are

$$F_z\{M, W\} \equiv \frac{2}{\nu}W'M' + \left(-\frac{1}{2\nu} + \frac{Z\nu}{R}\frac{R - r}{R - z}\right)WM - \frac{WM' - W'M}{|\mathbf{R} - \mathbf{r}|}, \tag{46}$$

$$F_{xy}\{M, W\} \equiv \frac{2}{\nu}W'M' + \left(-\frac{1}{2\nu} + \frac{Z\nu}{R}\frac{r + R}{r + z}\right)WM - \frac{WM' - W'M}{|\mathbf{R} - \mathbf{r}|}, \tag{47}$$

$F_x\{M, W\} = F_y\{M, W\} = F_{xy}\{M, W\}$ and $x$, $y$, $z$ are the components of vector $\mathbf{r}$.

The asymptotics of the radial unperturbed function $\chi_0(r_b)$ [see Eq. (40)] is governed mainly by the centrifugal potential. The atomic potential, which is determined by the polarization interaction $V(r_b) \longrightarrow -\alpha/2r_b^4$ at large distances, may be disregarded. The asymptotics of $\chi_0(r_b)$ is then[2]

$$\chi_0^{as} = N_0^{(1)}\left(1 + \frac{1}{\gamma r_b}\right)\exp(-\gamma r_b). \tag{48}$$

The coefficient $N_0^{(1)} = 0.112$ was calculated by numerically solving Eq. (40) using the model potential $V(r_b)$ from [26, 27].

If we use the semiclassical approximation (24) for the Whittaker functions [7], substitute them in (46) and (47), and expand on condition that $|\mathbf{r} - \mathbf{R}| \ll R$, then we obtain for $F_{z,\,xy}\{M, W\}$

$$F_z \approx F_{xy} \approx \left(1 + \frac{1}{\gamma_{scl}|\mathbf{r} - \mathbf{R}|}\right)\exp(-\gamma_{scl}|\mathbf{r} - \mathbf{R}|), \tag{49}$$

here, the constant $\gamma_{scl}$ in the semiclassical approximation is defined in (25). Expression (49), to within a constant factor, coincides with (48) if the ionic-state energy is given by the zero-order approximation (26) at $\gamma_{scl} = \gamma$. Thus, the constructed functions (41)–(44), (46), and (47) satisfy the necessary condition: at large internuclear separations [see condition (12)], they coincide with the unperturbed functions near the negative ion.

The region of large distances from the nucleus of the negative ion mainly contributes to the normalization integral, but the contribution from asymptotics (48) and (49) diverges when $r_b \longrightarrow 0$. Therefore, the specific expressions for the three components of the two-center ionic wave functions for $L = 1$ were chosen to be

$$\Phi_z(\mathbf{r}) = N_z\frac{\cos\theta_b}{r_b}\hat{F}_z\{M, W\}, \tag{50}$$

$$\Phi_x(\mathbf{r}) = N_{xy}\frac{\sin\theta_b\cos\hat{\varphi}_b}{r_b}\hat{F}_{xy}\{M, W\}, \tag{51}$$

$$\Phi_y(\mathbf{r}) = N_{xy}\frac{\sin\theta_b\sin\varphi_b}{r_b}\hat{F}_{xy}\{M, W\}, \tag{52}$$

where the following relation was used for each of the three components:

$$\hat{F}\{M, W\} = \begin{cases} \dfrac{2\pi\chi_0(r_b)}{\gamma\Gamma(1 - Zn_-)}, & r_b \le r_b^0 \\ F\{M, W\}, & r_b \ge r_b^0. \end{cases} \tag{53}$$

The function $\chi_0$ was determined by numerically solving Eq. (40). The joining point $r_b^0 = 25a_0$ was chosen in such a way that the solution of Eq. (40) for $r_b \ge r_b^0$ was close to the asymptotic expression (48). The model atomic potential $V(r_b)$ for Ca was taken from [26, 27].

The normalization constants $N_{x,\,y,\,z}$ of ionic functions (50)–(52),

$$N_z = \left(2\pi\int_0^\infty dr_b\int_0^\pi \hat{F}_z^2\cos^2\theta_b\sin\theta_b d\theta_n\right)^{-1/2}, \tag{54}$$

$$N_{xy} = \left(\pi\int_0^\infty dr_b\int_0^\pi \hat{F}_{xy}^2\sin^3\theta_b d\theta_b\right)^{-1/2}, \tag{55}$$

were calculated by using expansions of the Whittaker functions in terms of positive powers of the argument [44] and a numerical solution of Eq. (40).

Similar to the case of $L = 0$, the ionic wave function for $L = 1$ can be written as

$$\Phi_{x,\,y,\,z}(\mathbf{R}, \mathbf{r}) = N_0^{(1)}\frac{\sqrt{3\pi}}{\gamma}B_{x,\,y,\,z}(E)\left.\frac{\partial G(\mathbf{r}, \mathbf{r}', E)}{\partial\{x', y', z'\}}\right|_{\mathbf{r}' = \mathbf{R}}, \tag{56}$$

$$B_{x,\,y,\,z}(E) \equiv \frac{\gamma N_{x,\,y,\,z}(E)}{N_0^{(1)}\sqrt{3/4\pi}\,\Gamma(1 - Z\nu)}. \tag{57}$$

The numerically calculated functions $N(E)$ and $B(E)$ (using the zero-order approximation for energy $E$) are similar to these functions for $L = 0$ (see Figs. 1 and 2).

---

[2] Note that (48) is an exact solution of Eq. (40) at $V(r_b) = 0$.

Using the spectral representation of Green's function (6), we write a similar expansion for the ionic function:

$$\Phi_{x,y,z}(\mathbf{R},\mathbf{r}) = \sum_n C_n^{x,y,z}(E)\Psi_n^{x,y,z}(\mathbf{R},\mathbf{r}), \tag{58}$$

here, the normalized wave functions

$$\Psi_n^{x,y,z}(\mathbf{R},\mathbf{r}) = \left(Q_n^{x,y,z}(R)\right)^{-1/2} \times \sum_{lm}\psi^*_{nlm}(\mathbf{r})\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\{x,y,z\}} \tag{59}$$

are linear combinations of the Coulomb wave functions $\psi_{nlm}(\mathbf{r})$ with the same principal quantum number $n$ and, hence, with the same energy. The expansion coefficients and normalization factors are

$$C_n^{x,y,z}(E) = -\frac{N_0^{(1)}\sqrt{3\pi}}{\gamma}\frac{B_{x,y,z}(E)\sqrt{Q_n^{x,y,z}(R)}}{E_n - E}, \tag{60}$$

$$Q_n^z(R) \equiv \sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial z}\right|^2_{z=R} = \sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial R}\right|^2,$$

$$Q_n^{x,y}(R) \equiv \sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\{x,y\}}\right|^2_{x=y=0} \tag{61}$$

$$= \frac{1}{R^2}\sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\theta}\right|^2_{\theta=0},$$

with the last two sums with subscripts $x$ and $y$ being equal.

For a complete basis of adiabatic wave functions to be constructed (see the next section), function (59) must be expressed in terms of the sums of Coulomb functions $\psi_{nlm}$ in the coordinate system with the $z'$ axis perpendicular to the collision plane and with the $x'$ axis directed either along the impact parameter (rectilinear trajectories) or along the vector of the smallest separation (Coulomb trajectories, time $t = 0$). This system of $x'$, $y'$, $z'$ coordinates does not rotate during a collision. Note that the sum $\sum_{lm}\psi^*_{nlm}(\mathbf{r})\psi_{nlm}(\mathbf{R})$ does not depend on the coordinate-system orientation (see above). Therefore, to pass to the new system, we need only to redetermine the derivatives. Since the $x$ axis of the rotating coordinate system used in this section coincides with the $z$ axis of the new nonrotating coordinate system, it is clear that $(d/dx) = (d/dz')$. For the other two pairs of coordinates, we have the following obvious relations:

$$\frac{d}{dz} = \cos\phi_R\frac{d}{dx'} + \sin\phi_R\frac{d}{dy'},$$

$$\frac{d}{dy} = -\sin\phi_R\frac{d}{dx'} + \cos\phi_R\frac{d}{dy'}.$$

Expressing the derivatives with respect to $x'$ and $y'$ in terms of the derivatives with respect to spherical coordinates $R$, $\theta$, and $\phi$, we obtain the function $\Psi_n^{x,y,z}(\mathbf{R},\mathbf{r})$ in the new nonrotating coordinate system:

$$\Psi_n^x(\mathbf{R},\mathbf{r}) = \sum_{lm}\psi^*_{nlm}(\mathbf{r})J^x_{nlm}(\mathbf{R}),$$

$$J^x_{nlm}(\mathbf{R}) \equiv \frac{\partial\psi_{nlm}(\mathbf{R})/\partial\theta_R}{R\sqrt{Q_n^x(R)}},$$

$$\Psi_n^y(\mathbf{R},\mathbf{r}) = \sum_{lm}\psi^*_{nlm}(\mathbf{r})J^y_{nlm}(\mathbf{R}),$$

$$J^y_{nlm}(\mathbf{R}) \equiv \frac{\partial\psi_{nlm}(\mathbf{R})/\partial\phi_R}{R\sqrt{Q_n^y(R)}}, \tag{62}$$

$$\Psi_n^z(\mathbf{R},\mathbf{r}) = \sum_{lm}\psi^*_{nlm}(\mathbf{r})J^z_{nlm}(\mathbf{R}),$$

$$J^z_{nlm}(\mathbf{R}) \equiv \frac{\partial\psi_{nlm}(\mathbf{R})/\partial R}{\sqrt{Q_n^z(R)}},$$

where $\theta_R$ and $\phi_R$ are the spherical angles of vector $\mathbf{R}$; the polar angle $\theta_R$ is constant and equal to $\pi/2$ during the entire collision, because the $z'$ axis is perpendicular to the collision plane. The normalization factors $Q_n^{x,y,z}(R)$ are

$$Q_n^x(R) = \frac{1}{R^2}\sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\theta_R}\right|^2,$$

$$Q_n^2(R) = \sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial R}\right|^2, \tag{63}$$

$$Q_n^y(R) = \frac{1}{R^2}\sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\phi_R}\right|^2 = \frac{1}{R^2}\sum_{lm}m^2|\psi_{nlm}(\mathbf{R})|^2$$

and do not depend on the coordinate system (see Sect. 4).

Since the wave functions $\Psi_n^{x,y}(\mathbf{R},\mathbf{r})$ belong to the states with nonzero components of the orbital angular momentum along the $\mathbf{R}$ axis, these functions are zero if vector $\mathbf{r}$ is directed along vector $\mathbf{R}$. Indeed, the azimuthal angles of these vectors are $\pi/2$ in this case. The associated Legendre polynomials in (7) and their derivatives at $\theta = \pi/2$ are [44]

$$N_{lm}P_l^{|m|}(\cos\theta)\big|_{\cos\theta=0} = N_{lm}\frac{2^{|m|}}{\sqrt{\pi}} \times \frac{\Gamma[(l+|m|+1)/2]}{\Gamma[1+(l-|m|)/2]}\sin\left[\frac{\pi}{2}(l+|m|+1)\right], \tag{64}$$

$$N_{lm}\frac{dP_l^{|m|}(\cos\theta)}{d\cos\theta}\bigg|_{\cos\theta = 0} = N_{lm}\frac{2^{|m|+1}}{\sqrt{\pi}} \times \frac{\Gamma[1+(l-|m|)/2]}{\Gamma[(l-|m|+1)/2]}\sin\left[\frac{\pi}{2}(l+|m|)\right]. \qquad (65)$$

We see from the definition of $\Psi_n^x(\mathbf{R}, \mathbf{r})$ that for $\mathbf{r} \parallel \mathbf{R}$, this function is proportional to the product of expressions (64) and (65). The latter product, in turn, is proportional to the product of sines

$$\sin\left[\frac{\pi}{2}(l+|m|+1)\right]\sin\left[\frac{\pi}{2}(l+|m|)\right] = 0, \qquad (66)$$

which is zero for any integer $l$ and $m$. Thus, in the sum defining $\Psi_n^x$, either the function $\psi_{nlm}$ or its derivative $\partial\psi_{nlm}/\partial\theta$ are zero at $\theta_r = \theta_R = \pi/2$.

The sum defining $\Psi_n^y(\mathbf{R}, \mathbf{r})$ in (62) is also zero for $\mathbf{r} \parallel \mathbf{R}$, because the azimuthal angles of these vectors are equal in this case, $\phi_r = \phi_R$. This sum can be written as

$$\sum_{lm}\psi^*_{nlm}(\mathbf{r})\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\phi_R} = \frac{i}{2\pi}\sum_{l=0}^{n-1} f_{nl}(r)f_{nl}(R) \times \sum_{m=-l}^{m=+l} m\left|N_{lm}P_l^{|m|}(0)\right|^2 = 0. \qquad (67)$$

The expression under the modulus sign in the last sum depends only on the absolute value $|m|$. After multiplication by $m$, the contributions from $+|m|$ and $-|m|$ to this sum cancel out for any $|m|$, and, hence, the entire sum is zero.

Functions (62) are mutually orthogonal; the corresponding integrals are proportional to expressions (66) and (67), which are zero. The functions $\Psi_n^{x,y}(\mathbf{R}, \mathbf{r})$ for $L = 1$ are also orthogonal to $\Psi_n(\mathbf{R}, \mathbf{r})$ for $L = 0$, because they belong to the states with different components of the angular momentum along the internuclear axis. The integrals of $\Psi_n\Psi_n^x$ and $\Psi_n\Psi_n^y$ are proportional to the product of sines (66) and to the sum (67), respectively.

## 3. PROPERTIES OF THE COULOMB GREEN'S FUNCTION AND A COMPLETE BASIS OF ADIABATIC WAVE FUNCTIONS

A Coulomb system possesses a symmetry [43], which manifests itself in the degeneracy of its energy levels. This symmetry affects the properties of Coulomb Green's function, whose spectral expansion (6) can be written as

$$\tilde{G}(\mathbf{R}, \mathbf{r}, E) = \sum_n \frac{(Q_n(R))^{1/2}\Psi_n(\mathbf{R}, \mathbf{r})}{E_n - E}, \qquad (68)$$

where $\Psi_n(\mathbf{R}, \mathbf{r})$ are given by formula (32).

For each principal quantum number $n$, linear combinations can be constructed from $n^2$ wave functions of degenerate states, which will be the wave functions in the new representation. One of these functions is $\Psi_n(\mathbf{R}, \mathbf{r})$ in the spectral expansion of the ionic function (31). The remaining functions, designated as $\Psi_{nlm}(\mathbf{R}, \mathbf{r})$,

$$\Psi_{nlm}(\mathbf{R}(t), \mathbf{r}) = \sum_{l', m'} C^{n, l, m}_{n, l', m'}(\mathbf{R})\psi_{n, l', m'}(\mathbf{r}), \qquad (69)$$

are absent from (31). The functions $\Psi_{nlm}(\mathbf{R}, \mathbf{r})$ must be orthogonal between themselves and orthogonal to $\Psi_n$. Consequently, the degenerate Coulomb basis can be reconstructed in such a way that only one of the new functions, $\Psi_n(\mathbf{R}, \mathbf{r})$, will be present in the spectral expansion of Green's function (68), while $\Psi_{nlm}(\mathbf{R}, \mathbf{r})$ will not be present in this expansion.

In our problem, since there are no $\Psi_{nlm}(\mathbf{R}, \mathbf{r})$ in the spectral expansion of Green's function, the ionic term interacts only with one covalent state, while the energies of the remaining degenerate states $\Psi_{nlm}$ do not change and are equal to the unperturbed Coulomb energies, $E_n^0 = -Z^2/2n^2$. Nevertheless, the latter states are not absolutely passive. They can be populated during captures (1) and (2), because their wave functions $\Psi_{nlm}(\mathbf{R}(t), \mathbf{r})$ are time dependent and because the matrix element of the time derivative between them and $\Psi_n(\mathbf{R}, \mathbf{r})$ is nonzero. Thus, the complete basis of adiabatic states includes both $\Psi_n$ and $\Psi_{nlm}$ states, and our goal now is to construct the functions $\Psi_{nlm}$.

The possibility of reconstructing the Coulomb basis of eigenfunctions to study electron scattering by a system composed of many small potential wells was explored in [46], but the specific algorithm for constructing an orthonormal basis was not discussed in previous papers [33–36, 46].

To find the reconstructed orthonomal Coulomb basis of eigenfunctions, we assume that one of Coulomb functions (7), $\Psi_{n\lambda m}$, is orthogonal to $\Psi_n$ from the outset, so that $J_{n\lambda\mu}(\mathbf{R}) = 0$. Let us consider the combinations

$$L = 0, \quad \Psi_{nlm}(\mathbf{R}, \mathbf{r}) = \psi_{nlm}(\mathbf{r}) - J_{nlm}(\mathbf{R})[\Psi_n(\mathbf{R}, \mathbf{r}) + \psi_{n\lambda\mu}(\mathbf{r})], \quad \{lm\} \neq \{\lambda\mu\}, \qquad (70)$$

$$L = 1, \quad \Psi^{x,y,z}_{nlm}(\mathbf{R}, \mathbf{r}) = \psi_{nlm}(\mathbf{r}) - J^{x,y,z}_{nlm}(\mathbf{R})[\Psi^{x,y,z}_n(\mathbf{R}, \mathbf{r}) + \psi_{n\lambda\mu}(\mathbf{r})]. \qquad (71)$$

Each of these functions is orthogonal to $\Psi_n$ ($L = 0$), or to $\Psi_n^{x,y,z}$ ($L = 1$). For the mutual orthogonality, for example, of $\Psi_{nlm}(\mathbf{R}, \mathbf{r})$ to be established, we must calculate the integral of their product

$$\int\Psi_{nlm}\Psi_{nl'm'}d\mathbf{r} = \delta_{ll'}\delta_{mm'} - 2J_{nlm}J^*_{nl'm'} + J_{nlm}J^*_{nl'm'}\int\left|\Psi_n(\mathbf{R}, \mathbf{r}) + \psi_{n\lambda\mu}(\mathbf{r})\right|^2 d\mathbf{r}. \qquad (72)$$

Since $\psi_{n\lambda\mu}(\mathbf{r})$ is orthogonal to $\Psi_n(\mathbf{R}, \mathbf{r})$, the integral on the right-hand side of (72) is equal to 2 ($\psi_{n\lambda\mu}$and $\Psi_n$ are normalized to unity), and the sum of the second and third terms in (72) is zero. The same is true for $L = 1$; therefore, each of the functions (70) and (71) is normalized to unity and orthogonal to all the other functions:

$$\int \Psi^*_{nlm}(\mathbf{R}, \mathbf{r})\Psi_{nl'm'}(\mathbf{R}, \mathbf{r})d\mathbf{r} = \delta_{ll'}\delta_{mm'}, \qquad (73)$$

$$\int (\Psi^{x, y, z}_{nlm}(\mathbf{R}, \mathbf{r}))^*\Psi^{x, y, z}_{nl'm'}(\mathbf{R}, \mathbf{r})d\mathbf{r} = \delta_{ll'}\delta_{mm'}. \qquad (74)$$

At $l = \lambda$ and $m = \mu$, the equality $\Psi_{n\lambda\mu} = \psi_{n\lambda\mu}$ holds for the functions (70), because $\psi_{n\lambda\mu}$ is orthogonal to $\Psi_n$ ($L = 0$), and the second term in (70) is zero. Therefore, the function (70) with orbital quantum numbers $\lambda$ and $\mu$ is not orthogonal to the functions (70) with $l \neq \lambda$ and $\mu \neq m$. This is also true for $L = 1$. Consequently, for a given principal quantum number $n$, the number of orthonormal functions $\Psi_{nlm}$ or $\Psi^{x, y, z}_{nlm}$ is $n^2 - 1$. Together with $\Psi_n$, or $\Psi^{x, y, z}_n$, the total number of functions is $n^2$, as must be the case.

The proposed orthogonalization method, formulas (70) and (71), is general. It is based on the existence of $\psi_{n\lambda\mu}$, which is orthogonal to the active-state function $\Psi_n$ from the outset. For this method to be applicable to our problem, it must be shown that $\psi_{n\lambda\mu}$ actually exists.

In the introduced coordinate system with the $z$ axis perpendicular to the collision plane, the polar angle of vector $\mathbf{R}$ is a constant during the collision and is equal to $\theta_R = \pi/2$. The functions $\psi_{nlm}(\mathbf{R})$ and their derivatives with respect to $x$ and $y$ are proportional to the associated Legendre polynomials at $\cos\theta = 0$, while the derivatives with respect to $z$ are proportional to the derivatives of these polynomials. These polynomials and their derivatives at $\theta = \pi/2$ are given by (64) and (65), respectively. We thus see that $P_l^{|m|}(0)$ as well as $\psi_{nlm}(\mathbf{R})$ and its derivatives with respect to $x$ and $y$ are zero at even $l + |m| + 1$, while the derivative with respect to $z$ is zero at even $l + |m|$ (or odd $l + |m| + 1$). The integrals $J_{nlm}(\mathbf{R})$, formulas (32a) and (62), are zero at these orbital quantum numbers $l$ and $m$. Consequently, the Coulomb functions $\psi_{nlm}(\mathbf{r})$ with even $l + |m| + 1$ are orthogonal to $\Psi_n$ and $\Psi^{x, y}_n$ but not orthogonal to $\Psi^z_n$. At odd $l + |m| + 1$, $\psi_{nlm}$ are orthogonal to $\Psi^z_n$ and not orthogonal to $\Psi_n$ and $\Psi^{x, y}_n$.

Thus, in our problem, the function $\psi_{n\lambda\mu}(\mathbf{r})$ exists and is not unique. For each $n$, the number of such functions is approximately half the number of all degenerate states, i.e., $\approx n^2/2$.

As we see, the solution of the orthogonalization problem depends on the choice of the coordinate system. In the coordinate system we chose, this solution is simplest, because the polar angle $\theta_R$ of vector $\mathbf{R}$ is $\pi/2$ during the entire collision for both rectilinear and curvilinear Coulomb trajectories. The trajectory must be plane, which is the case for central forces. The proposed method of constructing a complete orthonomal basis is based on the specific form of active-state functions $\Psi_n(\mathbf{R}, \mathbf{r})$.

Each principal quantum number $n$ has its own set of orthonormal functions $\{\Psi_n, \Psi_{nlm}\}$. Any function from set $n$ is orthogonal to any function from set $n'$ ($n' \neq n$), because these sets are constructed from different sets of Coulomb functions, respectively, from $\psi_{n'lm}$ and $\Psi_{nlm}$, which are orthogonal between themselves.

The functions $\Psi_{nlm}$ given by (70) are zero at $\mathbf{r} = \mathbf{R}$: $\Psi_{nlm}(\mathbf{r} = \mathbf{R}) = 0$. This fact is determined by the specific form of active-state function $\Psi_n$ and by its orthogonality to $\Psi_{nlm}$; i.e., it results from the degeneracy of Coulomb energy levels attributable to Coulomb-field symmetry [43, 47].

The ionic wave function (31) is constructed from the wave functions of only active states. Consequently, the wave function of any passive state $\psi_{nlm}(\mathbf{R}, \mathbf{r})$ is orthogonal to the ionic function (31). By contrast, any active-state function $\Psi_n(\mathbf{R}, \mathbf{r})$ ($L = 0$), is not orthogonal to the ionic function $\Psi(\mathbf{R}, \mathbf{r})$ [see formula (31)]. The integral of their product is equal to the coefficient $C_n(E)$ [see formula (33)]. Based on the same method used to construct the orthonormal system of functions $\{\Psi_n, \Psi_{nlm}\}$ given by formulas (70) and (71), we construct the functions

$$\begin{aligned} L = 0, \quad \tilde{\Psi}_n(\mathbf{R}, \mathbf{r}) &= \Psi_n(\mathbf{R}, \mathbf{r}) \\ &- C_n(E)[\Phi(\mathbf{R}, \mathbf{r}) + \psi_{n\lambda'\mu'}(\mathbf{r})], \end{aligned} \qquad (75)$$

$$\begin{aligned} L = 1, \quad \tilde{\Psi}^{x, y, z}_n(\mathbf{R}, \mathbf{r}) &= \Psi^{x, y, z}_n(\mathbf{R}, \mathbf{r}) \\ &- C^{x, y, z}_n(E)[\Phi_{x, y, z}(\mathbf{R}, \mathbf{r}) + \psi_{n\lambda'\mu'}(\mathbf{r})], \end{aligned} \qquad (76)$$

where $\psi_{n\lambda'\mu'}$ is orthogonal to $\Psi_n$, $\Phi$ ($L = 0$) or to $\Psi^{x, y, z}_n$, $\Phi_{x, y, z}$ ($L = 1$) (the orbital quantum numbers $\lambda'$ and $\mu'$ differ from $\lambda$ and $\mu$). The functions $\tilde{\Psi}_n$ and $\tilde{\Psi}^{x, y, z}_n$ are orthogonal between themselves and orthogonal to the ionic functions $\Phi$ and $\Phi_{x, y, z}$ and to the functions of passive states $\Psi_{nlm}$ or $\Psi^{x, y, z}_{nlm}$ ($\lambda \neq \lambda'$, $\mu \neq \mu'$). In addition, the functions (75) and (76) are normalized, because

$$\int |\tilde{\Psi}_n(\mathbf{R}, \mathbf{r})|^2 d\mathbf{r} = \int |\Psi_n(\mathbf{R}, \mathbf{r})|^2 d\mathbf{r} = 1, \qquad (77)$$

and a similar relation holds for $L = 1$.

Next, we must calculate the adiabatic matrix elements of the time derivative. Since the $z$ component of vector $\mathbf{R}(t)$ is zero,

$$\frac{d}{dt} = \dot{R}\frac{\partial}{\partial R} + \dot{\phi}_R\frac{\partial}{\partial \phi_R}.$$

The matrix element between $\Psi_{nlm}$ and $\Psi_n$ is then given by

$$L = 0,$$

$$V_n^{nlm}(\mathbf{R}) \equiv \int \Psi_{nlm}(\mathbf{R}, \mathbf{r}) \frac{d\Psi_n(\mathbf{R}, \mathbf{r})}{dt} d\mathbf{r}$$

$$= \frac{dJ_{nlm}(\mathbf{R})}{dt} = \frac{1}{\sqrt{Q_n(R)}} \tag{78}$$

$$\times \left[ im\dot{\phi}_R \psi_{nlm}(\mathbf{R}) + \dot{R}\left( \frac{d\psi_{nlm}(\mathrm{R})}{dR} + Z\psi_{nlm}(\mathbf{R}) \frac{\psi_{n0}^2(R)}{Q_n(R)} \right) \right],$$

and does not depend on quantum numbers λ and μ. When calculating the matrix element (78), we used the expression for the derivative

$$\frac{dQ_n}{dR} = -2Z\psi_{n0}^2(R),$$

which follows from the definition of $Q_n(R)$ in formula (38). For $L = 1$, these matrix elements are

$$L = 1,$$

$$V_{n,\,nlm}^{x,\,y,\,z}(\mathbf{R}) = \int \Psi_{nlm}^{x,\,y,\,z}(\mathbf{R}, \mathbf{r}) \frac{\partial \Psi_n^{x,\,y,\,z}(\mathbf{R}, \mathbf{r})}{\partial t} d\mathbf{r} \tag{79}$$

$$= \left( \dot{\phi}_R \frac{\partial}{\partial \phi_R} + \dot{R} \frac{\partial}{\partial R} \right) J_{nlm}^{x,\,y,\,z}(\mathbf{R})$$

and do not depend on quantum numbers λ and μ either. The derivatives with respect to azimuthal angle $\phi_R$ are

$$\frac{\partial J_{nlm}^z(\mathbf{R})}{\partial \phi_R} = im \frac{\dfrac{\partial \psi_{nlm}(\mathbf{R})}{\partial R}}{\sqrt{Q_n^z(R)}},$$

$$\frac{\partial J_{nlm}^x(\mathbf{R})}{\partial \phi_R} = im \frac{\dfrac{\partial \psi_{nlm}(\mathbf{R})}{\partial \theta_R}}{R\sqrt{Q_n^x(R)}}, \tag{80}$$

$$\frac{\partial J_{nlm}^y(\mathbf{R})}{\partial \phi_R} = -\frac{m^2 \psi_{nlm}(\mathbf{R})}{R\sqrt{Q_n^y(R)}},$$

and the derivatives with respect to $R$ are

$$\frac{\partial J_{nlm}^x(\mathbf{R})}{\partial R} = \frac{1}{\sqrt{Q_n^x(R)}}$$

$$\times \left\{ Z \frac{Q_n(R) - \psi_{n0}^2(R)}{2R^3 Q_n^x(R)} - \frac{1}{R^2} + \frac{1}{R}\frac{\partial}{\partial R} \right\} \frac{\partial \psi_{nlm}(\mathbf{R})}{\partial \theta_R}, \tag{81}$$

$$\frac{\partial J_{nlm}^y(\mathbf{R})}{\partial R} = \frac{im}{\sqrt{Q_n^y(R)}}$$

$$\times \left\{ Z \frac{Q_n(R) - \psi_{n0}^2(R)}{2R^3 Q_n^y(R)} - \frac{1}{R^2} + \frac{1}{R}\frac{\partial}{\partial R} \right\} \psi_{nlm}(\mathbf{R}), \tag{82}$$

$$\frac{\partial J_{nlm}^z(\mathbf{R})}{\partial R} = \frac{-2}{\sqrt{Q_n^z(R)}}$$

$$\times \left\{ E_n + \frac{Z}{R} - \frac{l(l+1)}{2R^2} + \left( \frac{1}{R} - Z\frac{(\psi'_{n0}(R))^2}{2Q_n^z(R)} \right) \frac{\partial}{\partial R} \right\} \psi_{nlm}(\mathbf{R}). \tag{83}$$

The sums $Q_n^{x,\,y,\,z}(R)$ and their derivatives with respect to $R$ are calculated in the next section.

## 4. THE SUMS OF PRODUCTS OF THE COULOMB WAVE FUNCTIONS OVER DEGENERATE STATES

In this section, we calculate the sums of products of the Coulomb wave functions with negative energy, which are present in the wave functions of covalent states constructed in the preceding section. The products are summed over orbital quantum numbers $l$ and $m$. Based on the analysis of Green's function by Hostler and Pratt [37, 38], the authors of [39–41] calculated the sum

$$\hat{Q}_n(\mathbf{r}, \mathbf{R}) \equiv \sum_{l,\,m} \psi_{nlm}^*(\mathbf{r}) \psi_{nlm}(\mathbf{R})$$

$$= \frac{4Z^2}{n^2} \frac{\phi'_{n0}(\tau_y)\phi_{n0}(\tau_x) - \phi_{n0}(\tau_y)\phi'_{n0}(\tau_x)}{\tau_x - \tau_y}, \tag{84}$$

$$\tau_{x,\,y} = \frac{Z}{n}[r + R \pm |\mathbf{r} - \mathbf{R}|],$$

which is equal to the wave function (32), to within the normalization factor. In his papers [47] devoted to four-dimensional symmetry of the hydrogen atom, Fock studied a similar sum, but for the wave functions in momentum representation.

By analyzing a Taylor expansion of the sum (84), we can obtain the sums of products of the derivatives of arbitrary orders of Coulomb eigenfunctions with respect to the absolute value of $\mathbf{R}$ at $\mathbf{r} = \mathbf{R}$:

$$Q_n^{i,\,j}(R) = \sum_{l,\,m} \frac{d^i \psi_{nlm}^*(\mathbf{R})}{dR^i} \frac{d^j \psi_{nlm}(\mathbf{R})}{dR^j}. \tag{85}$$

To calculate the sum (85), let us consider the limit of relation (84) for $\mathbf{r} \longrightarrow \mathbf{R}$ when point $\mathbf{r}$ moves along vector $\mathbf{R}$ and when

$$\tau_x = \frac{2ZR}{n} = \text{const}, \quad \tau_y = \frac{2Zr}{n},$$

$$\tau_x - \tau_y = \frac{2Z(R - r)}{n}.$$

The expansion of the sum from (84) in a Taylor series in powers $(r - R)^n$ is

$$\sum_{l,m} \psi^*_{nlm}(\mathbf{r})\psi_{nlm}(\mathbf{R}) \approx \sum_{l,m} |\psi_{nlm}(\mathbf{R})|^2 + (r - R)\sum_{l,m} \frac{d\psi_{nlm}(\mathbf{R})}{dR}\psi^*_{nlm}(\mathbf{R}) + \ldots + \frac{(r-R)^k}{k!}\sum_{l,m} \frac{d^k\psi_{nlm}(\mathbf{R})}{dR^k}\psi^*_{nlm}(\mathbf{R}) + \ldots, \tag{86}$$

and the expansion of the right-hand part of (84) is

$$\frac{\phi'_{n0}(r)\phi_{n0}(R) - \phi_{n0}(r)\phi'_{n0}(R)}{R - r} = \left(\frac{d\phi_{n0}(R)}{dR}\right)^2 - \phi_{n0}(R)\frac{d^2\phi_{n0}(R)}{dR^2} + \ldots + \frac{(r-R)^{k-1}}{k!} \times \left(\phi'_{n0}(R)\frac{d^k\phi_{n0}(R)}{dR^k} - \phi_{n0}(R)\frac{d^{k+1}\phi_{n0}(R)}{dR^{k+1}}\right) + \ldots . \tag{87}$$

Comparing these expansions and equating the terms at the same powers $(r - R)^k$ yields

$$Q_n^{(0,j)}(R) \equiv \sum_{lm} \psi^*_{nlm}(\mathbf{R})\frac{d^j\psi_{nlm}(\mathbf{R})}{dR^j} = \frac{1}{j+1}(\phi_{n0}^{(j+1)}(R)\phi_{n0}^{(1)}(R) - \phi_{n0}^{(j+2)}(R)\phi_{n0}(R)), \tag{88}$$

where $\phi^{(j)} \equiv d^j\phi/d^jR$. For the special cases $j = 0$ and 1, we obtain the two previously calculated sums [39–41] from (88): the sum (38) and its derivative

$$\frac{dQ_n(R)}{dR} = -2Z\psi^2_{n0}(R).$$

Calculating the higher derivatives of $\phi^j_{n0}$ with respect to $R$ when differentiating the wave equation (18), we find the sums $Q_n^{(0,j)}$ for the special cases $j = 2$–5:

$$Q_n^{(0,2)}(R) \equiv \sum_{lm} \psi^*_{nlm}(\mathbf{R})\frac{d^2\psi_{nlm}(\mathbf{R})}{dR^2}$$

$$= \sum_l \frac{2l+1}{4\pi} f_{nl}(R)\frac{d^2 f_{nl}}{dR^2} = -\frac{2}{3}\left(E_n + \frac{Z}{R}\right)Q_n^{(0,0)}(R) + \frac{2Z}{3R^3}(2\phi^2_{n0}(R) - R\phi_{n0}(R)\phi'_{n0}(R)), \tag{89}$$

$$Q_n^{(0,3)}(R) \equiv \sum_{lm} \psi^*_{nlm}(\mathbf{R})\frac{d^3\psi_{nlm}(\mathbf{R})}{dR^3}$$

$$= \sum_l \frac{2l+1}{4\pi} f_{nl}(R)\frac{d^3 f_{nl}}{dR^3} = \frac{Z}{R^2}\left\{Q_n^{(0,0)}(R) + \left[2\left(E_n + \frac{Z}{R}\right) - \frac{3}{R^2}\right]\phi^2_{n0}(R) + \frac{2}{R}\phi_{n0}(R)\phi'_{n0}(R)\right\}, \tag{90}$$

$$Q_n^{(0,4)}(R) \equiv \sum_{lm} \psi^*_{nlm}(\mathbf{R})\frac{d^4\psi_{nlm}(\mathbf{R})}{dR^4}$$

$$= \sum_l \frac{2l+1}{4\pi} f_{nl}(R)\frac{d^4 f_{nl}}{dR^4} = \frac{4}{5}\left[\left(E_n + \frac{Z}{R}\right)^2 - \frac{3Z}{R^3}\right]Q_n^{(0,0)}(R) - \frac{32Z}{5R^3}\left[E_n + \frac{3Z}{2R} - \frac{3}{2R^2}\right]\phi^2_{n0}(R) + \frac{8Z}{5R^2}\left[E_n + \frac{Z}{R} - \frac{9}{2R^2}\right]\phi_{n0}(R)\phi'_{n0}(R), \tag{91}$$

$$Q_n^{(0,5)}(R) \equiv \sum_{lm} \psi^*_{nlm}(\mathbf{R})\frac{d^5\psi_{nlm}(\mathbf{R})}{dR^5}$$

$$= \sum_l \frac{2l+1}{4\pi} f_{nl}(R)\frac{d^5 f_{nl}}{dR^5}$$

$$= -\frac{4Z}{R^2}\left\{\left[E_n + \frac{Z}{R} - \frac{2}{R^2}\right]Q_n^{(0,0)}(R) + \frac{1}{R}\left[2E_n + \frac{3Z}{R} - \frac{8}{R^2}\right]\phi_{n0}(R)\phi'_{n0}(R) + \left[\left(E_n + \frac{Z}{R}\right)\left(E_n + \frac{Z}{R} - \frac{7}{R^2}\right) - \frac{5}{R^3}\left(Z - \frac{2}{R}\right)\right]\phi^2_{n0}(R)\right\}. \tag{92}$$

Differentiating equality (88) with respect to $R$, we derive the relation

$$\sum_{lm} \psi^{*(1)}_{nlm}(\mathbf{R})\psi^{(j)}_{nlm}(\mathbf{R}) + \sum_{lm} \psi^*_{nlm}(\mathbf{R})\psi^{(j+1)}_{nlm}(\mathbf{R}) = \frac{1}{j+1}(\phi^{(2)}_{n0}\phi^{(j+1)}_{n0} - \phi_{n0}\phi^{(j+3)}_{n0})\Big|_R . \tag{93}$$

The second sum in the left-hand part of this relation is equal to the sum (88) with the change $j \longrightarrow j + 1$, so the sums for $i = 1$ are

$$Q_n^{(1,j)}(R) \equiv \sum_{lm} \frac{d\psi^*_{nlm}}{dR}\frac{d^j}{dR^j}\psi_{nlm}(\mathbf{R}) = \frac{\phi^{(2)}_{n0}(R)\phi^{(j+1)}_{n0}(R)}{j+1} - \frac{\phi^{(1)}_{n0}(R)\phi^{(j+2)}_{n0}(R)}{j+2} - \frac{\phi_{n0}(R)\phi^{(j+3)}_{n0}(R)}{(j+1)(j+2)}. \tag{94}$$

A similar differentiation of Eq. (94) with respect to $R$ yields the sums $Q_n^{(i,\,j)}$ for $i = 2$:

$$Q_n^{(2,\,j)}(R) \equiv \sum_{lm} \frac{d^2\psi^*_{nlm}}{dR^2}\frac{d^j}{dR^j}\psi_{nlm}(\mathbf{R})$$

$$= \frac{\phi_{n0}^{(3)}(R)\phi_{n0}^{(j+1)}(R)}{j+1} - \frac{j\phi_{n0}^{(2)}(R)\phi_{n0}^{(j+2)}(R)}{(j+1)(j+2)} \tag{95}$$

$$-\frac{2\phi_{n0}^{(1)}(R)\phi_{n0}^{(j+3)}(R)}{(j+1)(j+3)} - \frac{2\phi_{n0}(R)\phi_{n0}^{(j+4)}(R)}{(j+1)(j+2)(j+3)}.$$

By continuing the differentiation, we can calculate the sum $Q_n^{(i,\,j)}$ for any $i$ and $j$, but the result rapidly becomes unwieldy. Below, we give the sums for the most interesting particular values of $i$ and $j$. We have

$$Q_n^{(1,\,1)}(R) = \sum_{lm}\left|\frac{d\psi_{nlm}(\mathbf{R})}{dR}\right|^2 = \sum_{l}\frac{2l+1}{4\pi}\left(\frac{df_{nl}(R)}{dR}\right)^2$$

$$= \frac{2}{3}\left(E_n + \frac{Z}{R}\right)Q_n^{(0,\,0)}(R) \tag{96}$$

$$+\frac{2Z}{3R^3}(\phi_{n0}^2(R) - 2R\phi_{n0}(R)\phi'_{n0}(R)),$$

for $i = j = 1$ and

$$Q_n^{(2,\,2)}(R) = \sum_{lm}\left|\frac{d^2\psi_{nlm}(\mathbf{R})}{dR^2}\right|^2 = \sum_{l}\frac{2l+1}{4\pi}\left(\frac{d^2f_{nl}(R)}{dR^2}\right)^2$$

$$= \frac{4}{5}\left\{\left[\left(E_n + \frac{Z}{R}\right)^2 + \frac{2Z}{R^3}\right]Q_n^{(0,\,0)}(R)\right.$$

$$-\frac{Z}{R^3}\left[8\left(E_n + \frac{Z}{R}\right) - \frac{Z}{R} - \frac{2}{R^2}\right]\phi_{n0}^2(R) \tag{97}$$

$$\left.+\frac{2Z}{R^2}\left(E_n + \frac{Z}{R} - \frac{2}{R^2}\right)\phi_{n0}(R)\phi'_{n0}(R)\right\},$$

for $i = j = 2$.

Differentiating (96) with respect to $R$ yields a simple expression for the sum with $i = 1$ and $j = 2$:

$$Q_n^{(1,\,2)}(R) \equiv \sum_{lm}\frac{d\psi^*_{nlm}}{dR}\frac{d^2\psi_{nlm}}{dR^2}$$

$$= \frac{1}{2}\frac{d}{dR}Q_n^{(1,\,1)}(R) = -Z(\psi'_{n0}(R))^2. \tag{98}$$

Differentiating (97) with respect to $R$ also yields a simple expression for $i = 2$ and $j = 3$:

$$Q_n^{(2,\,3)}(R) \equiv \sum_{lm}\frac{d^2\psi^*_{nlm}(\mathbf{R})}{dR^2}\frac{d^3\psi_{nlm}(\mathbf{R})}{dR^3}$$

$$= \frac{1}{2}\frac{d}{dR}Q_n^{(2,\,2)}(R) = -Z(\psi''_{n0}(R))^2 - \frac{4Z^2}{5R^4}(\phi_{n0}^2(R))'. \tag{99}$$

Integrating (98) and (99), we write $Q_n$ $(\equiv Q_n^{(0,\,0)})$, $Q_n^{(1,\,1)}$, and $Q_n^{(2,\,2)}$ as

$$Q_n(R) = 2Z\int_R^{\infty}\psi_{n0}^2(r)dr,$$

$$Q_n^{(1,\,1)}(R) = 2Z\int_R^{\infty}(\psi'_{n0})^2dr, \tag{100}$$

$$Q_n^{(2,\,2)}(R) = 2Z\int_R^{\infty}\left((\psi''_{n0}(r))^2 + \frac{4Z}{5r^4}(\phi_{n0}^2(r))'\right)dr.$$

The sums $Q_n$, $Q_n^{(1,\,1)}$, and $Q_n^{(2,\,2)}$ are shown in Fig. 3 as a function of internuclear separation $R$. They have no zeros at finite $R$. All the calculated sums are of interest not only in analyzing collisions between negative and positive ions, but also in the physics of highly excited Rydberg states [48].

For the system $A^- + B^{Z+}$, where the negative ion has the orbital angular momentum $L = 1$, the sums of products of the Coulomb functions (62), which define the wave functions of active covalent states, are calculated in a way similar to that for $L = 0$. It is necessary to analyze the limit of Coulomb Green's function for $E \longrightarrow E_n$. We obtained for the functions (62)

$$\Psi_n^z(\mathbf{R},\mathbf{r}) \equiv \sum_{l,\,m}\frac{\psi^*_{nlm}(\mathbf{r})}{\sqrt{Q_n^z(R)}}\frac{d\psi_{nlm}(\mathbf{R})}{dR}$$

$$= -\frac{4Z^2\cos\theta_b}{n^2|\mathbf{r}-\mathbf{R}|\sqrt{Q_n^z(R)}} \tag{101}$$

$$\times\left[\frac{\phi'_{n0}(\tau_y)\phi_{n0}(\tau_x) - \phi_{n0}(\tau_y)\phi'_{n0}(\tau_x)}{\tau_y - \tau_x}\right.$$

$$\left.+\phi'_{n0}(\tau_x)\phi'_{n0}(\tau_y) + \left(-\frac{1}{4} + \frac{n^2(R-r)}{2ZR(R-z)}\right)\phi_{n0}(\tau_x)\phi_{n0}(\tau_y)\right],$$

$$\Psi_n^{\{x,\,y\}}(\mathbf{R},\mathbf{r}) \equiv \sum_{l,\,m}\frac{\psi^*_{nlm}(\mathbf{r})}{\sqrt{Q_n^{x,\,y}(R)}}\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\{x,\,y\}}\Bigg|_{\mathbf{r}'=\mathbf{R}}$$

$$= -\frac{4Z^2\sin\theta_b\{\cos\theta_b, \sin\theta_b\}}{n^2|\mathbf{r}-\mathbf{R}|\sqrt{Q_n^{x,y}(R)}} \tag{102}$$
$$\times\left[\frac{\phi'_{n0}(\tau_y)\phi_{n0}(\tau_x)-\phi_{n0}(\tau_y)\phi'_{n0}(\tau_x)}{\tau_y-\tau_x}\right.$$
$$\left.+\,\phi'_{n0}(\tau_x)\phi'_{n0}(\tau_y)+\left(-\frac{1}{4}+\frac{n^2(r+R)}{2ZR(r+z)}\right)\phi_{n0}(\tau_x)\phi_{n0}(\tau_y)\right].$$

These functions have no poles at **r** = **R**.

To calculate the sums (63), many expansion terms for these functions must be analyzed in the limit $\mathbf{r} \longrightarrow \mathbf{R}$, which complicates the problem. Below, we therefore calculate the sums (63) by a different method.

The sum $Q_n^z(R)$, formula (63), is equal to the sum $Q_n^{(1,1)}$ calculated above, formulas (96) and (100). The sums $Q_n^{x,y}$, formula (63), are equal and can be written as two expressions:

$$Q_n^{x,y}(R) = \sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\{x,y\}}\right|_{z=0} \tag{103}$$
$$= \frac{1}{R^2}\sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\theta}\right|^2_{\theta=\pi/2} = \frac{1}{R^2}\sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\theta}\right|^2_{\theta=0}.$$

The values θ = π/2 and 0 correspond to the cases where the *z* axis is perpendicular to the collision plane and directed along *R*, respectively. The summation result does not depend on the choice of *z*-axis direction. We use the coordinate system with the *z* axis directed along vector **R**. It follows from the representation of the associated Legendre polynomial as the full hypergeometric function [44],

$$P_l^{|m|}(\cos\theta) = \frac{(-1)^{|m|}(l+|m|)!\sin^{|m|}(\theta)}{2^{|m|}|m|!(l-|m|)!} \tag{104}$$
$$\times F(|m|-l;\, l+|m|+1;\, |m|+1;\, \sin^2(\theta/2))$$

that the derivative with respect to angle θ at θ = 0 is nonzero only for the angular momentum components *m* = ±1 and is

$$\left.\frac{dP_l^1(\cos\theta)}{d\theta}\right|_{\theta=0} = -\frac{l(l+1)}{2}. \tag{105}$$

Using this result and the expression for normalization factor $N_{lm}$, (9), we transform the sum (103) to

$$Q_n^{x,y}(R) = \frac{1}{R^2}\sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\theta}\right|^2_{\theta=0} \tag{106}$$
$$= \frac{1}{8\pi R^2}\sum_l(2l+1)l(l+1)f_{nl}^2(R).$$

Calculating this sum can be reduced to calculating $\sum_{lm}l(l+1)|\psi_{nlm}(\mathbf{R})|^2$, which, using the summation theorem for spherical functions (35), can be written as

$$Q_n^{(l,l+1)}(R) \equiv \sum_{lm}l(l+1)|\psi_{nlm}(\mathbf{R})|^2 \tag{107}$$
$$= \frac{1}{4\pi}\sum_l(2l+1)l(l+1)f_{nl}^2(R).$$

The sums (106) and (107) differ by the *l*- and *m*-independent factor $(2R)^{-2}$. Having calculated (107), we therefore also determine the sum (106).

To calculate the sum (107), we express $l(l+1)f_{nl}$ in terms of $f_{nl}$, $df_{nl}/dR$, and $d^2f_{nl}/dR^2$ from the wave equation (8). Thereafter, (106) and (107) can be expressed in terms of the sums that have already been calculated above and take the form

$$Q_n^{(l,l+1)}(R) = \sum_{lm}l(l+1)|\psi_{nlm}(\mathbf{R})|^2$$
$$= \frac{1}{4\pi}\sum_l(2l+1)l(l+1)f_{nl}^2(R) \tag{108}$$
$$= \frac{4R^2}{3}\left(E_n+\frac{Z}{R}\right)Q_n(R)-\frac{2Z}{3R}[\phi_{n0}^2(R)+R\phi_{n0}(R)\phi'_{n0}(R)],$$

$$Q_n^{x,y}(R) = \frac{1}{R^2}\sum_{lm}\left|\frac{\partial\psi_{nlm}(\mathbf{R})}{\partial\theta}\right|^2_{\theta=0}$$
$$= \frac{1}{8\pi R^2}\sum_l(2l+1)l(l+1)f_{nl}^2(R) \tag{109}$$
$$= \frac{2}{3}\left(E_n+\frac{Z}{R}\right)Q_n(R)-\frac{Z}{3R^3}[\phi_{n0}^2(R)+R\phi_{n0}(R)\phi'_{n0}(R)].$$

The sum $Q_n^{x,y}$, formula (109), is shown in Fig. 4 as a function of internuclear separation *R* in comparison with $Q_n^z(R)$ [or $Q_n^{(1,1)}(R)$], formula (96). The two sums are positive for all *R*. However, whereas $Q_n^z$ decreases in steps in the classically permitted range of distances *R*, $Q_n^{x,y}(R)$ have indistinct bends instead of steps. Most of the other sums given above also exhibit a peculiar behavior as functions of *R*. All the sums were calculated both by using the derived quadratic forms of $\phi_{n0}$ and by direct summation over *l* using the Coulomb functions $f_{nl}$, (9). The results have always been in close agreement.

## 5. THE APPROXIMATION OF A DISTANT TERM CROSSING

In the preceding sections, the energy *E* and internuclear separation *R* of the $A^- + B^{Z+}$ system were assumed to be independent parameters. In this section, we calcu-

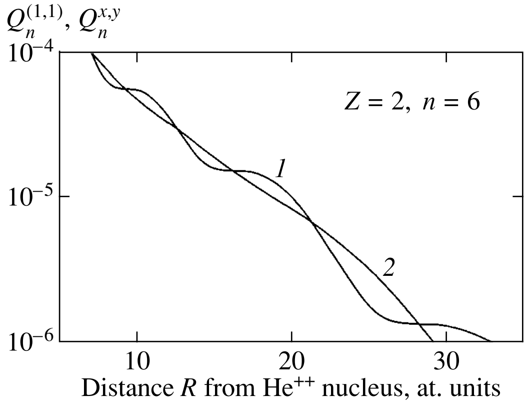

**Fig. 4.** Comparison of the sums $Q_n^{(1,1)}$ and $Q_n^{x,y}$ as functions of internuclear separation $R$: *1*—$Q_n^{(1,1)}(R)$, formulas (96), (100), and *2*—$Q_n^{x,y}(R)$, formula (109).

late the energy $E$ as a function of internuclear separation $R$. In the zero-order approximation, this energy follows the law (26), from which we derive the distances $R_n$ of term crossings:

$$\varepsilon_0 - \frac{Z}{R_n} = E_n, \longrightarrow R_n = \frac{Z}{\varepsilon_0 - E_n}. \tag{110}$$

For internuclear separations $R \approx R_n$ near distant term crossings, where $R_n$ is larger than the size of the corresponding Coulomb orbit for a covalent state, the matrix element between the ionic and covalent states is small compared to the difference between the Coulomb terms, $Z^2/n^3$. In this case, the two-level approximation may be used for the adiabatic wave function:

$$\begin{aligned}\Phi(\mathbf{R}, \mathbf{r}) &= \sum_{lm} C_n(E)\Psi_n(\mathbf{R}, \mathbf{r}) \\ &\approx B(E)\Phi_0(|\mathbf{R} - \mathbf{r}|) + C_n(E)\Psi_n(\mathbf{R}, \mathbf{r}).\end{aligned} \tag{111}$$

In this approximation, the ionic state $\Phi_0$ interacts with the active adiabatic state $\Psi_n(\mathbf{R}, \mathbf{r})$ that belongs to only one principal quantum number $n$. Separating the resonant term $\Psi_n$ from expansion (31) does not affect the form of $\Phi_0$, because the coefficients $C_n(E)$ are small far from a quasi-crossing. Using relations (33) and (60) between the coefficients $C_n$ and the function $B$, we obtain after normalizing the function (111)

$$\begin{gathered}1 = B^2(E) + C_n^2(E) = B^2(E)\left(1 + \frac{\Delta E_n^2(R)}{(E - E_n)^2}\right), \\ |E - E_n| \ll |E_n - E_{n\pm 1}| \approx Zn^{-3},\end{gathered} \tag{112}$$

because the overlap integral $\langle\Phi_0|\Psi_n\rangle$ is zero due to the orthogonality of $\Psi_n$ and $\Psi_{n'}$ at $n \neq n'$.

For negative ions with orbital angular momenta $L = 0$ and 1, $\Delta E_n(R)$ are

$$\Delta E_n(R) \equiv 2\pi N_0\sqrt{Q_n(R)}, \quad L = 0, \tag{113}$$

$$\Delta E_n^{x,y,z}(R) \equiv \frac{N_0^{(1)}\sqrt{3\pi}}{\gamma}\sqrt{Q_n^{x,y,z}(R)}, \quad L = 1. \tag{114}$$

Equalities (113) and (114) follow from formulas (33) and (60), while the sums $Q_n(R)$ and $Q_n^{x,y,z}(R)$ were calculated in the preceding section.

We derive the function $B(E)$ from (112). Substituting it in (33) and (60) yields $C_n(E)$ as functions of energy near term crossings:

$$B_n(E) = (E - E_n)[(E - E_n)^2 + \Delta E_n^2(R)]^{-1/2}, \tag{115}$$

$$C_n(E) = -\Delta E_n(R)[(E - E_n)^2 + \Delta E_n^2(R)]^{-1/2}, \tag{116}$$

$$\begin{gathered}C_{n,n'}(E) = -\Delta E_{n'}(R)\frac{(E - E_n)}{(E_n - E_{n'})} \\ \times [(E - E_n)^2 + \Delta E_n^2(R)]^{-1/2}, \\ E \approx E_n, \quad n' \neq n.\end{gathered} \tag{117}$$

Expressions (115) and (116) are the zero-order approximation of a distant crossing, while (117) is the next, first approximation. We see from (115) and (116) that near a quasi-crossing, where $|E - E_n| \approx |\Delta E_n(R_n)|$, the coefficients $C_n$ and $B_n$ are of the order of 1 in absolute value, while for the other covalent states $n' \neq n$ which the ionic term does not cross in the range of distances under consideration, the coefficients $C_{n,n'}(R_n)$ are small:

$$n' \neq n, \quad |C_{n,n'}(R_n)| \propto \left|\frac{\Delta E_{n'}(R_n)B_n}{E_n - E_{n'}}\right| \ll 1.$$

Let us write relations (115)–(117) by using the zero-order approximation for energy (26). For any $n$, the difference $E_0(R) - E_n$ can be written as

$$E_0(R) - E_n = Z\frac{R - R_n}{RR_n}, \tag{118}$$

which is an exact expression, not an approximation. Substituting (118) in (115)–(117) yields $B_m$, $C_n$, and $C_{n'}$ as functions of internuclear separation $R$ rather than energy:

$$B_n(R) = (R - R_n)[(R - R_n)^2 + \Delta R_n^2(R)]^{-1/2}, \tag{119}$$

$$C_n(R) = -\Delta R_n(R)[(R - R_n)^2 + \Delta R_n^2(R)]^{-1/2}, \tag{120}$$

$$\begin{gathered}C_{n'}(R) = -\Delta R_{n'}(R)\frac{(E_0(R) - E_n)}{(E_n - E_{n'})} \\ \times [(R - R_n)^2 + \Delta R_n^2(R)]^{-1/2},\end{gathered} \tag{121}$$

where

$$\Delta R_n(R) \equiv \frac{RR_n}{Z}\Delta E_n(R). \tag{122}$$

Note that at $R = R_n$,

$$\Delta R_n(R_n) = \frac{\Delta E_n(R_n)}{F_n},$$

where the force

$$F_n = \frac{Z}{R_n^2}.$$

Representing the coefficients $B$ and $C$ as functions of $R$ allows the two-level approximation under consideration to be compared with the exact calculation described in Section 1. In Fig. 5, the function $B(R)$ calculated from formula (119) of the two-level approximation is compared with the result of its exact calculation at $R \approx R_4$. The agreement between the results is excellent, so the two-level approximation proves to be very close to the exact calculation. For $R \approx R_5$ and $R \approx R_3$, the two-level approximation is equally close to the exact calculation.

The coefficients $B$ and $C$ in (115)-(117) are functions of energy. To determine the system's energy as a function of internuclear separation $R$, note that when the two-level approximation is considered in detail, the adiabatic wave functions $\psi_{1,2}$ prove to be linear combinations of the wave functions $\phi_{1,2}$ of the zero-order approximation [31, 43]:

$$\begin{gathered}\psi_1 = a\phi_1 + b\phi_2, \quad \psi_2 = -b\phi_1 + a\phi_2, \\ a = \left[\frac{\sqrt{x^2+\Delta^2}+x}{2\sqrt{x^2+\Delta^2}}\right]^{1/2}, \\ b = \left[\frac{\sqrt{x^2+\Delta^2}-x}{2\sqrt{x^2+\Delta^2}}\right]^{1/2},\end{gathered} \tag{123}$$

where, in our notation, the difference between the diagonal matrix elements $x$ and twice the nondiagonal matrix element $\Delta$ are

$$\begin{gathered}x \equiv E_1^0 + V_{11} - E_2^0 - V_{22} = E_0(R) - E_n, \\ \Delta \equiv 2|V_{0n}|.\end{gathered}$$

The equality $B_n = -C_n = 1/\sqrt{2}$ at $E - E_n = \Delta E_n(R)$ follows from (115) and (116). In turn, it follows from (123) that these coefficients correspond to a minimum energy difference between two quasi-crossing terms, when $x = 0$.

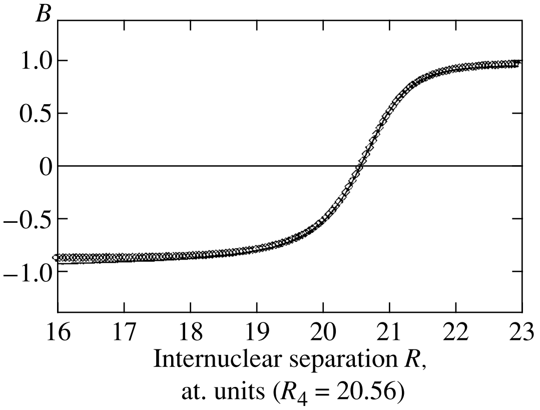

**Fig. 5.** Function $B$ versus internuclear separation $R$ for the $H^- + He^{++}$ system near the crossing of the ionic term with the covalent $n = 4$ term: the solid lines and diamonds represent, respectively, the exact numerical calculation and the approximation of a distant crossing, formulas (119) and (122).

Thus, for the $A^- + B^{Z+}$ system under study, we can write an expression for the nondiagonal matrix element,

$$V_{0n} = \langle \Phi_0 | V | \Psi_n \rangle = \Delta E_n(R), \tag{124}$$

and an expression for the energies of two quasi-crossing terms,

$$\begin{aligned}&(E_\pm(R) - E_n) \\ &= \frac{1}{2}\left(E_0(R) - E_n \pm \sqrt{(E_0(R) - E_n)^2 + 4\Delta E_n^2(R)}\right),\end{aligned} \tag{125}$$

because the diagonal matrix element for the ionic state is

$$\hat{H}_{00} = E_0(R) = \varepsilon_0 - \frac{Z}{R}, \tag{126}$$

and because the shifts of covalent terms may be disregarded by assuming that $\hat{H}_{nn} \approx E_n$. Substituting the expressions for energy (125) in (115) and (116), we derive relations (123).

Equation (115) for energy levels contains $\Delta E_n(R)$, which, according to (113) and (114), are expressed in terms of the sums of products of the Coulomb wave functions $Q_n$ and $Q_n^{x,y,z}$ calculated in the preceding section. These sums are shown in Figs. 3 and 4. As follows from these figures, $Q_n$, $Q_n^{x,y,z}$, and, hence, the nondiagonal matrix elements $V_{0n}$ are nonzero at all finite internuclear separations $R$, both for negative ions with a zero orbital angular momentum, $L = 0$, and for positive ions with $L = 1$.

Thus, the procedure for normalizing Coulomb Green's function as the wave function and analyzing the results allow us to calculate the nondiagonal matrix elements and to determine the behavior of terms for each crossing; i.e., this procedure allows the behavior of the $A^- + B^+$ system to be completely described.

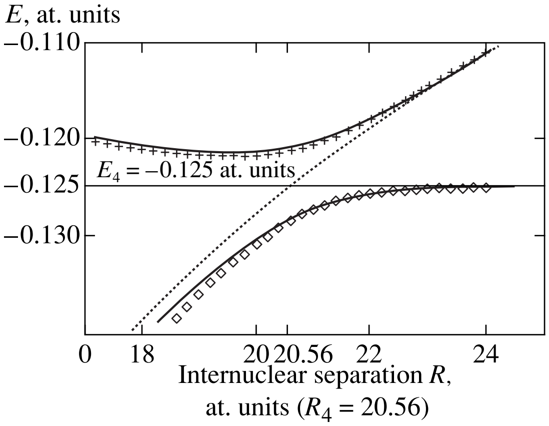

**Fig. 6.** Energy terms of the H$^-$ + He$^{++}$ system near the quasi-crossing of the ionic term with the covalent $n = 4$ term. A comparison of the results of an exact calculation obtained when solving a transcendental equation by using the logarithmic derivative of Green's function [33, 35] (solid lines) with the approximation of a distant term crossing (crosses and diamonds). The dotted line represents the ionic term in the zero-order approximation, $E_0(R) = \varepsilon_0 - Z/R$.

**Table 1.** Orbital sizes $r_n$, quasi-crossing positions $R_n$, and term splittings $\delta E_n = 2\Delta E_n(R_n)$ at quasi-crossings for the H$^-$ + H$^+$ system

| $n$ | $r_n$, at. units | $R_n$, at. units | $\delta E_n(R_n)$, at. units |
|---|---|---|---|
| 1 | 2.0 | 2.117 | $1.652^{-1}$ |
| 2 | 8.0 | 10.279 | $1.876^{-2}$ |
| 3 | 18.0 | 35.921 | $2.318^{-4}$ |
| 4 | 32.0 | 283.005 | $7.123^{-27}$ |

**Table 2.** Orbital sizes $r_n$, quasi-crossing positions $R_n$, and term splittings $\delta E_n = 2\Delta E_n(R_n)$ at quasi-crossings for the H$^-$ + He$^{++}$ system

| $n$ | $r_n$, at. units | $R_n$, at. units | $\delta E_n(R_n)$, at. units |
|---|---|---|---|
| 1 | 1.0 | 1.01 | $5.108^{-1}$ |
| 2 | 4.0 | 4.23 | $1.059^{-1}$ |
| 3 | 9.0 | 10.28 | $3.126^{-2}$ |
| 4 | 16.0 | 20.56 | $7.429^{-3}$ |
| 5 | 25.0 | 38.25 | $7.179^{-4}$ |
| 6 | 36.0 | 71.84 | $5.089^{-6}$ |
| 7 | 49.0 | 152.67 | $3.556^{-12}$ |
| 8 | 64.0 | 566.01 | $5.158^{-50}$ |

The system of adiabatic states breaks up into two groups. The first group includes the states that quasi-cross the ionic term. The second group includes the covalent states whose energies are above the energy $\varepsilon_0$ of the unperturbed negative ion and which are not involved in quasi-crossings. In the two-level approximation used for each crossing, the adiabatic energies of the first group are

$$E_n^{n+1}(R) = \frac{1}{2}\begin{cases} E_0(R) + E_{n+1} - D_{n+1}(R), & R \ge R_n^{n+1} \\ E_0(R) + E_n + D_n(R), & R \le R_n^{n+1}, \end{cases} \tag{127}$$

$$D_n(R) \approx \sqrt{x_n^2(R) + \Delta_n^2(R)},$$

$$x_n(R) \equiv E_0(R) - E_n, \quad x_n(R_n) = 0,$$

$$\Delta_n(R) \equiv 2\Delta E_n(R), \quad R_n^{n+1} \equiv \frac{R_n + R_{n+1}}{2}.$$

The two energy branches were joined at point $R_n^{n+1}$ between the two closest crossings $R_n$ and $R_{n+1}$.

The energy $E_n^{n+1}(R)$ is close to the Coulomb energy $E_n$ for $R < R_{n+1}$ and to $E_{n+1}$ for $R > R_{n+1}$. Between the crossings for $R_n < R < R_{n+1}$, this energy is close to the ionic energy of the zero-order approximation, $E_0(R)$ [see (26)]. Figure 6 shows the energies calculated from formula (127) in the approximation of a distant crossing for $n = 4$ ($R_4 = 20.56$) in comparison with the results of an exact calculation obtained when solving a transcendental equation using the logarithmic derivative of Green's function [33, 35]. The approximation of a distant term crossing is seen to be very close to the exact result. Tables 1 and 2 give the internuclear separations $R_n$ at which the ionic term crosses the covalent term and the term separations at quasi-crossings for H$^-$ + H$^+$ and H$^-$ + He$^{++}$ collisions.

The wave functions of the first group of states are

$$\Phi_n^{n+1}(\mathbf{R}, \mathbf{r}) = \begin{cases} C_{n+1}^-(R)\Phi_0(|\mathbf{R} - \mathbf{r}|) + C_{n+1}^+(R)\Psi_{n+1}(\mathbf{R}, \mathbf{r}), \\ R \ge R_n^{n+1} \\ C_{n+1}^+(R)\Phi_0(|\mathbf{R} - \mathbf{r}|) + C_n^-(R)\Psi_n(\mathbf{R}, \mathbf{r}), \\ R \le R_n^{n+1}, \end{cases} \tag{128}$$

where

$$C_n^{\pm}(R) \equiv \sqrt{\frac{D_n(R) \pm x_n(R)}{2D_n(R)}},$$

$$B_n^{n+1}(R) = \begin{cases} C_{n+1}^-(R), & R \ge R_n^{n+1} \\ C_n^+(R), & R \le R_n^{n+1}. \end{cases} \tag{129}$$

The coefficients $C_{n,n'}(R)$ describing the contributions of the states with principal quantum numbers $n' \neq n$ to the function $\Phi_n^{n+1}$ are

$$C^{\pm}_{n,n'}(R) = \mp\frac{\Delta E_{n'}(R_n)}{E_n - E_{n'}}C^{\pm}_n(R), \quad (n' \neq n). \tag{130}$$

The function $\Phi_n^{n+1}$ is close to $\Psi_n$ at $R < R_n$, because $|C_n^-| \approx 1$ and $C_n^+ \approx 0$ in this range. It is given by a linear combination of $\Psi_n$ and $\Phi_0$ near the crossing $R \approx R_n$ and by a linear combination of $\Phi_0$ and $\Psi_{n+1}$ near the crossing $R \approx R_{n+1}$. Between the crossings, $R_n < R < R_{n+1}$, $\Phi_n^{n+1}$ is close to the unperturbed wave function of the negative ion $\Phi_0$, when the amplitude $B_n^{n+1}(R)$ is close to unity and all $C_n^{\pm}$ are small. At $R > R_{n+1}$, $\Phi_n^{n+1}$ is close to $\Psi_{n+1}$, because $|C^+_{n+1}| \approx 1$ and $C^-_{n+1} \approx 0$ in this range.

The matrix elements of the derivative with respect to $R$ for two quasi-crossing active states are

$$\begin{aligned} V^n_{n-1}(R(t)) &\equiv \left(\frac{\partial}{\partial R}\right)^{n,n+1}_{n-1,n} \\ \equiv \int \Phi^n_{n-1}(\mathbf{R},\mathbf{r})\frac{\partial\Phi_n^{n+1}(\mathbf{R},\mathbf{r})}{\partial t}d\mathbf{r} &= C_n^-\frac{dC_n^+}{dR} - C_n^+\frac{dC_n^-}{dR} \\ &= \frac{Z}{R^2}\frac{2\pi N_0}{\sqrt{Q_n(R)}}\frac{Q_n(R) + x_n(R)\phi^2_{n0}}{x_n^2 + 4(2\pi N_0)^2 Q_n(R)}. \end{aligned} \tag{131}$$

These matrix elements are nonzero in narrow ranges $\Delta R_n$ near the points of quasi-crossing $R_n$, and their absolute values at $R = R_n$ are large:

$$\begin{aligned} (V^n_{n-1}(R(t)))_{\max} &= \left(\frac{\partial}{\partial R}\right)^{n,n+1}_{n-1,n}\bigg|_{R=R_n} \\ &= \frac{Z}{4R_n^2\Delta E_n(R_n)}. \end{aligned} \tag{132}$$

Near the $n$th crossing, the matrix elements between the active crossing states and the other active states $\Psi_{n'}$ that are not involved in the $n$th crossing, i.e., for $n' \neq n$, are also at a maximum. Using the coefficients $C_{n'}$, formula (121), we obtain

$$\begin{aligned} U^{n'}_{n-1}(R(t)) &= \left(\frac{\partial}{\partial R}\right)^{n'}_{n-1,n} = \frac{dC^-_{n,n'}}{dR} = -\frac{\Delta E_{n'}(R_n)}{E_n - E_{n'}}\frac{dC_n^-}{dR}, \\ U^{n'}_n(R(t)) &= \left(\frac{\partial}{\partial R}\right)^{n'}_{n,n+1} \\ &= \frac{dC^+_{n,n'}}{dR} = \frac{\Delta E_{n'}(R_n)}{E_n - E_{n'}}\frac{dC_n^+}{dR}, \end{aligned} \tag{133}$$

$$\begin{aligned} \frac{dC_n^{\pm}(R)}{dR} &= \pm 4\sqrt{2}\pi^2 N_0^2\frac{Z}{R^2} \\ &\times\frac{Q_n(R) + x_n(R)\phi^2_{n0}(R)}{D_n^{5/2}(R)\sqrt{D_n(R) \pm x_n(R)}}. \end{aligned} \tag{134}$$

The maximum absolute values of these matrix elements are reached at $R = R_n$,

$$\begin{aligned} &\frac{dC^{\pm}_{n,n'}(R)}{dR}\bigg|_{R=R_n} \\ &= \pm\frac{\Delta E_{n'}(R_n)}{\sqrt{2}(E_n - E_{n'})}\frac{Z}{4R_n^2\Delta E_n(R_n)}. \end{aligned} \tag{135}$$

These values are a factor of $\Delta E_{n'}(R_n)/(\sqrt{2}(E_n - E_{n'}))$ smaller than the maximum values (132) of matrix elements (131) between the wave functions of active states. The matrix elements (131) and (133) are shown in Figs. 7 and 8.

The wave functions of the extreme adiabatic states $\Phi_0^1$ and $\Phi_{n_m}^{n_m+1}$ ($n_m$ is the highest Coulomb level crossed by the ionic term) are given by different relations than (128) and (129). These formulas contain only the upper row with $n = 0$ for the extreme lower state $\Phi_0^1$ and only the lower row with $n = n_m$ for the extreme upper state $\Phi_{n_m}^{n_m+1}$.

The adiabatic energies of the states with principal quantum numbers $n > n_m$ are equal to the unperturbed Coulomb energies $E_n$. The wave functions of both passive and active states for these $n$ at all separations $\mathbf{R}$ are, respectively, $\Psi_{nlm}(\mathbf{R},\mathbf{r})$ and $\hat{\Psi}_n(\mathbf{R},\mathbf{r})$, formulas (70), (71) and (75), (76).

To conclude this section, we give asymptotic limits for the term splitting (125), which are valid for distant quasi-crossings when $R_n$ is more than twice the size $r_n$ of the corresponding Coulomb orbit:

$$L = 0, \quad \delta E_n^{as}(R_n) = 2\Delta E_n^{as}(R_n) \approx \sqrt{2\gamma n}A_s(R_n), \tag{136}$$

$$\begin{aligned} L = 1, \quad \delta E^{as}_{n,z}(R_n) &= 2\Delta E^{as}_{n,z}(R_n) \\ &\approx \sqrt{3Z}\frac{N_0^{(1)}}{\gamma}\sqrt{\frac{Z}{n}}A_s(R_n), \end{aligned} \tag{137}$$

$$\begin{aligned} \delta E^{as}_{n,xy}(R_n) &= 2\Delta E^{as}_{n,xy}(R_n) \\ &\approx \frac{N_0^{(1)}}{\gamma}\sqrt{\frac{3n(n-1)}{2}}\frac{A_s(R_n)}{R_n}, \end{aligned} \tag{138}$$

$$A_s(R_n) \equiv \left(\frac{Z}{n}\right)^{3/2}\left(\frac{2^n}{n!}\right)\left(\frac{ZR_n}{n}\right)^{n-1}\exp\left(-\frac{ZR_n}{n}\right).$$

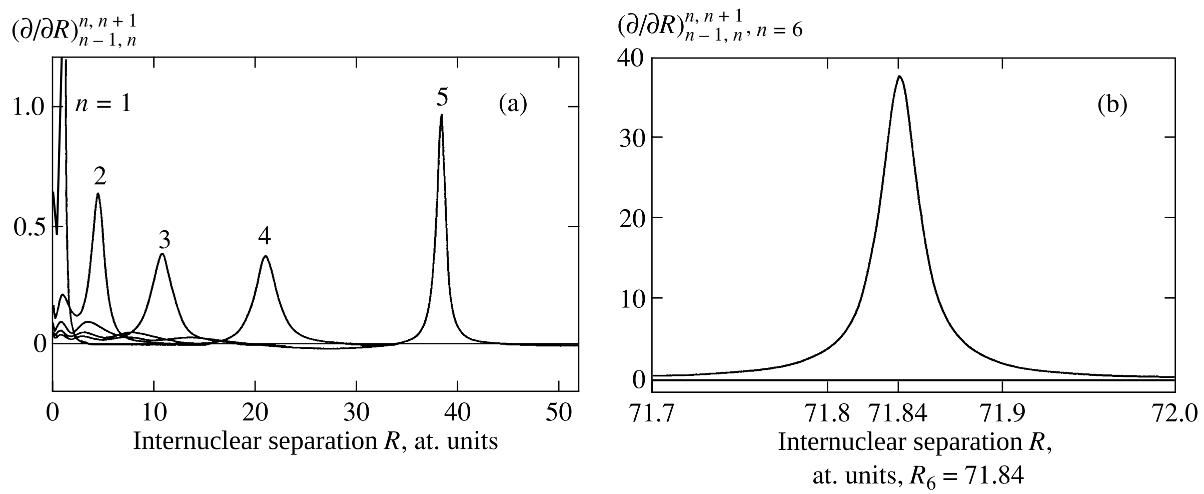

**Fig. 7.** Adiabatic matrix elements, formula (131), versus internuclear separation $R$ between the ionic state and the covalent states with (a) $n = 1, 2, 3, 4, 5$ and (b) $n = 6$ crossed by the ionic term.

We see from (136)–(138) that the asymptotics $\delta E^{as}_{n,\,xy}$ contains an additional negative power $R^{-1}$ compared to $\delta E^{as}_{n}(R_n)$ and $\delta E^{as}_{n,\,z}(R_n)$, because the asymptotics of the exchange matrix element is proportional to $R^{-|m|}$ [49].

## 6. THE SYSTEM OF CLOSE-COUPLING EQUATIONS AND THE ELECTRON CAPTURE CROSS SECTION

The evolution of a collisional system with time is described by the temporal Schrödinger equation for the wave function of an electron with coordinate $r$:

$$i\frac{\partial\Psi_{\text{tot}}(\mathbf{R},\mathbf{r},t)}{\partial t} = H_{\text{tot}}(t)\Psi_{\text{tot}}(\mathbf{R},\mathbf{r},t). \qquad (139)$$

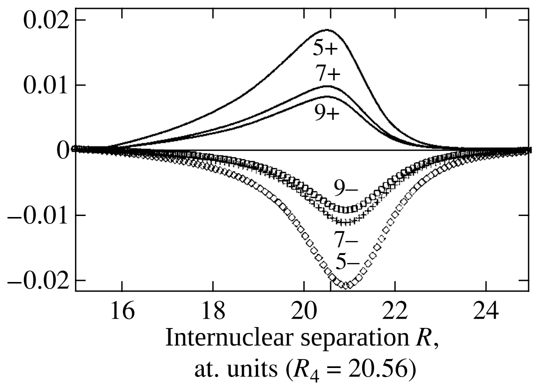

**Fig. 8.** Adiabatic matrix elements versus internuclear separation $R$ between the ionic state near its crossing with the covalent $n = 4$ term and the other states with $n = 5$, 7, and 9, formulas (133) and (134). The plus and minus signs refer, respectively, to one of the quasi-crossing states at $R \approx R_4$ and to the other state.

In this equation, the total two-center Hamiltonian $H_{\text{tot}}$ of the system depends on time $t$, because the atomic nuclei are assumed to be moving along a classical trajectory which is determined by the time dependence of internuclear separation $\mathbf{R}(t)$.

We expand the full wave function of the $A^- + B^+$ system in terms of the wave functions of the adiabatic states introduced above:

$$\Psi_{\text{tot}}(\mathbf{R},\mathbf{r},t) = \sum_{nlm}^{N_{\max}} b_{nlm}(t)\Psi_{nlm}(\mathbf{R},\mathbf{r}) \times \exp\left(-i\int^{t} E_n(t')dt'\right), \qquad (140)$$

with $N_{\max}$ exceeding $n_{\max}$; i.e., (140) includes the covalent states whose binding energies are lower than those of the unperturbed negative ion. These states are populated through the coefficients $C_{n,\,n'}$ in the expansion of the ionic wave function. The sum in (140) contains active and passive states.

The system of adiabatic equations for the coefficients $b_{nlm}(t)$ that results after substituting expansion (140) in Eq. (139) [50] is

$$\frac{db_{nlm}}{dt} = \sum_{n'l'm'}^{N_{\max}} b_{n'l'm'}(t)\left(\frac{\partial}{\partial t}\right)^{n'l'm'}_{nlm} \exp(i\phi_{n,\,n'}(t)), \qquad (141)$$

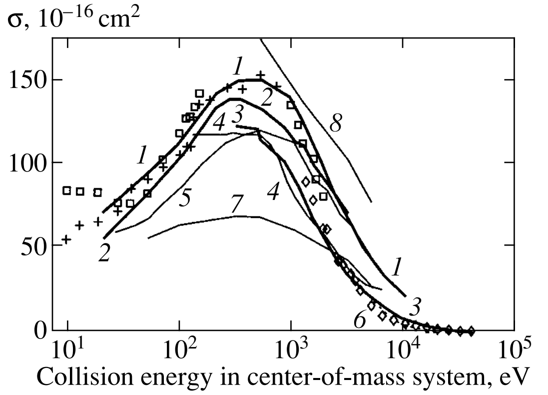

**Fig. 9.** The total cross section for electron capture in the $H^- + H^+ \longrightarrow H + H^*(n)$ collision. Theory: curve *1*—our calculation, curve *2*—from [1], curve *3*—Ermolaev's calculation [2], curve *4*—from [5], curve *5*—from [6], dotted curve *6*—the calculation of Shingal *et al.* [4], curve *7*—the calculation of Bates and Lewis [11], and curve *8*—the calculation of Dalgarno *et al.* [12]. Experiment: squares and crosses—merged-beam measurements of Brouillard's group [13], respectively, for the $H^- + H^+$ and $H^- + D^+$ collisions (the results for the $H^- + D^+$ collision are shown for the same collision velocity as that for $H^- + H^+$); and diamonds—crossed-beam measurements of Salzborn's group [16].

where the phases $\phi_{n',n}(t)$ are

$$\phi_{n',n}(t) \equiv \int_{t_0}^{t} (E_n^{n+1}(R(t')) - E_{n'}^{n'+1}(R(t')))dt'. \tag{142}$$

The system of equations (141) includes all the active and passive states with principal quantum numbers $n \le N_{\max}$, with the maximum principal quantum number varying in the range $N_{\max} = 9$–14. System (141) was solved numerically by the methods described in Hemming's book [51]. We used the matrix elements calculated above. The trajectories $\mathbf{R}(t)$ corresponded to the Coulomb field of attraction between $A^-$ and $B^{Z+}$. The partial ($\sigma_i$) and total ($\sigma_{\text{tot}}$) electron capture cross sections were calculated using the formula

$$\sigma_i = 2\pi\int_0^\infty P_i(\rho)\rho d\rho, \tag{143}$$

where $P_i$ is the probability of populating state $i$, $\sigma_{\text{tot}} = \sum_i \sigma_i$.

Figure 9 shows the total cross section for electron capture in the $H^- + H^+ \longrightarrow H + H^*(n)$ collision. The discrepancy of theoretical data is large. If we exclude the early results (curves *7* and *8*), then our calculated cross section will be the largest and the closest to the experimental results. This is because of the higher population of the $n = 3$ state in our calculation attributable to the coupling of the two crossing $n = 2$ states with the $n = 3$ states. This coupling is given by the matrix elements (133).

Figure 10 shows the total cross section for the $H^- + He^{++} \longrightarrow H + He^+(n)$ collision. As in the previous case, our calculated cross section exceeds other theoretical cross sections and is closest to the experimental cross section. The cause of the disagreement with other theoretical calculations is the same. Because of the nonadiabatic coupling between the crossing states and the states that are not involved in a given crossing, the population of more highly excited covalent levels significantly increases. The $n = 4$, 5, 6, and 7 states are most populated.

Figure 11 shows the total cross section for electron capture in the $Ca^- + He^{++} \longrightarrow Ca + He^+(n)$ collision that we calculated by disregarding spin-orbit splitting

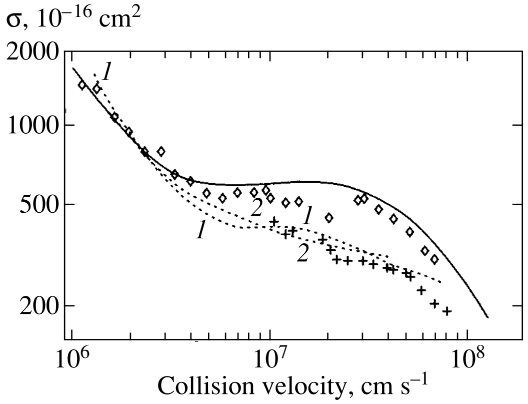

**Fig. 10.** The total cross section for electron capture in the $H^- + He^{++} \longrightarrow H + He^+(n)$ collision. Theory: the solid line represents our calculation; dotted curves *1* and *2* represent variational calculations [17, 19]. Experiment: crosses—from [18] and diamonds—from [17].

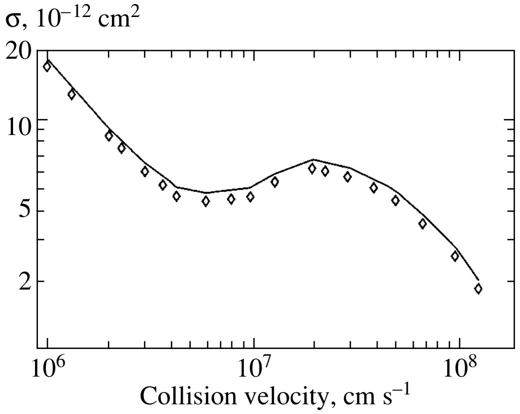

**Fig. 11.** Our calculated total cross section for electron capture in the $Ca^- + He^{++} \longrightarrow Ca + He^+(n)$ collision. The solid curve represents the cross section for the state with a zero component of the orbital angular momentum along the internuclear axis, $m = 0$; the diamonds are for $|m| = 1$.

of the binding energies for $\mathrm{Ca}^-$. We took the binding energy of this negative ion to be 0.022 eV, which is the mean of two actual binding energies: $E_{3/2} = 0.01973$ eV and $E_{1/2} = 0.02455$ eV (see, e.g., [27]). Figure 11 shows the two cross sections corresponding to two components of the orbital angular momentum $L = 1$ for $\mathrm{Ca}^-$ along the internuclear axis **R**. The states with the angular momentum components $|m| = 1$ and 0 have different term splittings at quasi-crossings, so the electron capture cross sections for them are different. The capture cross section for this ion is very large, because its binding energy is very low. Distances of $200–500a_0$ mainly contribute to the cross section. The covalent states $n = 14–19$ are most populated.

## ACKNOWLEDGMENTS

This study was supported by the Council for Grants of the President of Russia and State Support for Leading Scientific Schools (grant no. 00-15-96526).

## REFERENCES

1. V. Sidis, C. Kubach, and D. Fussen, Phys. Rev. A **27**, 2431 (1983).
2. A. M. Ermolaev, J. Phys. B **21**, 81 (1988); **25**, 3133 (1992).
3. D. Fussen and C. Kubach, J. Phys. B **19**, L31 (1986).
4. R. Shingal, B. H. Bransden, and D. R. Flower, J. Phys. B **18**, 2485 (1985).
5. L. F. Errea, C. Harel, P. Jimeno, *et al.*, Phys. Rev. A **54**, 967 (1996); Nucl. Instrum. Methods Phys. Res. B **98**, 335 (1995).
6. F. Borondo, A. Macias, and A. Riera, Phys. Rev. Lett. **46**, 420 (1981); Chem. Phys. **81**, 303 (1983); Chem. Phys. Lett. **100**, 63 (1983).
7. M. Chibisov, F. Brouillard, J. P. Chenu, *et al.*, J. Phys. B **30**, 991 (1997).
8. M. I. Chibisov, R. K. Janev, F. Brouillard, *et al.*, J. Phys. B **31**, 2795 (1998).
9. A. A. Radtsig and B. M. Smirnov, Zh. Éksp. Teor. Fiz. **60**, 521 (1971) [Sov. Phys. JETP **33**, 282 (1971)].
10. B. M. Smirnov, *Negative Ions* (Atomizdat, Moscow, 1978; McGraw-Hill, New York, 1982).
11. D. R. Bates and J. T. Lewis, Proc. Phys. Soc. London, Sect. A **68**, 173 (1955).
12. A. Dalgarno, G. A. Victor, and P. Blanchard, Air Force Cambridge Research Laboratory Report, No. 71-0342 (1971).
13. S. Szücs, M. Karemera, M. Terao, and F. Brouillard, J. Phys. B **17**, 1613 (1984).
14. B. Peart, M. A. Bennet, and K. Dolder, J. Phys. B **18**, L439 (1985).
15. B. Peart and D. A. Hayton, J. Phys. B **25**, 5109 (1992).
16. W. Schön, S. Krüdener, F. Melchert, *et al.*, J. Phys. B **20**, L759 (1987).
17. M. Terao, S. Szücs, M. Cherkani, *et al.*, Europhys. Lett. **1** (3), 123 (1986).
18. B. Peart and M. A. Bennet, J. Phys. B **19**, L321 (1986).
19. M. Terao, C. Harel, A. Salin, and R. J. Allan, Z. Phys. D **7**, 319 (1988).
20. M. H. Cherkani, S. Szücs, M. Terao, *et al.*, J. Phys. B **24**, 209 (1991).
21. M. H. Cherkani, S. Szücs, H. Hus, and F. Brouillard, J. Phys. B **24**, 2367 (1991).
22. B. Peart and P. M. Wilkins, J. Phys. B **19**, L515 (1986).
23. J. P. Chenu, A. Naji, K. Olamba, *et al.*, in *Abstracts of Contributed Papers of the XX International Conference on the Physics of Electronic and Atomic Collisions, ICPEAC, Vienna, Austria, 1997,* TU p. 187.
24. R. K. Janev, M. I. Chibisov, and F. Brouillard, Phys. Scr. **60**, 1573 (1999).
25. C. Desfrançois, Phys. Rev. A **51**, 3667 (1995).
26. I. I. Fabrikant, J. Phys. B **31**, 2921 (1998).
27. I. I. Fabrikant and M. I. Chibisov, Phys. Rev. A **61**, 022718 (2000).
28. C. Desfrançois, H. Abdoul-Carime, N. Khelifa, and J. P. Schermann, Phys. Rev. Lett. **73**, 2436 (1994); C. Desfrançois, B. Baillon, J. P. Schermann, *et al.*, Phys. Rev. Lett. **72**, 48 (1994); C. Desfrançois, H. Abdoul-Carime, and J.-P. Schermann, Int. J. Mod. Phys. **10**, 1339 (1996).
29. R. N. Compton, H. S. Carman, Jr., C. Desfrançois, *et al.*, J. Chem. Phys. **105**, 3472 (1996).
30. M. Reicherts, T. Roth, A. Gopalan, *et al.*, Europhys. Lett. **40**, 129 (1997).
31. B. M. Smirnov, Dokl. Akad. Nauk SSSR **161** (1), 92 (1965) [Sov. Phys. Dokl. **10**, 218 (1965)].
32. B. M. Smirnov, Dokl. Akad. Nauk SSSR **173**, 316 (1967) [Sov. Phys. Dokl. **12**, 242 (1967)].
33. I. V. Komarov, P. A. Pogorelyĭ, and A. S. Tibilov, Opt. Spektrosk. **27**, 198 (1969).
34. L. P. Presnyakov, Phys. Rev. A **2**, 1720 (1970).
35. T. M. Kereselidze and M. I. Chibisov, Zh. Éksp. Teor. Fiz. **68**, 12 (1975) [Sov. Phys. JETP **41**, 6 (1975)].
36. Yu. N. Demkov and V. N. Ostrovskii, *Zero-Range Potentials and Their Applications in Atomic Physics* (Leningr. Gos. Univ., Leningrad, 1975; Plenum, New York, 1988).
37. L. Hostler and R. H. Pratt, Phys. Rev. Lett. **10**, 469 (1963).
38. L. Hostler, J. Math. Phys. **5**, 591 (1964).
39. M. I. Chibisov, A. M. Ermolaev, F. Brouillard, and M. Cherkani, Zh. Éksp. Teor. Fiz. **117**, 313 (2000) [JETP **90**, 276 (2000)].
40. M. I. Chibisov, A. M. Ermolaev, F. Brouillard, and M. Cherkani, Phys. Rev. Lett. **84**, 451 (2000).
41. M. H. Cherkani, F. Brouillard, and M. I. Chibisov, J. Phys. B **34**, 39 (2001).
42. A. I. Baz', Ya. B. Zel'dovich, and A. M. Perelomov, *Scattering, Reactions and Decays in Nonrelativistic Quan-*

*tum Mechanics* (Nauka, Moscow, 1971, 2nd ed.; Israel Program for Scientific Translations, Jerusalem, 1966).

43. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1975; Pergamon, New York, 1977).
44. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Fizmatgiz, Moscow, 1962; Academic, New York, 1980).
45. Yu. N. Demkov and I. V. Komarov, Zh. Éksp. Teor. Fiz. **50**, 286 (1966) [Sov. Phys. JETP **23**, 189 (1966)].
46. I. A. Eganova and M. I. Shirokov, Preprint No. R4-5438, OIYaI (Joint Inst. for Nuclear Research, Dubna, 1970); Ann. Phys. (Leipzig), Folge 7, **21** (5/6) (1968).
47. V. A. Fok, Izv. Akad. Nauk SSSR, Otd. Mat. Estestv. Nauk, No. 2, 169 (1935); V. A. Fock, Z. Phys. **98**, 145 (1935).
48. V. S. Lebedev and I. L. Beigman, *Physics of Highly Excited Atoms and Ions* (Springer-Verlag, Berlin, 1998).
49. B. M. Smirnov, *Atomic Collisions and Elementary Processes in Plasma* (Atomizdat, Moscow, 1968).
50. E. E. Nikitin and S. Ya. Umanskii, *Nonadiabatic Transitions in Slow Atomic Collisions* (Atomizdat, Moscow, 1979); *Theory of Slow Atomic Collisions* (Springer-Verlag, Berlin, 1984).
51. R. W. Hamming, *Numerical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1962; Nauka, Moscow, 1972).

*Translated by V. Astakhov*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 278–294.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 315–332.*
*Original Russian Text Copyright © 2001 by Popov.*

## ATOMS, SPECTRA, RADIATION

# Multiphoton Atomic Ionization in the Field of a Very Short Laser Pulse

**V. S. Popov**

*Institute of Theoretical and Experimental Physics, State Scientific Center of the Russian Federation,*
*ul. Bol'shaya Cheremushkinskaya 25, Moscow, 117259 Russia*

Received February 23, 2001

**Abstract**—Closed analytic expressions are derived for the probability of multiphoton atomic and ionic ionization in a variable electric field $\mathcal{E}(t)$, which are applicable for arbitrary Keldysh parameters γ. Dependencies of the ionization probability and photoelectron pulse spectrum on the shape of a very short laser pulse are analyzed. Examples of pulse fields of various forms, including a modulated light pulse with a Gaussian or Lorentz envelope, are considered in detail. The interference effect in the photoelectron energy spectrum during atomic ionization by a periodic field of a general form is examined. The range of applicability of the adiabatic approximation in the multiphoton ionization theory is discussed. The imaginary time method is used in the calculations, which allows the probability of particle tunneling through oscillating barriers to be effectively calculated. *© 2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION AND STATEMENT OF THE PROBLEM

Studies of physical phenomena under extreme conditions, including those in strong external fields, have always been of considerable interest. As is well known, the Einstein law for the photoelectric threshold breaks down in the field of a strong light wave and multiphoton atomic and molecular ionization becomes possible.

The production of increasingly intense electromagnetic fields is associated with the shortening of a laser pulse, whose duration becomes comparable to an optical period (see, e.g., [1–3] and references therein) and whose spectrum contains many higher harmonics. Since multiphoton ionization is a highly nonlinear process [4-8], it cannot be reduced to the sum of contributions from individual harmonics. The ionization probability and the pulse spectrum of the emerging photoelectrons significantly depend on the shape of the pulse field, particularly when the external-field frequency ω exceeds the tunneling frequency $\omega_t$. Elucidating this dependence becomes necessary for analyzing experimental data and is the subject of our study (see also [9, 10]).

To this end, consider the problem of atomic-level ionization in a variable electric field (spatially uniform and linearly polarized),

$$\mathcal{E}(t') = F\varphi(t), \quad t = \omega t', \tag{1}$$

which arbitrarily depends on time $t'$. Here, $F$ is the field amplitude, ω is the characteristic field frequency, and $t$ is the dimensionless time. Regarding the function φ specifying the pulse shape, we assume that $|\varphi(t)| \le \varphi(0) = 1$ for $-\infty < t < \infty$ (i.e., $t = 0$ is the time of field maximum, when the electron emerges from under the barrier [6]) and $\varphi(-t) = \varphi(t)$; note that the latter condition is imposed only for the convenience of calculations.

For our calculations, we use the semiclassical imaginary time method (ITM) [6, 7, 11], which gives a clear description of the particle tunneling through any smooth and rapidly oscillating barriers. In this case, the subbarrier trajectories that formally satisfy the classical equations of motion (but with imaginary time!) are introduced. The principal (exponential) factor in the ionization probability is determined by the so-called extreme subbarrier trajectory, on which the imaginary part of the action function, Im$S$, reaches a minimum (and which, according to Feynmann [12, 13], determines the most probable particle tunneling path). Next, for the pulse spectrum of the emerging electrons to be derived, we must consider a bundle of classical, nearly extreme trajectories and calculate the quadratic correction to Im$S$ proportional to $(\mathbf{p} - \mathbf{p}_{\max})^2$ on them (see [7, 11] for more details; the ITM is also presented in the monograph [14]).

Below, we assume the following conditions to be satisfied:

$$K_0 = \frac{I}{\hbar\omega} \gg 1, \quad \epsilon = \frac{e\hbar F}{(2I)^{3/2}m^{1/2}} = \frac{F}{\kappa^3 F_a} \ll 1, \quad \mathrm{Im}S \gg \hbar, \tag{2}$$

they ensure that the semiclassical approximation is applicable to multiphoton processes. Here, $K_0$ is the multiquantum parameter, ϵ is the reduced electric field, $I = \kappa^2 me^4/2\hbar^2$ is the atomic-level ionization potential, and κ is the (dimensionless) momentum characteristic of a bound state.[1] The tunneling process significantly depends on Keldysh parameter $\gamma$ [4],

$$\gamma = \frac{\omega}{\omega_t} = \frac{\omega\sqrt{2mI}}{eF}, \qquad (2')$$

where $\omega_t$ is the tunneling frequency in a static field $F$ (below, we use atomic units, $\hbar = m = e = 1$, where $m$ is the electron mass, $F_a = m^2e^5/\hbar^4 = 5.14 \times 10^9$ V cm$^{-1}$).

Our goal is to investigate the atomic-level ionization probability and the photoelectron pulse spectrum as functions of pulse-field shape (1) and Keldysh parameter.[2] In the next section, we give basic equations that describe this dependence (in the semiclassical approximation). We formulate an algorithm that allows the auxiliary function $\chi(z)$ to be calculated from the specified shape of the external field $\varphi(t)$ whereupon the problem reduces to quadratures. In Section 3, we consider model examples of fields $\varphi(t)$, for which all calculations are performed analytically. The results of our numerical calculations, including those for an amplitude-modulated electromagnetic wave with a Gaussian or Lorentz envelope, are presented in Section 4. The tunneling interference effect in the photoelectron energy spectrum is examined in Section 5, and the range of applicability of the adiabatic approximation in the multiphoton ionization theory is considered in Section 6. Our results are briefly discussed in the final Section 7. Details of our calculations, including the derivation of approximation (17) for large $\gamma$, auxiliary formulas, and asymptotics, are given in the Appendices.

The results presented below have been partly announced in [9, 10]. It should be noted that the theory of multiphoton ionization by very short laser pulses was also considered by Keldysh [3], who, in particular, analytically and numerically analyzed (by a different method) soliton-like, Gaussian, and Lorentz pulses. To compare our results with those from [3], it is useful to bear in mind that Keldysh [3] used the following notation: $\Omega = 1/K_0$, $\mathcal{E} = 2\epsilon$, $\lambda = 1/\gamma$, and $f'(x) = \varphi(x)$ with $x = \omega t' \equiv t$.

## 2. BASIC EQUATIONS

For field (1), it is clear from physical considerations that the extreme trajectory is one-dimensional and directed along the field. Solving the classical equations of motion and calculating Im$S$ yields the multiphoton ionization probability (to within a preexponential factor):

$$w_i(\mathbf{p}) \equiv \frac{dW_i}{d\mathbf{p}} \propto \exp\left\{-\frac{2}{3\epsilon}g(\gamma) - \frac{\kappa}{F}\left[b_1(\gamma)(p_\parallel - p_{\max})^2 + b_2(\gamma)p_\perp^2\right]\right\}, \qquad (3)$$

where $p_\parallel$ and $p_\perp$ are the longitudinal (along the field) and transverse momenta of the emerging electron, respectively; $p_{\max}$ is its most probable momentum:[3]

$$g(\gamma) = \frac{3}{2}\int_0^1 \chi(\gamma z)(1 - z^2)dz,$$

$$b_1(\gamma) = -\gamma\frac{db_2}{d\gamma},$$

$$b_2(\gamma) = \int_0^1 \chi(\gamma z)dz = \gamma^{-1}\tau(\gamma), \qquad (4)$$

$$p_{\max} = \frac{F}{\omega}\int_0^\infty \varphi(t)dt$$

($p_\perp$ is the integral of motion, and $F/\omega$ is the characteristic momentum of the oscillatory electron motion in the wave field). This includes the function $\chi(z)$, which depends on pulse shape (1); after its derivation, the problem reduces to quadratures. This function can be specified parametrically:

$$\chi = \frac{1}{\tilde{\varphi}(\tau)}, \quad z = h(\tau) = \int_0^\tau \tilde{\varphi}(t)dt, \qquad (5)$$

where $\tilde{\varphi}(\tau) \equiv \varphi(it)$. The latter equation also defines (implicitly) the inverse function $\tau = h^{-1}(z)$; in this case,[4]

$$\chi(z) = \tau'(z), \qquad (6)$$

and the initial time of the subbarrier electron motion is

$$t_0' = i\omega^{-1}\tau(\gamma), \quad \tau(\gamma) = h^{-1}(\gamma). \qquad (6')$$

The derivation of Eqs. (3)–(6) is omitted here. It is based on the ITM and is a generalization of the calculations in [7] for the special (but important for applications) case of a monochromatic laser field. In this paper,

---

[1] For the ground level of a hydrogen atom, $\kappa = 1$. For the ground states of neutral atoms, $\kappa$ ranges from 0.535 for Cs ($I = 3.89$ eV) to 1.344 for He ($I = 24.59$ eV); see [15] and Table 1 in [11]. For weakly bound states, this parameter can be appreciably smaller than unity. Thus, for example, $I = 0.754$ eV and $\kappa = 0.235$ for a negative hydrogen ion and $I = 0.077$ eV and $\kappa = 0.075$ for He$^-$. In these cases, the reduced fields $\epsilon$ and the ionization probabilities significantly increase (at fixed electric field $F$).

[2] For monochromatic laser emission, the momentum, energy, and angular distributions of photoelectrons during multiphoton ionization have already been considered previously [6–8]. Recently, they have been analyzed in detail [16–8] over the entire range of $\gamma$, including the general case of elliptic polarization.

[3] Assuming that $p_{\max} = 0$ in (3), we obtain the pulse spectrum of the electrons at the time of their emergence form under the barrier ($t = 0$). Its recalculation to the distribution of emerging photoelectrons ($t \longrightarrow +\infty$) in finite kinetic energies is a separate problem, which has been addressed by many authors (see, e.g., [19–23]). Here, we only note that expression (4) for $p_{\max}$ applies only to very short (with a duration $T \lesssim 1$ ps) laser pulses).

[4] In (5), $\tau$ is the parameter whose elimination yields an explicit dependence of $\chi$ on $z$. Note that $h(\tau)$ differ from the vector potential $A(t')$ continued to the complex plane only by the factor $i\omega/cF$.

we apply these formulas to pulse fields of various forms and discuss the effect of laser-pulse shortening on the total ionization probability and on the photoelectron spectrum. Note that examples of calculating (by the ITM) the preexponential factor, including those for time-constant electric and magnetic fields of arbitrary magnitudes and directions, can be found in [7, 11].

In the tunneling limit, $\gamma \ll 1$ (low-frequency laser field), ionization occurs at times close to $t = 0$. Assuming for $t \longrightarrow 0$ that

$$\varphi(t) = 1 - \frac{a_2}{2!}t^2 + \frac{a_4}{4!}t^4 - \frac{a_6}{6!}t^6 + \ldots, \quad a_2 > 0, \tag{7}$$

we find from (5) that [see also (A.3) and (A.4) in Appendix A]

$$\chi(z) = \sum_{n=0}^{\infty}(-1)^n\frac{\chi_n}{(2n)!}z^{2n}, \qquad \chi_0 = 1, \quad \chi_1 = a_2, \quad \chi_2 = 10a_2^2 - a_4, \tag{7'}$$

whereupon, using (4), we obtain expansions that allow the adiabatic corrections to be easily calculated:

$$g(\gamma) = 1 + \sum_{n=1}^{\infty}(-1)^n g_n\gamma^{2n}, \qquad g_n = \frac{3}{(2n+1)!(2n+3)}\chi_n, \tag{8}$$

$$b_1(\gamma) = \sum_{n=0}^{\infty}(-1)^n\frac{\chi_{n+1}}{(2n+1)!(2n+3)}\gamma^{2n+2}, \qquad b_2(\gamma) = \sum_{n=0}^{\infty}(-1)^n\frac{\chi_n}{(2n+1)!}\gamma^{2n}. \tag{8'}$$

The last two formulas can also be written as

$$b_1 = \frac{2}{3}\sum_{n=1}^{\infty}(-1)^{n-1}n(2n+3)g_n\gamma^{2n}, \qquad b_2 = \frac{1}{3}\sum_{n=0}^{\infty}(-1)^n(2n+3)g_n\gamma^{2n}, \tag{8''}$$

where $g_n$ are the coefficients of series (8) for the function $g(\gamma)$. To a first approximation in $\gamma^2$, the dependence of multiphoton ionization probability (3) on pulse shape is universal:

$$w_i(\mathbf{p}) \propto \exp\left\{-\left[\frac{2\kappa^3}{3F}\left(1 - \frac{1}{10}\tilde{\gamma}^2\right) + \frac{\kappa}{F}\left(\frac{1}{3}\tilde{\gamma}^2(p_\parallel - p_{\max})^2 + p_\perp^2\right)\right]\right\}, \qquad \gamma \ll 1, \tag{9}$$

where $\tilde{\gamma} = \sqrt{a_2}$ and $a_2 = -\varphi''(0)$ is the curvature of the pulse near its apex. In this case, $\Delta p_\perp \sim \sqrt{\epsilon}\,\kappa \ll \kappa$, $\Delta p_\parallel \sim \gamma^{-1}\Delta p_\perp \sim \sqrt{\epsilon}\,E/\omega \ll p_0$, and $p_0 \sim F/\omega$ is the characteristic momentum of the electron oscillatory motion in the wave field (for $t > 0$). In our case, the longitudinal electron momentum is much larger than the transverse one, which is explained by the possibility of electron acceleration along a slowly varying electric field $\mathcal{E}(t)$.

In the other limit (rapidly varying fields), it is convenient to rewrite (3) in a different form:

$$w_i(\mathbf{p}) \propto \exp\left\{-\left[\frac{\kappa^2}{\omega}f(\gamma) + \frac{1}{\omega}(c_1p_\parallel^2 + c_2p_\perp^2)\right]\right\}, \qquad \gamma \gg 1, \tag{10}$$

where $\kappa^2/\omega = 2K_0 \gg 1$,

$$f(\gamma) = \frac{2}{3}\gamma g(\gamma) = \int_0^{\gamma}\chi(u)\left(1 - \frac{u^2}{\gamma^2}\right)du, \tag{10'}$$

and $c_{1,2}(\gamma) = \gamma b_{1,2}(\gamma)$. In physical problems, the function $f(\gamma)$ for $\gamma \longrightarrow \infty$ either approaches a constant limit [if $\varphi(t)$ has the singularity $t = i$, $\tau_s$ at a finite distance from the real time axis in the complex plane] or increases logarithmically [see Table 1 in [9] and expansions (A.5)–(A.10)].

## 3. ANALYTIC MODELS

In the cases considered below, Eqs. (4)–(6) can be solved analytically. These model examples define the basic qualitative characteristics of the process under study and can also be of interest in their own right.

(1) $\varphi(t) = \cos t$ corresponds to monochromatic laser light. In this case, $\chi(z) = (1 + z^2)^{-1/2}$, the integrals in (4) are tabular and give

$$f(\gamma) = \left(1 + \frac{1}{2\gamma^2}\right)\operatorname{arcsinh}\gamma - \frac{\sqrt{1+\gamma^2}}{2\gamma}, \qquad c_1 = \operatorname{arcsinh}\gamma - \frac{\gamma}{\sqrt{1+\gamma^2}}, \qquad c_2 = \tau_0(\gamma) = \operatorname{arcsinh}\gamma, \quad p_{\max} = 0, \tag{11}$$

in close agreement with previous results [4-6].

(2) For $\varphi(t) = 1/\cosh^2 t$ (soliton-like pulse), we have $\chi(z) = (1 + z^2)^{-1}$, whence

$$f(\gamma) = (1 + \gamma^{-2})\arctan\gamma - \frac{1}{\gamma}, \qquad c_1 = \arctan\gamma - \frac{\gamma}{1+\gamma^2}, \quad c_2 = \arctan\gamma, \qquad p_{\max} = F/\omega. \tag{12}$$

(3) For a Gaussian pulse, $\varphi(t) = \exp(-t^2/2\sigma^2)$, $\chi(z)$ can be determined from the equation

$$\frac{1}{\chi}w(\sqrt{-\ln\chi}) = \frac{z}{\sqrt{2}\sigma}, \quad 0 < \chi \le 1, \qquad (13)$$

where $w(z)$ is the so-called Dawson integral [24] (see also Appendix B). At $\sigma = 1$, we have

$$\chi(z) = \begin{cases} 1 - \frac{1}{2}z^2 + \frac{7}{24}z^4 + \dots, & z \longrightarrow 0, \\ (z\sqrt{2\ln z})^{-1}[1 + O(1/\ln z)], & z \longrightarrow \infty, \end{cases} \qquad (13')$$

as a result, we obtain expansions (A.10).

(4) For a pulse field,

$$\varphi(t) = (1 + t^2)^{-\alpha}, \quad \alpha \ge 1, \qquad (14)$$

Eqs. (5) take the form

$$\chi = (1 - \tau^2)^{\alpha}, \quad z = \tau F(\alpha, 1/2; 3/2; \tau^2), \quad 0 \le \tau < 1, \qquad (5')$$

where $\tau$ is the parameter and $F(\dots) \equiv {}_2F_1(\dots)$ is the Gaussian hypergeometric function; hence

$$\chi(z) = 1 - \alpha z^2 + \frac{1}{6}(7\alpha^2 - 3\alpha)z^4 + \dots, \quad z \longrightarrow 0, \qquad (15)$$

$$\chi(z) \approx [2(\alpha - 1)z]^{-\alpha/(\alpha - 1)}, \quad z \longrightarrow \infty (\alpha > 1). \qquad (15')$$

As $\gamma \longrightarrow \infty$, the initial point of the subbarrier trajectory, $t_0 = i\tau_0(\gamma)$, approaches to the singularity $t_s = i$ of field (14):

$$\tau_0(\gamma) = \begin{cases} 1 - 2e^{-2\gamma} + \dots, & \alpha = 1, \\ 1 - k\gamma^{-1/(\alpha - 1)} + \dots, & \alpha > 1, \end{cases} \qquad (16)$$

where $k = [2^{\alpha}(\alpha - 1)]^{-1/(\alpha - 1)}$ (see Appendix C for more details). This example is typical of those cases where the nearest singularity $t_s = i\tau_s$ of field function $\varphi(t)$ lies at a finite distance from the real axis and is a pole ($\alpha = 1, 2, \dots$) or a power-type bifurcation point. In this case, (10) takes the asymptotic form

$$w_i(\mathbf{p}) \propto \exp\left\{-\frac{\tau_s}{\omega}[\kappa^2 + (p_{\parallel} - p_{\max})^2 + p_{\perp}^2]\right\}, \quad \gamma \gg 1, \qquad (17)$$

$$p_{\max} = \frac{\sqrt{\pi}\Gamma(\alpha - 1/2)}{2\Gamma(\alpha)}\frac{\kappa}{\gamma} \ll \kappa, \qquad (17')$$

which differs from the probability of ionization by monochromatic emission and significantly exceeds it.

Here,

$$W_i = \int w_i(\mathbf{p})d\mathbf{p} \propto \exp(-2K_0\tau_s),$$

while for a monochromatic field,

$$W_i \propto \exp(-2K_0\ln 2\gamma).$$

In this case, the photoelectron momentum distribution approaches an isotropic Gaussian distribution with the center at point $\mathbf{p}_{\max}$ in momentum space:

$$\Delta p_{\parallel} = \Delta p_{\perp} \sim \sqrt{\omega/\tau_s} \sim \kappa(K_0\tau_s)^{-1/2} \ll \kappa.$$

(5) At integer and half-integer $\alpha$, the hypergeometric function in (5') reduces to elementary functions; see formulas (C.1) and (C.2) in Appendix C. The case $\alpha = 3/2$ is particularly simple analytically: here, $z = \tau(1 - \tau^2)^{-1/2}$, $\chi(z) = (1 + z^2)^{-3/2}$, and we have[5]

$$f(\gamma) = \frac{1}{\gamma}\left(\sqrt{1 + \gamma^2} - \frac{\operatorname{arcsinh}\gamma}{\gamma}\right), \quad c_1 = c_2^3, \quad c_2 = \frac{\gamma}{\sqrt{1 + \gamma^2}}. \qquad (18)$$

This example has a kind of a duality: $\varphi(t) \equiv \chi(t)$ for it, which also holds for a pulse of the form $\varphi(t) = 1/\cosh t$.

The case $\alpha = 1$, i.e., $\varphi(t) = 1/(1 + t^2)$, corresponds to a Lorentz pulse shape, with

$$f(\gamma) = 1 - \frac{1}{\gamma^2}\int_0^{2\gamma}\frac{xdx}{e^x + 1} = \begin{cases} \frac{2}{3}\gamma - \frac{2}{15}\gamma^3 + \frac{4}{105}\gamma^5 - \dots, & \gamma \ll 1, \\ 1 - \frac{\pi^2}{12\gamma^2} + 2\gamma^{-1}\exp(-2\gamma), & \gamma \gg 1, \end{cases} \qquad (19)$$

$$c_1(\gamma) = \tanh\gamma - \frac{\gamma}{\cosh^2\gamma}, \quad c_2(\gamma) = \tanh\gamma,$$

$$p_{\max} = \frac{\pi F}{2\omega} = \frac{\pi\kappa}{2\gamma}.$$

At arbitrary $\alpha$, substituting $\tau = \sqrt{u/(1 + u)}$ in (5) yields the equation

$$z = \sqrt{u}F\left(\frac{1}{2}, \frac{3}{2} - \alpha; \frac{3}{2}; -u\right), \quad u = \chi^{-1/\alpha} - 1, \quad 0 < u < \infty, \qquad (20)$$

which explicitly defines the inverse function $z = z(\chi)$. The subsequent calculations using formulas (4) involve no difficulties. At $\alpha = 3/2$ and 1, expressions (18) and (19) readily

---

[5] Note that individual terms in the expressions for $f(\gamma)$ and $g(\gamma)$ become infinite when $\gamma \longrightarrow 0$ (in sum, they cancel out). Therefore, to obtain expansions in the adiabatic range $\gamma \ll 1$, it is more convenient to use (8) rather than exact formulas of type (11), (12), or (18).

follow from them, while at $\alpha = 2$ [a pole of the second order in $\varphi(t)$], we obtain

$$z = \frac{1}{2}(\sqrt{u(1+u)} + \operatorname{arcsinh}\sqrt{u}), \quad u = \chi^{-1/2} - 1. \tag{21}$$

(6) A generalization of the soliton pulse (12) is

$$\varphi(t) = (\cosh t)^{-\alpha}, \quad \alpha \geq 1. \tag{22}$$

Changing from $\tau$ to $\xi = \tan\tau$ in (5) yields

$$\chi = (1+\xi^2)^{-\alpha/2},$$

$$z = \int_0^{\xi}(1+x^2)^{(\alpha-2)/2}dx = \xi F\left(1-\frac{\alpha}{2}, \frac{1}{2}; \frac{3}{3}; -\xi^2\right),$$

$0 < \xi < \infty$. At $\alpha = 1$, considering that $F(1/2, 1/2; 3/2; -\xi^2) = \operatorname{arcsinh}\xi/\xi$, we derive $\chi(z) = 1/\cosh z$. For even $\alpha = 2, 4, \ldots$, the expression for $z(\xi)$ reduces to a polynomial: thus, $z = \xi$ and $\xi + (1/3)\xi^3$ at $\alpha = 2$ and 4.

(7) The preceding examples, except for example 1, belong to unidirectional pulses. Although any uniform field is the solution of the Maxwell equations, $\mathcal{E} \longrightarrow \mathcal{E}(t - x/c)$, and, hence, is (in principle) physically realizable, the integral $J = \int_{-\infty}^{\infty}\mathcal{E}(t)dt$ is either zero or numerically small for the fields commonly encountered in laser physics (see [25, 26]). As an example, consider

$$\varphi_\alpha(t) = \frac{d}{dt}\left[\frac{t}{(1+t^2)^{\alpha}}\right] = \frac{1-(2\alpha-1)t^2}{(1+t^2)^{\alpha+1}},$$

$$\alpha > 1/2,$$

$$\varphi_\alpha(t) = 1 - 3\alpha t^2 + \frac{5}{2}\alpha(\alpha+1)t^4 + \ldots, \quad t \longrightarrow 0, \tag{23}$$

$$\varphi_\alpha(t) \approx -(2\alpha-1)t^{-2\alpha} \longrightarrow 0, \quad t \longrightarrow \infty.$$

In this case, Eqs. (5) take the form

$$\xi = \frac{(1-\tau^2)^{\alpha+1}}{1+(2\alpha-1)\tau^2}, \quad z = \frac{\tau}{(1-\tau^2)^{\alpha}}. \tag{5''}$$

If $\alpha = 1/2$, then example (23) coincides with (18); at $\alpha = 1$, we have

$$\tau(\gamma) = \frac{2\gamma}{1+\sqrt{1+4\gamma^2}},$$

$$\chi(z) = \frac{2}{1+4z^2+\sqrt{1+4z^2}}, \tag{24}$$

$$f(\gamma) = \sqrt{1+\frac{1}{4\gamma^2}} - \frac{1}{\gamma} + \frac{\operatorname{arcsinh}2\gamma}{4\gamma^2},$$

while at arbitrary $\alpha$, the system of equations (5'') can be easily solved numerically.

For a monochromatic field $\varphi(t) = \cos t$, $J = 0$ as well. In that case, this integral has an unequivocal meaning when it is considered that the laser field, which is approximately uniform near the focus, is adiabatically switched off at infinity:

$$J = \lim_{\alpha\to 0}\int_0^{\infty}\exp\left[-\frac{(\alpha t)^2}{2}\right]\cos\omega t dt$$

$$= \lim_{\alpha\to 0}\sqrt{\frac{\pi}{2}}\alpha^{-1}\exp\left(-\frac{\omega^2}{2\alpha^2}\right) = 0.$$

Therefore, J ≈ 0 for $\alpha \ll \omega$ with an exponential accuracy.

(8) In the limit $\alpha \longrightarrow \infty$, examples 4 and 7 are equivalent (after the scaling $t \longrightarrow t/\sqrt{\alpha}$) to the pulses

$$\varphi(t) = \exp(-t^2),$$

$$(1-2t^2)\exp(-t^2) = \frac{d}{dt}[t\exp(-t^2)]. \tag{25}$$

Note that when the time scale is changed, $\varphi(t) \longrightarrow \varphi(\lambda t)$, the following scaling relations hold:

$$g(\gamma) \longrightarrow g(\lambda\gamma), \quad f(\gamma) \longrightarrow \lambda^{-1}f(\lambda\gamma),$$

$$c_{1,2}(\gamma) \longrightarrow \lambda^{-1}c_{1,2}(\lambda\gamma), \tag{26}$$

and $\chi(z) \longrightarrow \chi(\lambda z)$.

(9) To estimate the effect of higher harmonics in the laser pulse spectrum, we assume that

$$\varphi(t) = \frac{(1-\rho)^2\cos t}{1+\rho^2-2\rho\cos 2t}, \quad \rho_0 < \rho < 1, \tag{27}$$

where $\rho$ is the anharmonicity parameter,[6] $\rho_0 = -(\sqrt{2}-1)^2 = -0.1715$, with $\varphi(t+\pi) = -\varphi(t)$.

In this case,

$$\tau(\gamma) = a\operatorname{arcsinh}\{p^{-1}\tanh(p\gamma)\}, \tag{28}$$

$$\tau_s = \operatorname{arcsinh}(p^{-1}) = -\frac{1}{2}\ln\rho, \quad p = \frac{2\sqrt{\rho}}{1-\rho}. \tag{28'}$$

Function (27) has a pole at $\cos 2t = (\rho + \rho^{-1})/2$ or $t = i\tau_s$ (as $\rho \longrightarrow 0$, this point goes to infinity, and $\varphi(t) = \cos t$ is an integral function). As $\rho$ increases, the pulse becomes increasing sharp, turning into a sequence of $\delta$-shaped peaks of alternating polarity when $\rho \longrightarrow 1$:

$$\varphi(t) \approx \begin{cases} [1-\rho(1+4\sin^2 t)+O(\rho^2)]\cos t, & \rho \ll 1, \\ \sum\limits_{n=-\infty}^{\infty}(-1)^n\exp\left\{-\frac{1}{2}a_2(t-n\pi)^2\right\}, & \rho \longrightarrow 1, \end{cases} \tag{29}$$

---

[6] The intensities of adjacent odd harmonics in the pulse spectrum are related as $\rho^2$.

where $a_2 = (1 + 6\rho + \rho^2)/(1 - \rho)^2$, with $a_2 = 0$ at $\rho = \rho_0$. After simple, though cumbersome, calculations, we obtain

$$\chi(z) = \{\cosh(pz)\sqrt{1 + (1 + p^{-2})\sinh^2(pz)}\}^{-1}, \quad (30)$$

so that

$$\begin{cases} \chi(z) = 1 - (p^2 + 1/2)z^2 + \ldots, & z \longrightarrow 0, \\ \chi(z) \propto \exp(-2pz), & z \longrightarrow \infty. \end{cases} \quad (30')$$

Note that pulse (27) becomes a double-humped one for $\rho < \rho_0$, and $t = 0$ is not a point of $\varphi(t)$ maximum for it but a point of minimum.

(10) Finally, consider the following ansatz for the function $\chi$:

$$\chi(z) = (1 + z^2)^{-\mu}. \quad (31)$$

Calculating the integrals in (4) [26, 27] yields

$$\begin{aligned} g(\gamma) &= F\left(\frac{1}{2}, \mu; \frac{5}{2}; -\gamma^2\right) \\ &\equiv \frac{1}{1 + \gamma^2} F\left(\frac{1}{2}, \frac{5}{2} - \mu; \frac{5}{2}; \frac{\gamma^2}{1 + \gamma^2}\right), \\ b_1(\gamma) &= \frac{2}{3}\mu\gamma^2 F\left(\frac{3}{2}, \mu + 1; \frac{5}{2}; -\gamma^2\right), \\ b_2(\gamma) &= F\left(\frac{1}{2}, \mu; \frac{3}{2}; -\gamma^2\right). \end{aligned} \quad (32)$$

It can be shown [see (B.6)] that the pulse shape corresponding to (31) is characterized by the expansion

$$\varphi(t) = 1 - \mu t^2 + \frac{1}{6}(7\mu^2 - 3\mu)t^4 + \ldots, \quad (33)$$

for $t \longrightarrow 0$, while for $t \longrightarrow \infty$,

$$\varphi(t) \approx \begin{cases} 4\exp(-2t), & \mu = 1, \\ [2(\mu - 1)t]^{-\mu/(\mu - 1)}, & \mu > 1. \end{cases} \quad (33')$$

In particular, at $\mu = 1/2$, 1, and 3/2, we have $\varphi(t) = \cos t$, $1/\cosh^2 t$, and $(1 + t^2)^{-3/2}$ (the examples considered above), while $g(\gamma)$ has a remarkably simple form at $\mu = 5/2$:

$$\begin{aligned} g(\gamma) &= (1 + \gamma^2)^{-1/2}, \\ \tau(\gamma) &= \left(\gamma + \frac{2}{3}\gamma^3\right)(1 + \gamma^2)^{-3/2}. \end{aligned} \quad (34)$$

In that case, the coefficients $g_n$ from (8) are

$$g_n = \frac{3\Gamma(n + \mu)}{(2n + 1)(2n + 3)\Gamma(\mu)n!}. \quad (35)$$

Since $g_n \propto n^{\mu - 3}$ for $n \longrightarrow \infty$, series (8), (8'), and (8") converge at $|\gamma| < 1$. At $\mu = 1/2$, we derive an adiabatic expansion for the Keldysh function (11):

$$\begin{aligned} f(\gamma) &= \frac{2}{3}\left(\gamma - \frac{1}{10}\gamma^3 + \frac{9}{280}\gamma^5 - \frac{5}{336}\gamma^7 + \ldots\right), \\ f_n &= \frac{2}{3}g_n = \frac{(2n - 1)!!}{(2n + 1)(2n + 3)2^{n - 1}n!}, \end{aligned} \quad (36)$$

whose first terms were found previously [4–6].

On the other hand, all functions (32) decrease proportionally to $1/\gamma$ as $\gamma \longrightarrow \infty$:

$$\begin{aligned} & g(\gamma) \approx c\gamma^{-1}, \quad b_1(\gamma) \approx b_2(\gamma) \approx \frac{2}{3}c\gamma^{-1}, \\ & c = \frac{3\sqrt{\pi}\,\Gamma(\mu - 1/2)}{4\,\Gamma(\mu)}, \end{aligned} \quad (37)$$

in agreement with the behavior of the curves in Fig. 1.

The above examples demonstrate the ITM efficiency. An analytic solution is also possible for $\varphi(t) = (\cosh^2 t + \beta^2 \sinh^2 t)^{-1}$ and $[\cosh^2(\beta t) + (\sinh(\beta)/\beta)^2]^{-1}$ ($\beta$ is the parameter, $0 \le \beta < \infty$), for $\varphi(t) = \mathrm{cn}(t, q)$, where cn is the elliptic cosine, and others (see the table). For any pulse shape, including that taken directly from experimental data, $\chi(z)$ can be easily calculated numerically using the above equations, whereupon the problem reduces to quadratures.

## 4. NUMERICAL CALCULATIONS

Let us now discuss the results of our numerical calculations. The function $g(\gamma)$ for several pulse fields is presented in Fig. 1, which, for comparison, also shows this function for a monochromatic field (curve *1*). The notation is explained below. The curve numbers in Fig. 1 correspond to

$$\begin{aligned} & 1)\ \varphi = \cos t; \quad 2)\ \varphi = 1/\cosh^2 t; \\ & 3)\ \varphi = \exp(-t^2); \quad 4)\ \varphi = (1 + t^2)^{-3/2}; \\ & 5)\ \varphi = (1 + t^2)^{-1}; \quad 6)\ \varphi = (1 + t^4)\exp(-t^2); \\ & 7)\ \varphi = (1 + c_* t^4)\exp(-t^2). \end{aligned} \quad (38)$$

The time axis was scaled so that all pulses had the same curvature at the apex [$\varphi''(0) = -1$], which corresponds to a changeover from $\gamma$ to $\tilde{\gamma} = \sqrt{a_2}\gamma$, where $a_2$ is the coefficient in expansion (7). This is convenient for comparing variously shaped pulses: in the adiabatic range, the ionization probability does not depend on the form of $\varphi(t)$, see (9); such a dependence appears only beginning with terms of order $\gamma^4$. The function $f(\gamma)$ is shown in Fig. 2. We see from Figs. 1 and 2 that when passing from $\varphi(t) = \cos t$ to pulse fields of various forms concentrated in a finite time interval, $g(\gamma)$ and $f(\gamma)$

**Table**

| $\varphi(t)$ | $\chi(z)$ | $\tilde{g}_4 \times 100$ |
|---|---|---|
| $\cos t$ | $(1+z^2)^{-1/2}$ | 3.21 |
| $\cos^2 t$ | – | 2.86 |
| $1/\cosh^2 t$ | $(1+z^2)^{-1}$ | 2.14 |
| $(\cosh^2 t + \beta^2 \sinh^2 t)^{-1}$ | $[\cosh^2\beta z + (\sinh\beta z/\beta)^2]^{-1}$ | $1.43(1+\delta/2)$ |
| $[\cosh^2\beta t + (\sinh\beta t/\beta)^2]^{-1}$ | $(\cosh^2 z + \beta^2\sinh^2 z)^{-1}$ | $2.14(1-\delta/3)$ |
| $1/I_0(t)$ | – | 1.96 |
| $1/\cosh t$ | $1/\cosh z$ | 1.79 |
| $(1+t^2)^{-1}$ | $1/\cosh^2 z$ | 1.43 |
| $(1+t^2)^{-3/2}$ | $(1+z^2)^{-3/2}$ | 1.79 |
| $(1+t^2)^{-2}$ | see (21) | 1.96 |
| $\exp(-t^2)$ | see (13) | 2.50 |
| $(1-t^2)/(1+t^2)^2$ | $2/(1+4z^2+\sqrt{1+4z^2}$ | 2.38 |
| $(1-2t^2)e^{-t^2}$ | – | 2.98 |
| $\mathrm{cn}(t, q)$, $0 \le q \le 1$ | $[1+(\sinh qz/q)^2]^{-1/2}$ | $0.357(9-4q^2)$ |
| see (33) | $(1+z^2)^{-\mu}$ | $1.07(1+\mu^{-1})$ |

Note: Here, $\delta = 1/(1+\beta^2)$, $I_0(t)$ is the modified Bessel function, and cn is the elliptic cosine [24].

decrease, particularly at $\gamma > 1$ (rapidly varying fields). As a result, in view of conditions (2), the ionization probability increases sharply [because $\epsilon \ll 1$ and $K_0 \gg 1$; see formulas (3) and (10)]. Some of the curves in Fig. 1 refer to pulses of the type

$$\varphi(t) = (1+ct^4)\exp(-t^2), \quad a_2 = 2. \qquad (39)$$

At $c = 0$, the pulse is Gaussian (curve *3*); dashed curves *6* and *7* in Fig. 1 correspond to $c = 1$ and $c = c_* \approx 1.544$ (see Appendix B). In all the cases we considered, $g(\gamma)$ monotonically decreases with increasing Keldysh parameter $\gamma$.

We see from Fig. 2 that the behavior of $f(\gamma)$ at large $\gamma$ is directly related to the analytic properties of field function $\varphi(t)$ in the complex plane. More specifically, curves *1* and *3*, which correspond to the integral analytic functions $\varphi(t) = \cos t$ and $\exp(-t^2)$, rise as $\ln\gamma$ and

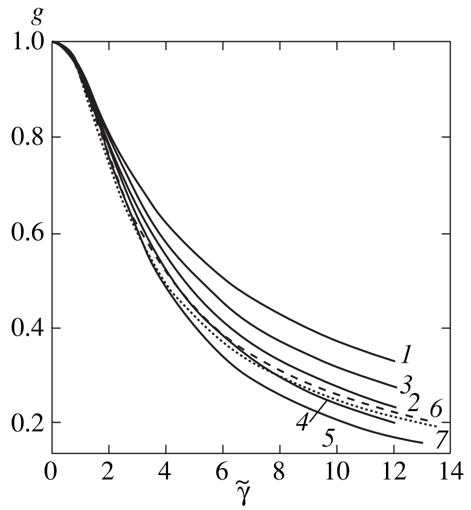

**Fig. 1.** Function $g(\gamma)$ for fields of form (1). The curve numbers *1–7* are explained in (38). The scaled variable $\tilde{\gamma} = \sqrt{a_2}\gamma$ is along the *x* axis.

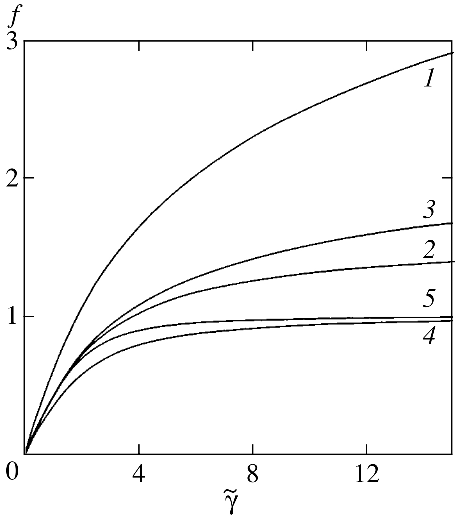

**Fig. 2.** Function $f(\gamma)$ from Eq. (10). The curve numbers are given in (38).

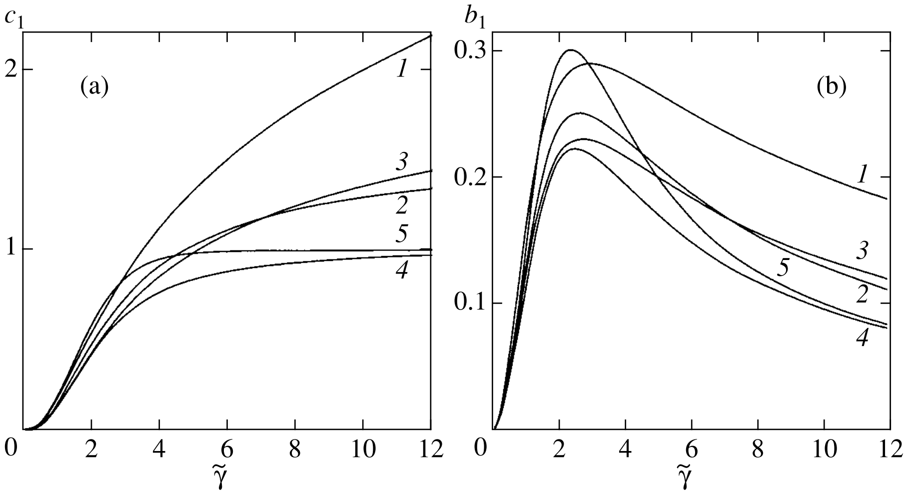

**Fig. 3.** (a) Coefficient $c_1(\gamma)$ for the same pulses as in the previous figure. Note that $c_1(\infty) = \pi/2$ for curve *2* and $c_1(\infty) = 1$ for curves *4* and *5*. (b) Coefficient $b_1(\gamma)$ in (3) for various pulses.

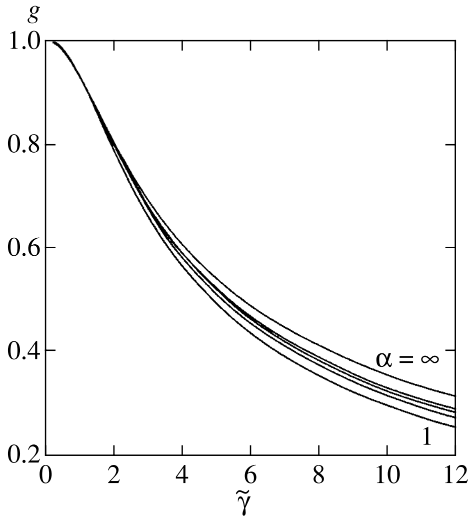

**Fig. 4.** Function $g(\gamma)$ for pulses of form (23) with $J = 0$. The curves correspond to $\alpha = 1, 3/2, 2, 5/2$, and $\infty$ (from bottom to top), $\tilde{\gamma} = \sqrt{6\alpha}\gamma$.

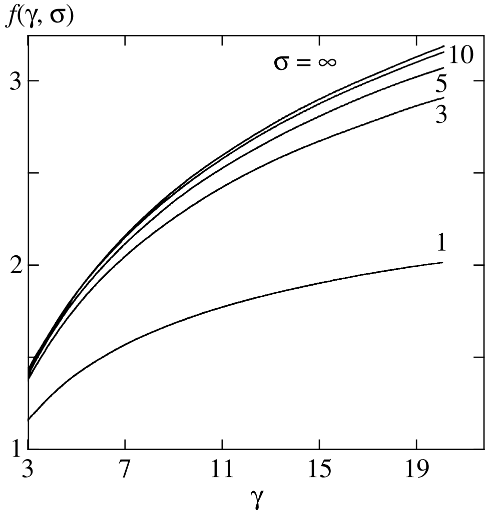

**Fig. 5.** Function $f(\gamma, \sigma)$ from (10) for a pulse field (40). The curves (from bottom to top) correspond to the following pulse-width parameters σ: $\sigma = 1, 3, 5, 10$, and $\infty$.

$\sqrt{\ln\gamma}$, while in the remaining cases, they approach a constant limit: $f(\gamma) \longrightarrow \tau_s$, where $t_s = i\tau_s$ is the position of the singularity of $\varphi(t)$ closest to the real axis: $\tau_s = \pi/2$ for curve *2* and $s = 1$ for curves *4* and *5*. This asymptotic limit [corresponding to formula (17) for the pulse spectrum] can be approached rapidly enough.

Plots of pulse-spectrum coefficient $c_1$ and coefficient $b_1$ against $\tilde{\gamma}$ are shown in Figs. 3a and 3b, respectively. The curves for $c_2(\gamma)$ are similar to those in Fig. 3, except the range of small γ, in which $c_1 \propto \gamma^3$ and $c_2 = \gamma + O(\gamma^3)$.

For pulses (14) and (23), we performed calculations at $\alpha = 1, 1.5, 2, 3$, and $\infty$ [$\alpha = \infty$ corresponds to passage to the limit (25)]. In both cases, $g(\tilde{\gamma})$ monotonically increases with exponent α, while the ionization probability decreases (Fig. 4). This may be because the weight of the high harmonics in $\varphi(t)$ with frequencies $\omega \gg 1/\tau_s$ decreases:

$$\int_0^\infty \frac{\cos\omega t}{[1+(t/\tau_s)^2]^\alpha}dt \propto (\omega\tau_s)^{-\alpha}\exp(-\omega\tau_s).$$

A comparison of Figs. 1 and 4 shows that there is no qualitative difference between unidirectional pulses of type (14) and pulses with the integral $J = 0$ for multiphoton ionization.

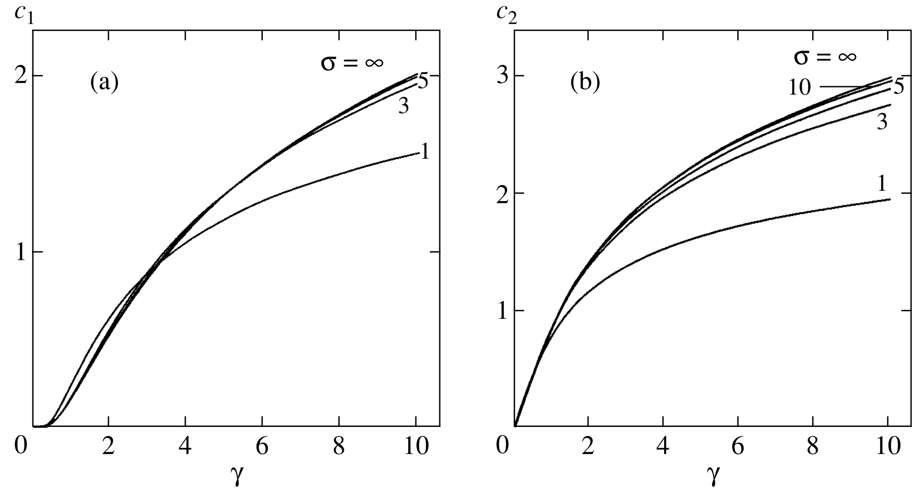

**Fig. 6.** Coefficients (a) $c_1$ and (b) $c_2$ of pulse spectrum (10) versus $\gamma$. The values of $\sigma$ in (40) are shown alongside the curves.

The next figures refer to a modulated electromagnetic pulse,

$$\varphi(t) = \exp\left(-\frac{t^2}{2\sigma^2}\right)\cos t, \qquad (40)$$

which is closer to an actual experiment. Here, at small $\gamma$,

$$g(\gamma) = 1 - \frac{1}{10\sigma^2}(1+\sigma^2)\gamma^2 + \frac{1}{280\sigma^4}(9\sigma^4 + 14\sigma^2 + 7)\gamma^4 + \dots . \qquad (41)$$

The pulse shortens with decreasing $\sigma$: its amplitude decreases by a factor $\delta \approx \exp(-2\pi^2/\sigma^2)$ in one period of the laser field, with $J \propto \exp(-\sigma^2/2) \longrightarrow 0$ for $\sigma \gg 1$. The functions $g(\gamma)$ and $f(\gamma)$ also decrease (see Fig. 5), causing the ionization probability to increase sharply. Physically, this is because the relative weight of high harmonics $\omega_n$ in the pulse spectrum increases. At $\omega_n > I$, these harmonics can ionize the atom even in the first order of perturbation theory (whereas for monochromatic light, there are no higher harmonics at all, and only the multiphoton ionization mechanism [4] remains). As we see from Fig. 5, this becomes noticeable at $\sigma \sim 5$–10 for field (40). Assuming (arbitrarily) the duration of a laser pulse to be the time during which its amplitude exceeds a fixed $\varepsilon$, we have for $\sigma \gg 1$

$$N \approx 0.1\sigma^2 \ln\frac{1}{\varepsilon};$$

i.e., $N \sim 0.2\sigma^2$ for $\varepsilon = 0.1$. Therefore, the light-pulse shortening begins to appreciably affect the ionization probability when the pulse spans $N \sim 5$–10 periods of the laser field. Regarding the shape of the pulse spectrum, it follows from Fig. 6a that the dependence of $c_1(\gamma, \sigma)$ on $\sigma$ may be disregarded if $\sigma \geq 3$ and, generally, it is less significant than for the function $f$, i.e., for the rate of atomic ionization. The same is also true for $c_2(\gamma, \sigma)$; see Fig. 6b.

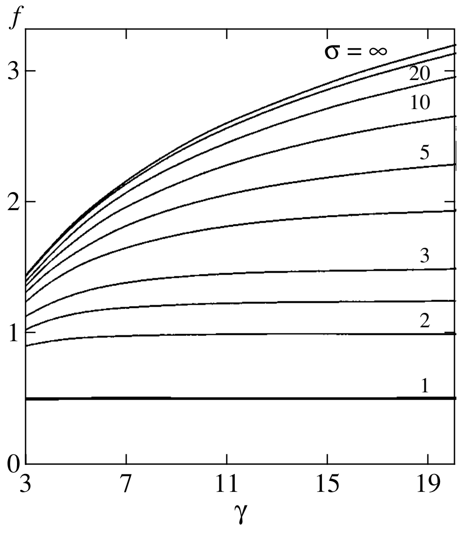

**Fig. 7.** Function $f(\gamma, \sigma)$ for a pulse with a Lorentz envelope (42). The curves (from bottom to top) correspond to $\sigma = 1$, 2, 2.5, 3, 4, 5, 6.67, 10, 20, and $\infty$ (monochromatic light).

Similar results were obtained for a Lorentz envelope, i.e., for[7]

$$\varphi(t) = [1 + (2t/\sigma)^2]^{-1}\cos t. \qquad (42)$$

As for (40), $f(\gamma, \sigma)$ decreases with decreasing $\sigma$ (i.e., with pulse shortening) at fixed $\gamma$ (Fig. 7). However, there is a qualitative difference between pulses (40) and (42) at large $\gamma$, which is related to the analytic properties of $\varphi(t)$. More specifically, $f(\gamma, \sigma)$ in Fig. 5 increases

---

[7] Here, $\sigma$ is equal to the width of the envelope at half its height. For field (40), its value is $2.35\sigma$.

(logarithmically) with $\gamma$, while for (42), it approaches a constant limit:[8] $f(\gamma, \sigma) \longrightarrow \tau_s = \sigma/2$ as $\gamma \longrightarrow \infty$, just as in the case of (14). This is because (40) is an integral function of $t$, while (42) has a pole at point $t_s = i\sigma/2$.

Finally, the function $g(\gamma)$ for a periodic field (27) calculated from Eqs. (4) and (30) are displayed in Fig. 8, which shows a sharp dependence on anharmonicity parameter $\rho$, with $g(\gamma)$ decreasing appreciably even at small $\gamma$. At $\rho$ close to unity, the ionization probability increases dramatically even at $\gamma \lesssim 1$.

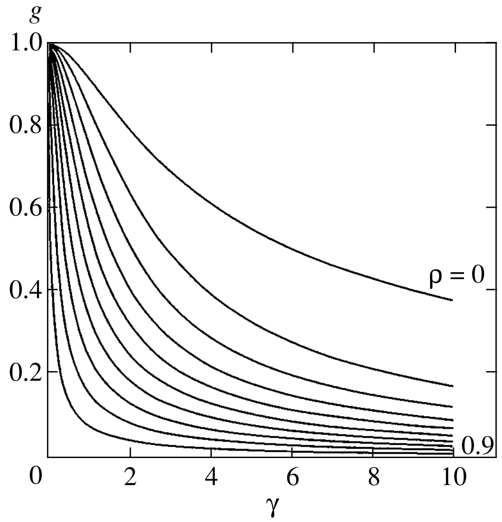

**Fig. 8.** Periodic field (27). $\rho = 0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8$, and $0.9$ for the curves from top to bottom.

Eqs. (4)–(6) allow all the quantities appearing in the semiclassical formulas (3) and (10) for $w_i(\mathbf{p})$ to be calculated for an arbitrary pulse $\mathscr{E}(t)$ and at any $\gamma$. This enables a detailed comparison of the multiphoton ionization theory with experiments in strong fields and under very short pulses.

## 5. TUNNELING INTERFERENCE IN THE ENERGY SPECTRUM

There is an interference effect in the photoelectron energy spectrum, which (for linearly polarized laser emission) was noted in [6] and has recently been studied experimentally [28, 29], where the phenomenon was referred to as the “tunneling interference”. In the case of a periodic field (1), for which $\varphi(t + T/2) = -\varphi(t)$, the equation for saddle points in the complex $t$ plane (or for the initial time in the ITM),

$$\mathbf{p}^2(t_k) = \left(p_\parallel + \frac{F}{\omega}\int_0^{t_k}\varphi(t)dt\right)^2 + p_\perp^2 = -\kappa^2, \qquad (43)$$

has the solutions

$$t_k \equiv \omega t'_k = k\pi + i\tau(\gamma\sqrt{1 + q_\perp^2} + i\gamma q_\parallel). \qquad (44)$$

Here, $k = 0, \pm 1, \pm 2, \ldots$, $\mathbf{q} = \mathbf{p}/\kappa$, $\mathbf{p}$ is the momentum of the electron as it emerges from under the barrier, $\tau(z)$ is the function introduced in (5) and (6), and $T$ is taken to be $2\pi$. At $p \ll \kappa$ and for $k = 0$, we have

$$\begin{aligned}\tau(\gamma, \mathbf{p}) &= -it_0 = \tau_0 + \frac{1}{2}\tau_1 q_\perp^2 + \tau_2 q_\parallel^2 \\ &\quad + i(\tau_1 - \tau_2 q_\perp^2 - \tau_3 q_\parallel^2)q_\parallel + O(q^4),\end{aligned} \qquad (45)$$

where $\tau_0 = \tau(\gamma)$ is the initial time of the subbarrier electron motion and

$$\tau_n(\gamma) = \frac{(-1)^{n+1}}{n!}\gamma^n\frac{d^n\tau(\gamma)}{d\gamma^n}, \quad n \geq 1. \qquad (45')$$

The amplitude $A_k$ of the electron transition from a bound state to a continuum state is determined by the action function $S$ calculated along the path from point $t_k$ to the real time axis (whereupon $\mathrm{Im}S(t)$ no longer changes). As we see from (45), the initial point $t_0$ for $p_\parallel \neq 0$ is displaced from the imaginary time axis, and the other saddle points $t_k$ are displaced similarly. In this case, a real shift in phase $\phi$ arises between the amplitudes $A_k$ and $A_{k+1}$ corresponding to the points $t_k$ and $t_{k+1}$ within one period of laser emission (for example, $A_0$ and $A_1$), which causes interference. Using the ITM [7], we obtain

$$\begin{aligned}A_k(t) &= \exp[iS_k(t)] \\ &= \exp\left\{-\frac{i}{2}\int_{t_k}^{t}[k^2 + \mathbf{p}^2(t')]dt'\right\},\end{aligned} \qquad (46)$$

$$\begin{aligned}A_{2k} &= \exp\left(-2\pi ik\frac{\varepsilon}{\omega}\right)A_0, \\ A_{2k+1} &= \exp\left[-i\left(\frac{\pi\varepsilon}{\omega} + \phi\right)\right]A_{2k}.\end{aligned} \qquad (47)$$

Here, the phase difference between $A_0$ and $A_{2k}$ accumulates as $t$ changes along the real time axis in $k$ periods ($\varepsilon$ is the system’s quasi-energy in a periodic field), and the phase $\phi$ arises when integrating (over imaginary time) the linear (in $p_\parallel$) term $2p_\parallel F\omega^{-1}h(\tau')$ that enters into $\mathbf{p}^2(t')$ in (46). To be more precise, the adjacent amplitudes $A_{2k}$ and $A_{2k+1}$ receive the phase factors $\exp(\pm i\phi/2)$. After the coherent addition of $A_k$, the transition probability becomes proportional to time[9] and the $\delta$ function expressing the energy conservation law for $n$-photon absorption emerges. Adding up the contributions from $2N$ saddle points (over

---

[8] As $\sigma$ increases, the range in which this asymptotics is established is displaced toward increasingly large $\gamma$, and we have $f(\gamma) = \ln 2\gamma - 1/2 + \ldots$ for $\sigma = \infty$.

[9] For a single pulse of the form $1/\cosh^2 t$ or $\exp(-t^2)$, only the total ionization probability over the entire time of pulse action can be determined.

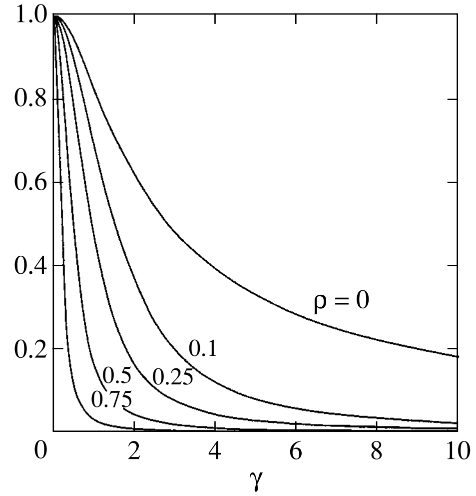

**Fig. 9.** Function $A(\gamma)$ determining the oscillation phase versus Keldysh parameter for a periodic field (27). The anharmonicity parameters $\rho$ = 0, 0.1, 0.25, 0.5, and 0.75 correspond to the curves (from top to bottom).

$N$ periods of the field), we pass to the probability of atomic ionization per unit time for $N \longrightarrow \infty$(see (B.7) and [30–33]):

$$w = \frac{\omega}{\pi}\sum_{n>\nu}\int\frac{d\mathbf{p}}{(2\pi)^3}[1+(-1)^n\cos\phi]w_i(\mathbf{p}) \times\delta\left(\frac{p^2}{2}-(n-\nu)\omega\right) \qquad (48)$$

(the linear regime $W_i \propto t$ that holds on time scales $\omega^{-1} \ll t \ll 1/w$); $w_i(\mathbf{p}) = |A_0|^2$ applies to a single pulse and is given by (3).

For ionization of the atomic $s$ level, the pulse spectrum during $n$-photon absorption takes the form

$$w(\mathbf{p}_n) \longrightarrow w(\mathbf{p}_n)[1+(-1)^n\cos\phi(\mathbf{p}_n)]/2. \qquad (49)$$

Here, $p_n = \sqrt{2\omega(n-\nu)}$, $\nu$ is the photoionization threshold (C.9), the multiphoton ionization probability $w(\mathbf{p})$ is defined in (3) or (10), the oscillation phase is

$$\phi(\mathbf{p}_n) = \frac{\kappa^2 p_\parallel}{F}[A(\gamma)+O(p^2/\kappa^2)], \qquad (50)$$

$$A = 2\gamma^{-2}\int_0^{\tau_0(\gamma)} h(\tau)d\tau = 1-\frac{1}{4}a_2\gamma^2+\ldots \quad (\gamma \longrightarrow 0), \qquad (51)$$

and the function $h(\tau)$ is defined in (5). In particular, $h(\tau) = \sinh\tau$ and $\tau_0 = \operatorname{arcsinh}\gamma$ for monochromatic laser light, whence [6]

$$A(\gamma) = 2\gamma^{-2}(\cosh\tau_0 - 1) = \frac{2}{1+\sqrt{1+\gamma^2}} \qquad (52)$$

(see the curve for $\rho = 0$ in Fig. 9). Since the characteristic oscillation phase is $\phi \sim \kappa^2 p_\parallel/F \sim 1/\gamma\sqrt{\epsilon}$ at small $\gamma$ and $\phi \sim \kappa p/\omega \sim \sqrt{K_0/\ln 2\gamma}$ for $\gamma \gg 1$, the number of oscillations is large (because $K_0 \gg 1$) even in the latter case.

It is easy to show, using (51), that for a periodic sequence of pulses of the form $\varphi(t) = 1/\cosh^2 t$ and $(1+t^2)^{-3/2}$, we have, respectively,

$$A(\gamma) = \frac{\ln(1+\gamma^2)}{\gamma^2} \text{ and } \frac{2}{1+\gamma^2+\sqrt{1+\gamma^2}} \qquad (53)$$

(in these cases, the phase $\phi$ for $\gamma \gtrsim \sqrt{K_0}$ is of the order of unity). Finally, for a periodic field (27), $h(\tau) = p^{-1}\operatorname{arccoth}(p\sinh\tau)$ and

$$A(\gamma) = \frac{2}{p\gamma^2}\int_0^{\tanh p\gamma}\frac{\operatorname{arccoth}x}{\sqrt{x^2+p^2}}dx, \quad p = \frac{2\sqrt{\rho}}{1-\rho} \qquad (54)$$

[at $\rho = 0$, we return to (52)]. As we see from Fig. 9, the oscillation phase rapidly decreases with increasing anharmonicity even at comparatively small $\gamma < 1$.

The energy spectrum is obtained by integrating expressions (3) and (49) over the photoelectron escape angles; as a result, the amplitude of the oscillating term significantly decreases [see formulas (B.14) and (B.15)]. This is in qualitative agreement with Fig. 1 from [28] (see the upper right part of this figure that refers to linearly polarized emission). Note that for $\varphi = \cos t$, the cubic (in momentum) term in the expansion of oscillation phase $\phi$ (at all $\gamma$) was also calculated [16].

The interference effect in the photoelectron pulse spectrum that arises when adding up the amplitudes $A_k$ was apparently first considered (in the multiphoton atomic ionization theory) in [6] (see formula (53) there and Section 8 in [16]). In a different physical situation (and for a different dispersion law $\varepsilon(\mathbf{p})$), similar phenomena are encountered in the semiconductor electrical breakdown theory [32], in the theory of charged boson and fermion pair production from vacuum in a variable electric field [30, 31], and in the problem of resonant atomic-level excitation in a strong electromagnetic field (in the two-level approximation [33]). Currently, the production of $e^+e^-$ pairs from vacuum in a strong electric field, which has previously been considered from a purely theoretical point of view [30, 31, 34–38], is again attracting attention, because projects are being developed to produce free-electron X-ray lasers based on the TESLA electron-positron collider and SLAC [39].

## 6. A REMARK ON THE ADIABATIC APPROXIMATION

The examples of pulse fields analyzed above allow the range of applicability of the adiabatic ($\gamma \ll 1$) approximation to be considered in the multiphoton ionization theory. After the scaling $t \longrightarrow \sqrt{a_2}\, t$ (see Section 4), we obtain

$$g(\gamma) = 1 - \frac{1}{10}\tilde{\gamma}^2 + \tilde{g}_4\tilde{\gamma}^4 + \ldots,$$
$$\tilde{g}_4 = \frac{10a_2^2 - a_4}{280a_2^2}, \quad \tilde{\gamma} = \sqrt{a_2}\gamma, \tag{55}$$

where $a_2$ and $a_4$ are the coefficients in (7), and similar expansions for the coefficients $b_1(\gamma)$ and $b_2(\gamma)$ of the pulse spectrum.

A dependence on specific pulse form $\varphi(t)$ begins to show up here in the terms of order $\gamma^4$. The coefficients $\tilde{g}_4$ (in contrast to $g_4$) depend only on pulse shape but not on pulse duration, and they are numerically small in all the cases we considered (see the table). Thus, for example, for (14) and (23), we have, respectively,

$$\tilde{g}_4 = \frac{1}{280}(7 - 3\alpha^{-1}) \text{ and } \frac{1}{168}(5 - \alpha^{-1}), \tag{56}$$

this coefficient ranges from 0.025 to 0.032 for field (40) at $0 < \sigma < \infty$ and from 0.014 to 0.032 for (42). This suggests that the range of applicability of the adiabatic approximation (which definitely holds for $\gamma \ll 1$) is generally extended up to $\gamma \gtrsim 1$, so the situation at $\gamma \sim 1$ is closer to the tunneling one rather than to the multiphoton one. This is also confirmed by the results of our numerical calculations presented in Figs. 1 and 4, in which the curves for different pulses are very close to each other at $\gamma \le 2$–3.

Note also that the asymptotics of $g_n$ [and the radius of convergence of the adiabatic expansions (8), (8'), and (8")] is determined by the nearest singularity of $\chi(z)$. If

$$\chi(z) \approx A(z^2 + a^2)^{-\mu}, \quad z \longrightarrow \pm ia, \tag{57}$$

then, in view of the expansion

$$(1 - z)^{-\mu} = \sum_n \frac{\Gamma(n + \mu)}{n!\Gamma(\mu)} z^n,$$

we have[10] for $n \longrightarrow \infty$

$$g_n \approx \frac{3A}{4\Gamma(\mu)} n^{\mu - 3} a^{-2(n + \mu)}, \quad n \longrightarrow \infty. \tag{58}$$

For monochromatic emission, $\chi(z) = (1 + z^2)^{1/2}$, i.e., $A = a = 1$, $\mu = 1/2$, and asymptotics (58) agrees with the expansion of exact coefficients (36):

$$f_n = \frac{2}{3} g_n = \frac{1}{2\sqrt{\pi}} n^{-5/2}\left[1 - \frac{17}{8n} + O\left(\frac{1}{n^2}\right)\right]. \tag{59}$$

Similarly, the parameters in (58) for pulse (14) are

$$\mu = \alpha, \quad a = \sqrt{\pi}\frac{\Gamma(\alpha - 1/2)}{2\Gamma(\alpha)}. \tag{60}$$

In both cases, power bifurcation points are the singularities of $g(\gamma)$ closest to zero. Thus, for example, we have for the Keldysh function (11)

$$g(\gamma) = \frac{3}{2\gamma} f(\gamma) = c_0 + c_1(1 + \gamma^2)$$
$$+ (1 + \gamma^2)^{3/2} + \ldots, \quad \gamma \longrightarrow \pm i$$

($c_0$ and $c_1$ are constants), in close agreement with (59).

## 7. CONCLUSION

(a) The equations of motion for a classical particle in a uniform field (1) can be integrated analytically. This allows formulas (3)-(5) for the atomic-level ionization probability, $W_i$, and for the photoelectron pulse spectrum, $w_i(\mathbf{p})$, to be derived (in the semiclassical approximation). The external electric field $\mathscr{E}(t')$ enters into these formulas via the function $\chi$; we formulated a simple algorithm for determining this function. Thereafter, the ionization probability can be calculated for an arbitrary pulse satisfying the applicability conditions for the semiclassical approximation, and, in many cases, an analytic solution can be obtained.

(b) We analyzed the dependence of the functions defining $w_i(\mathbf{p})$ on laser-pulse shape. The coefficients $c_{1,\,2}(\gamma)$ of the pulse spectrum were shown to be virtually independent of the duration of a very short pulse (40) if it spanned no fewer than three optical periods.

(c) We considered the effect of tunneling interference in a periodic laser field, which produces rapid oscillations in the photoelectron energy spectrum.

(d) The formulas derived above apply to the ionization of systems bound by short-range forces ($H^-$, $He^-$, etc.). As was shown in [9, 42], the Coulomb interaction between the emerging electron and the atomic core can be taken into account in terms of the semiclassical perturbation theory based on the Coulomb potential, which gives rise to a (large in magnitude) preexponential factor in the expression for the atomic-level ionization rate. Since the pulse-spectrum shape is determined mainly by the factors in the exponent, our results can be used not only to describe the ionization of $H^-$-type negative ions but also for neutral atoms (in any case, at $\gamma \lesssim 1$).

(e) Above, we imposed the condition $\varphi(-t) = \varphi(t)$, which actually means that $\varphi(t)$ is the real-valued analytic function $t^2$ whose nearest singularity on the semiaxis $-\infty < t^2 < 0$ lies at a finite distance from $t = 0$ [or has no singularities in a finite part of the $t$ plane at all, as in the case $\varphi = \cos t \equiv \cosh\sqrt{-t^2}$ or $\exp(-t^2)$]. All the functions considered above, $\varphi(t) = [(t^2 + a^2)(t^2 + b^2)]^{-1}$, $\mathrm{cn}(t, q)$, and many others satisfy this condition.

---

[10] Except for $\mu = 0, -1, -2, \ldots$ when $\chi(z)$ has a logarithmic singularity. The corresponding formulas can be found in [40, 41].

If this condition is not satisfied, then the imaginaries cannot be completely eliminated from the equations of subbarrier motion, and the formulas become more complex. However, the ionization probability $W_i$ can be calculated using the steepest descent method by determining saddle points of the action function in the complex plane. In this way, Keldysh [3] considered pulse fields $\varphi(t) = 3^{3/2}\sinh t/2\,\cosh^3 t$ and $t\exp[(1 - t^2)/2]$ (a soliton-like one-cycle pulse and a Gaussian one-cycle pulse [3]), for which $\varphi(-t) = -\varphi(t)$ and $J = 0$; the numerical factors were chosen here in such a way that $|\varphi(t_m)| = 1$ at the extrema. In these cases, oscillations are also predicted in the photoelectron spectrum, which owe their origin to the interference of two saddle points symmetric about the imaginary $t$ axis (and with equal Im$S$), a situation similar to that considered in Section 5 for a periodic field. Note that for fields of the form $\varphi(t) = 1/\cosh^2 t$ and $1/(1 + t^2)$, the exponential factors in the probability $W_i$ calculated in [3] and [9, 10] are equal, within the accuracy of the semiclassical approximation itself.

## ACKNOWLEDGMENTS

I wish to thank L.V. Keldysh for a helpful discussion at the initial stage of this study, for information about [3], and for a discussion of the results. I also wish to thank L.B. Okun, who pointed to [39], S.P. Goreslavsky, V.D. Mur, and S.V. Popruzhenko for discussions of various issues during the work and for pointing to [25–29], S.G. Pozdnyakov, who performed numerical calculations, and V.A. Gani and M.N. Markina for help in preparing the manuscript.

This study was supported in part by the Russian Foundation for Basic Research (project no. 01-02-16850).

## APPENDIX A

### *Expansions for Small and Large γ*

Substituting (7) in (5), we have

$$h(\tau) = \sum_{n=0}^{\infty}\frac{a_{2n}}{(2n+1)!}\tau^{2n+1} = z, \quad a_0 = 1. \tag{A.1}$$

Hence, using formulas for the inversion of a power series [27, 43], we obtain the expansion of $\tau$ in powers of $z$, whose coefficients are expressed in terms of $a_{2k}$. On the other hand,

$$\tau(z) = \sum_{n=0}^{\infty}(-1)^n\frac{\chi_n}{(2n+1)!}z^{2n+1}. \tag{A.2}$$

A comparison of these expressions gives formulas (7'). The next two coefficients are

$$\chi_3 = a_6 - 56a_4a_2 + 280a_2^3, \tag{A.3}$$
$$\chi_4 = -a_8 + 120a_6a_2 + 126a_4^2 - 4620a_4a_2^2 + 15400a_2^4$$

[using formulas from [43], we can also explicitly write out the coefficients $\chi_5$ and $\chi_6$ for an arbitrary field $\varphi(t)$, but these expressions are very cumbersome].

Thus, we have the following expansion for $z \longrightarrow 0$:

$$\chi(z) = 1 - \frac{1}{2}a_2z^2 + \frac{5}{12}(a_2^2 - 0.1a_4)z^4 - \frac{7}{18}\left(a_2^3 - \frac{1}{5}a_2a_4 + \frac{1}{280}a_6\right)z^6 + \dots, \tag{A.4}$$

whose substitution in (4) gives the expansions of $g(\gamma)$ and $b_{1,2}(\gamma)$ of the pulse spectrum in the adiabatic range.

Below, we also give a summary of asymptotics ($\gamma \longrightarrow \infty$) for the functions $f$, $c_1$, and $c_2$ in those cases where they can be obtained by expanding exact formulas of type (11) or (12).

For $\varphi(t) = \cos t$, it is clear that $\tau_s = \infty$ and

$$f(\gamma) = \ln 2\gamma - \frac{1}{2} + \frac{\ln 2\gamma}{2\gamma^2} + \dots, \tag{A.5}$$
$$c_1 = \ln 2\gamma - 1 + \frac{3}{4\gamma^2}, \quad c_2 = \ln 2\gamma + \frac{1}{4\gamma^2} + \dots$$

(here, the terms proportional to $\gamma^{-4}$ were discarded); for $\varphi(t) = 1/\cosh^2 t$ (soliton), we have $\tau_s = \pi/2$ and

$$f(\gamma) = \frac{\pi}{2} - 2\gamma^{-1} + \frac{\pi}{2\gamma^2} - \frac{2}{3\gamma^3} + \dots,$$
$$c_1 = \frac{\pi}{2} - 2\gamma^{-1} + O(\gamma^{-3}), \tag{A.6}$$
$$c_2 = \frac{\pi}{2} - \gamma^{-1} + \dots,$$

for $\varphi(t) = 1/\cosh t$,

$$f(\gamma) = \frac{\pi}{2} - \frac{\pi^3}{8\gamma^2} + 4\gamma^{-1}e^{-\gamma} + \dots, \tag{A.7}$$
$$c_1 \approx c_2 = \frac{\pi}{2} + O(e^{-\gamma}),$$

for $\varphi(t) = (1 + t^2)^{-3/2}$,

$$f(\gamma) = 1 - (\ln 2\gamma - 1/2)\gamma^{-2} + O(\gamma^{-4}), \tag{A.8}$$
$$c_1 = 1 - \frac{3}{2\gamma^2} + \dots, \quad c_2 = 1 - \frac{1}{2\gamma^2} + \dots,$$

for a Lorentz pulse, $\varphi(t) = (1 + t^2)^{-1}$,

$$\begin{gathered} f(\gamma) = 1 - \frac{\pi^2}{12\gamma^2} + 2\gamma^{-1}e^{-2\gamma} + \ldots, \\ c_1 \approx c_2 = 1 + O(e^{-2\gamma}) \end{gathered} \tag{A.9}$$

and, finally, for a Gaussian, $\varphi(t) = \exp(-t^2/2)$, we obtain

$$\begin{aligned} f(\gamma) &= \sqrt{2\ln\gamma} + \frac{\ln\ln\gamma}{\sqrt{2\ln\gamma}} + \ldots \\ &= \sqrt{2\ln(\gamma\ln\gamma)} + \ldots, \\ c_1(\gamma) &\approx c_2(\gamma) \approx \sqrt{2\ln\gamma}. \end{aligned} \tag{A.10}$$

Thus, for fields of type (14) and (23),

$$\lim_{\gamma \to \infty} \tau(\gamma) = \lim_{\gamma \to \infty} f(\gamma) = \tau_s, \tag{A.11}$$

where $t_s = i\tau_s$ is the singularity of field function $\varphi(t)$ closest to the real axis; the larger is the exponent $\alpha$, the more slowly this limit is approached.

## *APPENDIX B*

### *Auxiliary Functions*

(a) The function $w(x)$, or the Dawson integral, is defined as [24, 44]

$$\begin{aligned} w(x) &= \exp(-x^2)\int_0^x \exp(t^2)dt \\ &= \frac{\sqrt{\pi}}{2i}\operatorname{erf}(ix)\exp(-x^2), \end{aligned} \tag{B.1}$$

has the expansions

$$w(x) = \begin{cases} x - \frac{2}{3}x^3 + \frac{4}{15}x^5 - \ldots, & x \longrightarrow 0, \\ \frac{1}{2}x^{-1} + \frac{1}{4}x^{-3} + \frac{3}{8}x^{-5} + \ldots, & x \longrightarrow \infty, \end{cases} \tag{B.2}$$

and reaches a maximum of 0.54104 at $x_m = 0.9241\ldots$. Note the relation

$$\begin{gathered} w^{(n)}(x) = (-1)^n[H_n(x)w(x) - p_{n-1}(x)], \\ n = 1, 2, \ldots, \end{gathered} \tag{B.3}$$

where $p_{n-1}(x)$ is a polynomial of degree $n - 1$: $p_1 = 1$, $p_2 = 2x$, $p_3 = 4x^2 - 2x - 2$, etc.; and $H_n(x)$ is the Hermitian polynomial.

(b) Let us consider function (39). For $0 \le c \le 1$, it monotonically decreases at $t > 0$; for $c > 1$, minima ($\pm t_1$) and maxima ($\pm t_2$) appear in it, with $t_1^2 = 1 - \sqrt{1 - c^{-1}}$ and $t_2^2 = 1 + \sqrt{1 - c^{-1}}$. The height of the maxima increases with constant $c$, and $\varphi(\pm t_2) = \varphi(0) = 1$ at $c = c_* = 1.5441\ldots$. If $c > c_*$, then the electric field reaches the largest value not at zero but at $t = \pm t_2$. The dashed curves in Fig. 1 correspond to $c = 1$ and $c_*$, and curve *3* corresponds to $c = 0$ (Gaussian). Given that

$$\begin{aligned} h(\tau) &= \int_0^\tau e^{t^2}(1 + ct^4)dt \\ &= \left(1 + c\frac{\partial^2}{\partial\beta^2}\right)\int_0^\tau e^{\beta t^2}dt\Bigg|_{\beta = 1}, \end{aligned} \tag{B.4}$$

$$\int_0^\tau \exp(\beta t^2)dt = \frac{1}{\sqrt{\beta}}w(\sqrt{\beta}t)\exp(\beta\tau^2)$$

and using (B.3), we finally obtain

$$\begin{aligned} h(\tau) &= \left[\left(1 + \frac{3}{4}c\right)w(\tau)\right. \\ &\quad \left.+ c\left(\tau^2 - \frac{1}{2}\tau - \frac{1}{4}\right)\right]\exp(\tau^2), \end{aligned} \tag{B.5}$$

which determines $\chi(z)$ and $g(\gamma)$ for a pulse of form (39).

(c) The very short pulse $\varphi(t)$ corresponding to ansatz (31) can be specified parametrically:

$$\begin{gathered} \varphi = (1 + \xi^2)^{-\mu}, \quad t = \int_0^\xi (1 + x^2)^{\mu - 3/2}dx, \\ 0 < \xi < \infty, \end{gathered} \tag{B.6}$$

whence follow expansions (33) and (33') and, at $\mu = 1/2$, 1, and 3/2, the explicit expressions for $\varphi(t)$ given above (Section 3).

(d) When deriving formula (48), we used the relation

$$\begin{aligned} &\lim_{N \to \infty}\frac{1}{NT}\left|\sum_{k=0}^{2N-1} A_k\right|^2 \\ &= |A_0|^2 \lim_{N \to \infty}\frac{1}{NT}\left|\sum_{k=0}^{2N-1}\exp\{-i[k\alpha + (-1)^k\beta]\}\right|^2 \\ &= \omega|A_0|^2\sum_{n=-\infty}^{\infty}[1 + (-1)^n\cos2\beta]\delta(\alpha - n\pi) \end{aligned} \tag{B.7}$$

and the fact that $|A_k| = |A_0|$ for all $k$ in view of the condition $\varphi(t + T/2) = -\varphi(t)$. In our case, $\beta = \phi/2$ [see (50)] and

$$\begin{gathered} 2\alpha = \frac{1}{2}\int_0^T[\kappa^2 + \mathbf{p}^2(t')]dt' = \frac{\pi}{\omega}(p_\parallel^2 + p_\perp^2 + 2\omega\nu), \\ \alpha - n\pi = \frac{\pi}{\omega}\left[\frac{1}{2}p^2 - (n - \nu)\omega\right] = \frac{\pi}{2\omega}(p^2 - p_n^2), \end{gathered} \tag{B.8}$$

where ν is the photoionization threshold; for example,

$$\nu = K_0\left(1 + \frac{1}{2\gamma^2}\sum_{k \ge 1}\frac{f_k^2}{k^2}\right), \tag{B.9}$$

if $\varphi(t) = \sum_k f_k \cos kt$. The excess of ν over the multiquantum parameter $K_0$ is related to the energy of oscillatory motion of the emerging electron in the wave field.

(e) Two types of integrals are encountered when (49) is integrated over the photoelectron escape angles:

$$\begin{aligned}&\frac{1}{2}\int_0^{\pi}\exp\{-(a\sin^2\theta + b\cos^2\theta)\}\sin\theta d\theta \\ &\quad = e^{-b}\frac{w(\sqrt{a-b})}{\sqrt{a-b}},\end{aligned} \tag{B.10}$$

where $a = c_2 p_n^2/\omega$, $b = c_1 p_n^2/\omega$, $a > b$ [see (10) and (11)], and $w(x)$ is the Dawson function (B.1) and its generalization

$$\begin{aligned}&w(x, \lambda) \\ &= e^{-x^2}\int_0^x e^{t^2}\cos 2\lambda t dt = \mathrm{Re}[e^{2i\lambda x}w(x + i\lambda)]\end{aligned} \tag{B.11}$$

[here, $x$ and $\lambda$ are real, $\lambda = A(\gamma)\kappa^2 p_n/2F$]. The substitution $t = x\sqrt{1-s}$ gives the integral representation

$$\begin{aligned}&w(x, \lambda) \\ &= \frac{1}{2}x\int_0^1 e^{-sx^2}\cos(2\lambda x\sqrt{1-s})(1-s)^{-1/2}ds,\end{aligned} \tag{B.12}$$

whence

$$w(x, \lambda) = x - \frac{2}{3}(1 + \lambda^2)x^3 + O(x^5), \quad x \longrightarrow 0, \tag{B.13}$$

$$\begin{aligned}w &= \frac{1}{2x}\left[\cos 2\lambda x + \frac{\lambda}{x}\sin 2\lambda x\right. \\ &\quad \left.+ \frac{1 - 2\lambda^2}{2x^2}\cos 2\lambda x + \dots\right], \quad x \longrightarrow \infty,\end{aligned} \tag{B.14}$$

and for $\rho \equiv x^2 + \lambda^2 \longrightarrow \infty$, we have

$$w(x, \lambda) = \frac{x\cos 2\lambda x + \lambda\sin 2\lambda x}{2(x^2 + \lambda^2)} + O(\rho^{-3}). \tag{B.15}$$

Thus, this function rapidly oscillates and decreases at $\lambda x \gg 1$, which accounts for the significant reduction in oscillation amplitude in the electron energy spectrum [28] compared to formula (49), where this amplitude reaches 100%.

## *APPENDIX C*

### *Asymptotics of the Function f(γ) for γ ⟶ ∞*

Using [24, 27], we obtain

$$\begin{aligned}&F\left(\alpha, \frac{1}{2}; \frac{3}{2}; \tau^2\right) \\ &= \begin{cases}\operatorname{arccoth}\tau/\tau, & \alpha = 1, \\ (1-\tau^2)^{-1/2}, & \alpha = 3/2, \\ [\operatorname{arccoth}\tau + \tau/(1-\tau^2)]/2\tau, & \alpha = 2, \\ \left(1 - \frac{2}{3}\tau^2\right)(1-\tau^2)^{-3/2}, & \alpha = 5/2,\end{cases}\end{aligned} \tag{C.1}$$

and, from the recurrent relation,

$$\begin{aligned}&F\left(\alpha + 1, \frac{1}{2}; \frac{3}{2}; \tau^2\right) = \frac{1}{2\alpha} \\ &\times\left[(1-\tau^2)^{-\alpha} + (2\alpha - 1)F\left(\alpha, \frac{1}{2}; \frac{3}{2}; \tau^2\right)\right].\end{aligned} \tag{C.2}$$

For arbitrary $\alpha > 1$ and $x \longrightarrow 1$,

$$F\left(\alpha, \frac{1}{2}; \frac{3}{2}; x\right) = [2(\alpha - 1)(1 - x)^{\alpha - 1}]^{-1} + \dots,$$

therefore, (16) for the initial time $\tau_0(\gamma)$ directly follows from the equation $h(\tau_0) = \gamma$.

The function $\chi(z)$ has asymptotics (15') for $z \longrightarrow \infty$. If $\alpha < 3/2$ in (14), then $\alpha/(\alpha - 1) > 3$ and

$$\begin{aligned}&f(\gamma) = \tau_s - a\gamma^{-2} + \dots, \quad a = \int_0^{\infty}\chi(z)z^2 dz, \\ &\qquad \gamma \longrightarrow \infty\end{aligned} \tag{C.3}$$

(for example, $\tau_s = 1$ and $a = \pi^2/12$ for $\alpha = 1$), as well as

$$\begin{aligned}c_1(\gamma) &= \tau_s - \alpha c_{\infty}\gamma^{-\nu} + \dots, \\ c_2(\gamma) &= \tau_s - (\alpha - 1)c_{\infty}\gamma^{-\nu} + \dots,\end{aligned} \tag{C.4}$$

where $\chi(z) \approx c_{\infty}z^{-\alpha/(\alpha - 1)}$ and $\nu = 1/(\alpha - 1) > 2$. Hence follows the asymptotic formula (17).

If $\alpha = 3/2$, then $\chi(z) \propto z^{-3}$ and the integral (C.3) logarithmically diverges at the upper limit. Naturally, a correction proportional to $\ln\gamma/\gamma^2$ appears here [see (A.8)]. Finally, for $\alpha > 3/2$,

$$\begin{aligned} \tau(\gamma) &= \tau_s - k\gamma^{-\nu} + \ldots, \\ f(\gamma) &= \tau_s - k_1\gamma^{-\nu} + \ldots, \end{aligned} \qquad \text{(C.5)}$$

with

$$\begin{aligned} \nu &= \frac{1}{\alpha - 1} < 2, \\ k &= \frac{2\alpha - 2}{2\alpha - 3}[2^{\alpha}(\alpha - 1)]^{-1/(\alpha - 1)}, \\ E_1 &= \frac{2(\alpha - 1)}{2\alpha - 3}k. \end{aligned} \qquad \text{(C.6)}$$

We thus determined the asymptotics of $f(\gamma)$ for $\gamma \longrightarrow \infty$ for pulse (14) and established its relationship to the nearest singularity of $\varphi(t)$, which specifies the pulse shape, in the complex $t$ plane. We pass from (14) to the general case of a power singularity,

$$\varphi(t) \approx (1 + t^2/\tau_s^2)^{-\alpha}, \quad t \longrightarrow i\tau_s, \qquad \text{(C.7)}$$

by using the scaling relations (26).

In the opposite case, $\gamma \longrightarrow 0$, the higher orders of adiabatic expansions (8) and their radius of convergence depend on the singularity of $\chi(z)$ closest to zero. Thus, the analytic properties of $\chi$ and $\varphi$ manifest themselves at small and large $\gamma$, respectively.

## REFERENCES

1. N. Bloembergen, Rev. Mod. Phys. **71**, S283 (1999).
2. T. Brabec and F. Krausz, Rev. Mod. Phys. **72**, 545 (2000).
3. L. Keldysh, *Multiphoton Ionization by a Very Short Pulse* (in press).
4. L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1945 (1964) [Sov. Phys. JETP **20**, 1307 (1964)].
5. A. I. Nikishov and V. I. Ritus, Zh. Éksp. Teor. Fiz. **50**, 255 (1966) [Sov. Phys. JETP **23**, 168 (1966)].
6. A. M. Perelomov, V. S. Popov, and M. V. Terent'ev, Zh. Éksp. Teor. Fiz. **50**, 1393 (1966) [Sov. Phys. JETP **23**, 924 (1966)]; Zh. Éksp. Teor. Fiz. **51**, 309 (1966) [Sov. Phys. JETP **24**, 207 (1967)].
7. V. S. Popov, V. P. Kuznetsov, and A. M. Perelomov, Zh. Éksp. Teor. Fiz. **53**, 331 (1967) [Sov. Phys. JETP **26**, 222 (1968)].
8. A. I. Nikishov, Tr. Fiz. Inst. Akad. Nauk SSSR **111**, 152 (1979).
9. V. S. Popov, Laser Phys. **10**, 1033 (2000).
10. V. S. Popov, Preprint No. 39-00, ITÉF (Institute of Theoretical and Experimental Physics, Moscow, 2000); Pis'ma Zh. Éksp. Teor. Fiz. **73**, 3 (2001) [JETP Lett. **73**, 1 (2001)].
11. V. S. Popov, B. M. Karnakov, and V. D. Mur, Zh. Éksp. Teor. Fiz. **113**, 1579 (1998) [JETP **86**, 860 (1998)].
12. R. P. Feynman, Rev. Mod. Phys. **20**, 367 (1948).
13. R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965; Mir, Moscow, 1968).
14. A. I. Baz', Ya. B. Zel'dovich, and A. M. Perelomov, *Scattering, Reactions and Decays in Nonrelativistic Quantum Mechanics* (Nauka, Moscow, 1971, 2nd ed.; Israel Program for Scientific Translations, Jerusalem, 1966), p. 225.
15. A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Énergoatomizdat, Moscow, 1986; Springer-Verlag, Berlin, 1985).
16. V. S. Popov, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 493 (1999) [JETP Lett. **70**, 502 (1999)]; Zh. Éksp. Teor. Fiz. **118**, 56 (2000) [JETP **91**, 48 (2000)].
17. V. S. Popov, Preprint No. 17-00, ITÉF (Institute of Theoretical and Experimental Physics, Moscow, 2000).
18. V. D. Mur, S. V. Popruzhenko, and V. S. Popov, Zh. Éksp. Teor. Fiz. **119**, 893 (2001) [JETP **92**, 777 (2001)].
19. A. V. Gaponov and M. A. Miller, Zh. Éksp. Teor. Fiz. **34**, 242 (1958) [Sov. Phys. JETP **7**, 168 (1958)].
20. S. P. Goreslavskiĭ, N. B. Narozhnyĭ, and V. P. Yakovlev, Pis'ma Zh. Éksp. Teor. Fiz. **46**, 173 (1987) [JETP Lett. **46**, 219 (1987)].
21. S. P. Goreslavsky, N. B. Narozhny, and V. P. Yakovlev, J. Opt. Soc. Am. B **6**, 1752 (1989).
22. R. R. Freeman, P. H. Bucksbaum, and T. J. McIlrath, IEEE J. Quantum Electron. **24**, 1461 (1988).
23. J. H. Eberly, J. Javanainen, and K. Rzazewsky, Phys. Rep. **204**, 331 (1991).
24. *Handbook of Mathematical Functions,* Ed. by M. Abramowitz and I. Stegun (Dover, New York, 1965; Nauka, Moscow, 1979).
25. E. G. Bessonov, Zh. Éksp. Teor. Fiz. **80**, 852 (1981) [Sov. Phys. JETP **53**, 433 (1981)].
26. R. Grobe and M. V. Fedorov, Laser Phys. **3**, 265 (1993).
27. H. B. Dwight, *Tables of Integrals and Other Mathematical Data* (Dover, New York, 1947).
28. G. G. Paulus, F. Zacher, H. Walther, *et al.*, Phys. Rev. Lett. **80**, 484 (1998).
29. W. Becker, M. Kleber, A. Lohr, *et al.*, Laser Phys. **8**, 56 (1998).
30. V. S. Popov, Pis'ma Zh. Éksp. Teor. Fiz. **13**, 261 (1971) [JETP Lett. **13**, 185 (1971)]; Pis'ma Zh. Éksp. Teor. Fiz. **18**, 435 (1973) [JETP Lett. **18**, 255 (1973)]; Yad. Fiz. **19**, 1140 (1974) [Sov. J. Nucl. Phys. **19**, 584 (1974)].

31. M. S. Marinov and V. S. Popov, Fortschr. Phys. **25**, 373 (1977).
32. Yu. A. Bychkov and A. M. Dykhne, Zh. Éksp. Teor. Fiz. **58**, 1735 (1970) [Sov. Phys. JETP **31**, 928 (1970)].
33. D. F. Zaretskiĭ and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **66**, 537 (1974) [Sov. Phys. JETP **39**, 257 (1974)].
34. J. Schwinger, Phys. Rev. **82**, 664 (1951).
35. E. Brezin and C. Itzykson, Phys. Rev. D **2**, 1191 (1970).
36. N. B. Narozhnyĭ and A. I. Nikishov, Yad. Fiz. **11**, 1072 (1970) [Sov. J. Nucl. Phys. **11**, 596 (1970)]; Zh. Éksp. Teor. Fiz. **65**, 862 (1973) [Sov. Phys. JETP **38**, 427 (1973)].
37. V. S. Popov, Zh. Éksp. Teor. Fiz. **61**, 1334 (1971) [Sov. Phys. JETP **34**, 709 (1971)].
38. V. M. Mostepanenko and V. M. Frolov, Yad. Fiz. **19**, 885 (1974) [Sov. J. Nucl. Phys. **19**, 451 (1974)].
39. A. Ringwald, Preprint DESY 01-024; hep-ph/0103185 (2001).
40. E. C. Titchmarsh, *The Theory of Functions* (Oxford Univ. Press, London, 1939; Gostekhizdat, Moscow, 1951).
41. V. S. Popov, Zh. Éksp. Teor. Fiz. **47**, 2229 (1964) [Sov. Phys. JETP **20**, 1494 (1964)].
42. A. M. Perelomov and V. S. Popov, Zh. Éksp. Teor. Fiz. **52**, 514 (1967) [Sov. Phys. JETP **25**, 336 (1967)].
43. C. E. van Orstrand, Philos. Mag. **19**, 366 (1910).
44. K. A. Karpov, *Tables of w(z) Function on the Complex Plane* (Akad. Nauk SSSR, Moscow, 1954).

*Translated by V. Astakhov*

Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 295–300.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 333–339.
Original Russian Text Copyright © 2001 by Gridchin, Popov, Smirnova.

ATOMS, SPECTRA, RADIATION

# Angular Distributions of the Fragments from Coulomb Explosions of Diatomic Molecules in Intense Laser Fields

**V. V. Gridchin, A. M. Popov, and O. V. Smirnova***
*Moscow State University, Vorob'evy gory, Moscow, 119899 Russia*
**e-mail: smirnova@mics.msu.su*

Received February 28, 2001

**Abstract**—The angular distributions of the fragments from a Coulomb explosion of a diatomic heteronuclear molecule during multielectron dissociative ionization in a superintense field are considered in terms of classical mechanics. The patterns of angular distributions of the Coulomb explosion fragments are shown to differ in different ranges of laser pulse parameters. In particular, there are two distinct modes of fragment separation: separation in a Coulomb field and separation in the field of an effective "fragment+field" potential. The effective potential includes both the force of Coulomb repulsion between the fragments and the period-averaged force exerted on the system by the field; it can be determined by using the Kramers–Henneberger method. The limits of applicability of the Kramers–Henneberger method to the problem in question are discussed. These limits specify the range of field parameters in which the fragments fly apart in a direction perpendicular to the field for the initially arbitrary orientation of the molecular axis relative to the field. *© 2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

One of the most efficient methods for theoretically describing the dynamics of atomic systems in superintense fields is the Kramers–Henneberger method [1, 2].

The idea behind the Kramers–Henneberger method is to apply the following transformation [1] to the initial Hamiltonian of an atom in a laser field

$$\hat{H} = \frac{1}{2}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + V(\mathbf{r}), \quad (1)$$

where

$$\mathbf{A} = A_0\mathbf{e}_x\sin\omega t, \quad A_0 = -Ec/\omega,$$

$$S_{KH} = \exp\left(\frac{i}{c}\mathbf{p}\int_0^t \mathbf{A}(t')dt'\right)$$

$$\times\exp\left(-\frac{i}{2c^2}\int_0^t A^2(t')dt'\right).$$

This transformation reduces Hamiltonian (1) to

$$\hat{H}_{KH} = \frac{p^2}{2} + V(\mathbf{r} + \mathbf{e}_x a_e\cos\omega t), \quad (2)$$

where $a_e = E/\omega^2$ is the oscillation amplitude of a free electron in the laser field; and $E$ and $\omega$ are the field strength and frequency, respectively. Below, we use the atomic system of units, $m_e = \hbar = e = 1$. In the Kramers–Henneberger approximation, the time-dependent potential in Hamiltonian (2) is substituted with the Kramers–Henneberger period-averaged potential $V_{KH}(\mathbf{r}, a_e)$. This approximation is valid if the effect of the corrections

$$\delta V = V(\mathbf{r} + \mathbf{e}_x a_e\cos\omega t) - V_{KH}(\mathbf{r}, a_e)$$

is marginal. In this case, some quantities, for example, the ionization rate and polarizability, can be calculated by using the perturbation theory, and the system's exact quasi-energies are well approximated by the steady-state energies.

Presently, the properties of the Kramers–Henneberger potential, eigenfunctions, and eigenstates are well understood [3-7]. The Kramers–Henneberger potential virtually coincides with the initial atomic potential for $a/a_e \gg 1$, where $a$ is the scale size of the atomic potential. As the oscillation amplitude increases, $a_e > a$, the Kramers–Henneberger potential acquires a double-well structure and extends along the electric vector of an electromagnetic wave. Since the Kramers–Henneberger potential is the central object of the Kramers–Henneberger formalism, which characterizes the rearrangement of atomic states in weak high-frequency and strong fields and is used to describe adiabatic stabilization, probing the structure of this object in a real experiment is of considerable interest. Computer simulations that allow the structure of the Kramers–Henneberger potential to be determined are discussed in [8, 9].

Here, we show that the angular distributions of the fragments from Coulomb explosions of diatomic heteronuclear molecules during dissociative ionization in an intense laser field are determined by the structure of the Kramers–Henneberger potential.

## 2. MODELING DISSOCIATIVE IONIZATION

The dissociative ionization of molecules in intense laser fields has been studied extensively both experimentally [10–22] and theoretically [23–32]. A model in which the entire variety of the events constituting the pattern of dissociative ionization reduces to two consecutive events, (1) electron removal and (2) Coulomb explosion of the molecular ion produced by the electron removal, is commonly used in theoretical treatment to simplify this process. Thus, the evolution of dissociative ionization is determined by the competition of the above effects, which give different contributions to this process in different ranges of laser-pulse parameters.

The most important characteristic of dissociative ionization is the angular distribution of Coulomb explosion fragments. Experimental results suggest that the angular distributions of the fragments from molecular dissociation in a strong, linearly polarized field have a sharp peak in the direction of the field polarization axis (see, e.g., [12, 13, 15, 17, 19]). The sharp anisotropy in angular distributions is interpreted as resulting from the dynamic alignment of molecules [12, 33] or from a significant increase in dissociation cross sections with an increasing degree of molecular alignment with the field [15]. These processes obviously take place at the first stage of dissociative ionization.

In this paper, we consider the dissociation of a heteronuclear HA molecule (H is a hydrogen atom, and A is an atom of a different element, for example, deuterium, chlorine, bromine, and the like) by an optical pulse of intensity $P \geq 10^{19}$ W cm$^{-2}$ and duration 100 fs. One might expect the two-electron ($q$ + 1-electron) ionization of the molecule in such intense fields to take place virtually instantaneously. Therefore, it seems reasonable to focus attention on the second stage of dissociative ionization, the Coulomb explosion of the molecular ion $HA^{(q+1)+}$ with $q \geq 1$. Thus, we consider here the processes that can affect the formation of angular distributions at the stage of the Coulomb explosion.

We consider the dissociation problem in terms of classical mechanics. Under the assumption of sudden electron removal at time $t = 0$, the dynamics of ions, $H^+$ and $A^{(q+1)+}$, is described by the equations

$$\mu\frac{dz}{dt} = p_\rho, \quad \frac{dp_\rho}{dt} = -\frac{\partial}{\partial\rho}\left(\frac{1}{\sqrt{\rho^2+z^2}} + \frac{L_z^2}{2\mu\rho^2}\right),$$

$$\mu\frac{dz}{dt} = p_z, \quad \frac{dp_z}{dt} = -\frac{\partial}{\partial z}\frac{1}{\sqrt{\rho^2+z^2}} + \gamma E(t)$$

with the initial conditions

$$\rho(0) = R_0\sin\theta_0, \quad z(0) = R_0\cos\theta_0,$$

$$p_\rho = 0, \quad p_z = 0,$$

where

$$E(t) = \begin{cases} E_0\cos(\omega t)\sin^2\left(\dfrac{\omega t}{20}\right), & 0 \leq t \leq 5T_\omega, \\ E_0\cos\omega t, & 5T_\omega < t \leq 85T_\omega, \\ E_0\cos(\omega t)\sin^2\left(\dfrac{\omega t}{20}\right), & 85T_\omega < t \leq 90T_\omega, \\ 0, & t > 90T_\omega, \end{cases} \tag{3}$$

$$\gamma = \left(\frac{1}{M_1} - \frac{q}{M_2}\right)\mu, \tag{4}$$

$L_z$ is the $z$ component of the momentum moment, $T_\omega = 2\pi/\omega$, $M_1$ is the proton mass, $M_2$ is the mass of $A^{q+}$, $\mu$ is the reduced ion mass, $\rho$ and $z$ are the components of the vector that describes the relative motion of ions, $p_p$ and $p_z$ are the components of the momentum vector of the relative motion, $R_0$ is the equilibrium internuclear distance, and $\theta_0$ is the angle between the molecular axis and the field polarization vector. Here, we use cylindrical coordinates.

The angular distributions of the fragments from the Coulomb explosion of the HD molecule ($M_2 = 2M_1$, $q = 1$, $R_0 = 1.5$) for various field parameters are shown in Fig. 1. The angle

$$\theta_{\text{out}} = \lim_{t\to\infty}\theta(t),$$

$$\theta(t) = \begin{cases} \arctan(\rho(t)/z(t)), & z \geq 0, \\ \pi + \arctan(\rho(t)/z(t)), & z < 0, \end{cases}$$

which characterizes the direction of ion separation on completion of the field pulse, is measured from the $z$ axis. The centrifugal potential affects the angular distributions only slightly, at least for $0 \leq L_z^2 \leq L_{\max}^2$. For $L_{\max}$, we used an estimate

$$\frac{L_{\max}(L_{\max}+1)}{2\mu R_0^2} \approx kT,$$

$kT \approx 0.025$ eV. The pattern of the angular distributions significantly depends on field parameters and can be described in terms of the Kramers–Henneberger approximation.

Indeed, the Hamiltonian of our problem in the center-of-mass system, which describes the relative motion of the Coulomb explosion fragments, is

$$H = \frac{1}{2}\left(\frac{1}{M_1} + \frac{1}{M_2}\right)p^2 - \frac{\mathbf{pA}}{c}\left(\frac{1}{M_1} - \frac{q}{M_2}\right) + \frac{q}{r} + \frac{A^2}{2c^2}\left(\frac{1}{M_1} + \frac{q^2}{M_2}\right). \tag{5}$$

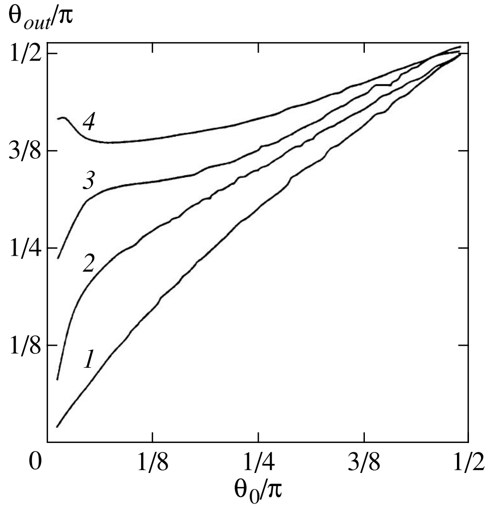

**Fig. 1.** The HD molecule. The changes in angular distributions when passing from Coulomb to the Kramers–Henneberger mode of ion separation. $\theta_{\rm out}$ is the angle between the direction of motion of the Coulomb explosion fragments on completion of the field pulse and the field polarization axis, $\theta_0$ is the angle between the molecular axis and the field polarization vector. $P = 10^{19}$ W cm$^{-2}$, $\omega = 1$ eV (*1*); $P = 2 \times 10^{20}$ W cm$^{-2}$, $\omega = 1$ eV (*2*); $P = 2 \times 10^{21}$ W cm$^{-2}$, $\omega = 1$ eV (*3*); and $P = 9 \times 10^{22}$ W cm$^{-2}$, $\omega = 9$ eV (*4*).

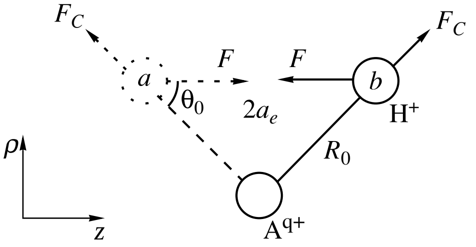

**Fig. 2.** To the discussion of ion separation dynamics in the field of the Kramers–Henneberger potential. $F_C$ is the force of Coulomb repulsion between H$^+$ and A$^{q+}$, $F$ is the force exerted on H$^+$ by the laser field.

After the Kramers–Henneberger transformation (it is also defined in classical mechanics [34–39]), Hamiltonian (5) takes the form

$$H = \frac{p_\rho^2 + p_z^2}{2\mu} + \frac{L_z^2}{2\mu\rho^2} + V_{KH}(\rho, z, a_e) + \sum_n V_n(\rho, z, a_e)e^{in\omega t}. \qquad (6)$$

The Hamiltonian in Eq. (6) is written in cylindrical coordinates, and

$$V_{KH}(\rho, z, a_e) = \frac{1}{2\pi}\int_0^{2\pi} \frac{q}{\sqrt{\rho^2 + (z + a_e\cos\varphi)^2}}d\varphi, \qquad (7)$$

$$V_n(\rho, z, a_e) = \frac{1}{2\pi}\int_0^{2\pi} \frac{q}{\sqrt{\rho^2 + (z + a_e\cos\varphi)^2}}e^{-in\varphi}d\varphi,$$

$$a_e = \frac{E}{\omega^2}\left(\frac{1}{M_1} - \frac{q}{M_2}\right).$$

In the Kramers–Henneberger approximation, the time-dependent term in Hamiltonian (6) can be disregarded. Thus, the separation dynamics of the Coulomb explosion fragments is determined by the structure of the Kramers–Henneberger potential and, hence, significantly depends on field parameters. The Kramers–Henneberger potential (7) is [40]

$$V_{KH}(\rho, z, a_e) = \frac{2}{\pi}q \times \frac{K\left\{\frac{1}{\sqrt{2}}\left[1 - \frac{\rho^2 + z^2 - a_e^2}{[(\rho^2 + (z + a_e)^2)(\rho^2 + (z - a_e)^2)]^{1/2}}\right]^{1/2}\right\}}{[(\rho^2 + (z + a_e)^2)(\rho^2 + (z - a_e)^2)]^{1/4}},$$

where $K$ is the complete elliptic integral of the first kind. For $a_e \ll 1$,

$$V_{KH} = \frac{1}{r} + O\left(\left(\frac{1}{a_e}\right)^2\right),$$

and the fragments fly apart in the Coulomb field with $\theta_{\rm out} = \theta_0$ (Fig. 1, curve *1*). Below, this mode of ion separation is called Coulomb mode. The ions in the field of the Kramers–Henneberger potential for

$$a_e > 1, \quad R_0/a_e < 1 \qquad (8)$$

fly apart in a direction perpendicular to the field ($\theta_{\rm out} = \pi/2$) at any $\theta_0$. This result is qualitatively explained by Fig. 2. For simplicity, we consider the case with $M_1 \ll M_2$. The field-induced oscillations of H$^+$ take place along straight line *ab*. Clearly, when averaged over the period for $R_0\cos\theta_0 \le a_e$, the $z$ component of the total force acting on H$^+$ is zero, while its $\rho$ component is nonzero. Thus, if the averaging procedure that underlies the Kramers–Henneberger method is valid, the fragments fly apart perpendicular to the direction of field polarization. Below, this mode of ion separation is called the Kramers–Henneberger mode. Note that, given the finite duration of the laser pulse front, the satisfaction of conditions (8) is not enough for the Kramers–Henneberger mode to be realized. The third condition that limits the range of admissible fields and frequencies stems from the fact that when the field intensity and frequency decrease, the duration $\tau$ of the part of the pulse with $a_e < 1$ increases, and the Kramers–Henneberger potential is close to the Coulomb potential. If the ions fly apart to a distance $a_\tau$ larger than $a_e$ in time $\tau$, then the formation of the double-peaked

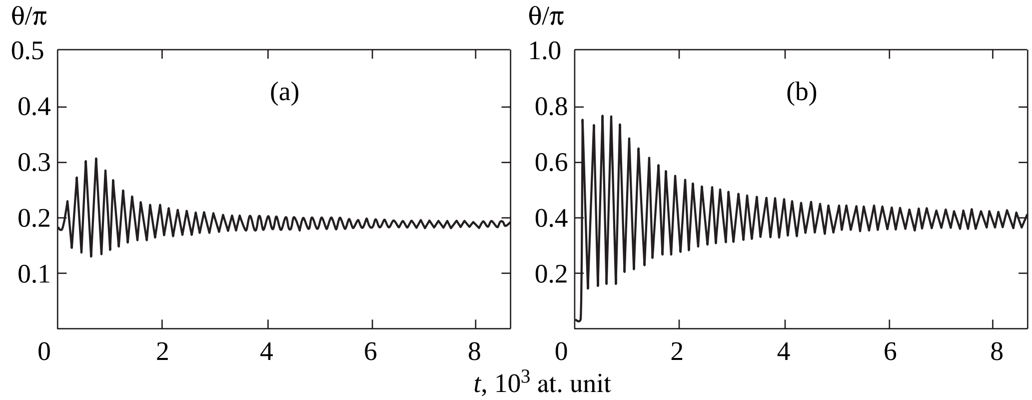

**Fig. 3.** Variations of θ with time: (a) in the Coulomb mode of ion separation, $P = 9 \times 10^{19}$ W cm$^{-2}$, $\omega = 1$ eV; and (b) in the Kramers–Henneberger mode, $P = 3.5 \times 10^{21}$ W cm$^{-2}$, $\omega = 1$ eV.

structure of the Kramers–Henneberger potential on time scales $t > \tau$ will no longer significantly affect the mode of ion separation, which will remain the Coulomb one as before. In contrast to conditions (8), the condition $a_\tau < a_e$ depends on the shape of the pulse envelope. In our case [see Eq. (3)], it is clear that

$$\tau = \frac{20}{\omega}\arcsin a_e^{-1/2} \approx \frac{20}{\omega}a_e^{-1/2}, \quad a_r \approx \frac{1}{2}\frac{\tau^2}{R_0^2\mu}.$$

Thus, the deviations from the Coulomb mode of nuclear separation (curves *2*–*4* in Fig. 1) are attributable to the formation of the double-peaked structure of the Kramers–Henneberger potential. Figure 1 traces the changes in angular distributions when passing from Coulomb (curve *1*) to the Kramers–Henneberger mode of ion separation (curve *4*). The variations of θ with time in the Coulomb and the Kramers–Henneberger modes of ion separation are shown in Fig. 3.

Let us consider the question of whether the averaging procedure, or the range of applicability of the Kramers–Henneberger approximation, is valid for our problem. The applicability of the Kramers–Henneberger approximation to the finite motion of a particle in the field of an attractive potential was considered in [39, 41]. However, these results cannot be extended to the infinite motion of a particle in the field of a repulsive potential. Note that the possibility of considering the dynamics of a particle in the field of a repulsive potential in terms of the Kramers–Henneberger approximation was pointed out in [39]. For the validity of the averaging procedure, it is important that the following two conditions be satisfied in our problem. First, the change in relative coordinate ρ in half the period must be small compared to the internuclear distance,

$$\Delta\rho \approx \frac{1}{2R_0^2\mu}\frac{\pi^2}{\omega^2}.$$

Second, the force exerted on $H^+$ by the field must be larger than the force of Coulomb repulsion; otherwise, $H^+$ oscillations about the force center $A^{q+}$ cannot be provided (see Fig. 3a).

Thus, the domain of field parameters in which the ions fly apart in the Kramers–Henneberger mode is given by

$$a_e > 1, \quad P > \frac{137}{8\pi}\xi^{-2}\left(1 - \frac{q}{m}\right)^{-2}\omega^4, \tag{9}$$

$$\frac{R_0}{a_e} < 1, \quad P > \frac{137}{8\pi}\xi^{-2}\left(1 - \frac{q}{m}\right)^{-2}R_0^2\omega^4, \tag{10}$$

$$\alpha_\tau < a_e, \quad P > 20^2\frac{137}{8\pi}\xi^{-1}\left(1 + \frac{1}{m}\right)\left(1 - \frac{1}{m}\right)^{-2}R_0^{-2}\omega^2, \tag{11}$$

$$\Delta\rho < R_0, \quad \omega > \frac{1}{\sqrt{2}}\left(1 + \frac{1}{m}\right)^{1/2}\pi\xi^{-1/2}R_0^{-3/2}, \tag{12}$$

$$E_0 > \frac{1}{R_0^2}, \quad P > \frac{137}{8\pi}R_0^{-4}. \tag{13}$$

Here, $\xi = 5.44 \times 10^{-4}$ is the electron-to-proton mass ratio and $m = M_2/M_1$. The satisfaction of conditions (9)–(11) is necessary to ensure the double-peaked structure of the Kramers–Henneberger potential during the ion separation; the satisfaction of conditions (12) and (13) is necessary for the averaging procedure to be valid. Figure 4 shows the domain of field parameters in which the Kramers–Henneberger mode is established for the HD molecule. Condition (13) is not reflected in Fig. 4, because it leads to the requirement $P > 1$, which is definitely satisfied in the entire domain. Straight line *5* in Fig. 4 corresponds to the conditions $\nu = c$, $\nu = E_0/\mu\omega$. In the region below this straight line, the Coulomb explosion dynamics can be described in terms of the nonrelativistic model used here. The diamonds mark the domain boundary constructed by simulating the Coulomb explosion of the above molecule. Note that for $M_1 \ll M_2$ (for example, the HCL molecule), the intensities required for the Kramers–Henneberger

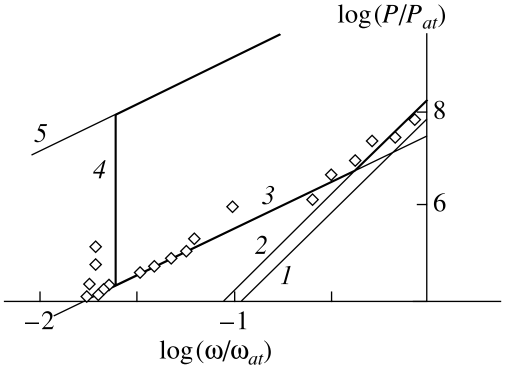

**Fig. 4.** The domain of field parameters (bounded by heavy lines) in which the Kramers–Henneberger ion separation mode is established for the HD molecule. The diamonds mark the domain boundary constructed by computer simulating Coulomb explosion of this molecule. $P_{at}$ and $\omega_{at}$ are the atomic units of field intensity and frequency, respectively. $a_e = 1$ (*1*), $R_0/a_e = 1$ (*2*), $a_\tau = a_e$ (*3*), $\Delta\rho = R_0$ (*4*), and $\nu = c$ (*5*).

mode to be achieved are an order of magnitude lower than those for the HD molecule. This is because the parameter $\gamma$ increases in this case approximately threefold [see Eq. (3)].

## 3. CONCLUSIONS

The angular distributions of the fragments from the Coulomb explosion of a diatomic heteronuclear molecule during dissociative ionization in a superintense laser field have been considered in terms of classical mechanics.

The patterns of angular distributions of the Coulomb explosion fragments have been shown to differ in different ranges of laser-pulse parameters. In particular, there are two distinct modes of ion separation: separation in a Coulomb field ($\theta_{out} = \theta_0$) and separation in the field of an effective ion + field potential ($\theta_{out} = \pi/2$). Based on the Kramers–Henneberger method, we determined the boundaries of these modes. The analytic estimates are in good agreement with computations (see Fig. 4).

Considering the above problem in terms of classical mechanics implies the following: (1) the realization of initial conditions close to the classical ones in a real experiment and (2) the classical dynamics of ions. A linearly polarized, molecule-aligning pulse of the intensity $P \approx 10^{13}$ Wcm$^{-2}$ can apparently be used to realize initial conditions close to the classical ones in a real experiment. The classical treatment of the ion dynamics is possible, because the parameter $\nu_s/\nu_d \approx \xi^{1/4}$, where $\nu_s$ is the spread velocity and $\nu_d$ is the drift velocity of the ion wave packet, is small.

## ACKNOWLEDGMENTS

We are grateful to the Russian Foundation for Basic Research (project nos. 00-02-16046 and 00-15-96554) and INTAS (grant no. 99-1495) for support.

## REFERENCES

1. H. A. Kramers, *Les Particles Elementaires, Report to the Eighth Solvay Conference, Brussels, 1950.*
2. W. C. Henneberger, Phys. Rev. Lett. **21**, 838 (1968).
3. M. Pont, N. R. Walet, M. Gavrila, and C. W. McCrudy, Phys. Rev. Lett. **61**, 939 (1988).
4. M. Pont and M. Gavrila, Phys. Lett. A **123**, 469 (1987).
5. M. Pont, Phys. Rev. A **40**, 5659 (1989).
6. M. Pont, N. R. Walet, and M. Gavrila, Phys. Rev. A **41**, 477 (1990).
7. Q. Su and J. H. Eberly, Phys. Rev. A **43**, 2474 (1991).
8. R. Grobe and J. H. Eberly, Phys. Rev. A **47**, 719 (1993).
9. E. A. Volkova, A. M. Popov, and O. V. Tikhonova, Zh. Éksp. Teor. Fiz. **109**, 1586 (1996) [JETP **82**, 853 (1996)].
10. L. Fransinski, K. Codling, P. Hatherly, *et al.*, Phys. Rev. Lett. **58**, 2424 (1987).
11. K. Boyer, T. S. Luk, J. S. Solem, and S. K. Rhodes, Phys. Rev. A **39**, 1186 (1989).
12. P. A. Hatherly, L. J. Fransinski, K. Codling, *et al.*, J. Phys. B **23**, L291 (1990).
13. D. T. Strickland, Y. Beaudoin, P. Dietrich, and P. B. Corkum, Phys. Rev. Lett. **68**, 2755 (1992).
14. W. T. Hill, J. Zhu, D. L. Hatten, *et al.*, Phys. Rev. Lett. **69**, 2646 (1992).
15. K. Codling and L. J. Fransinski, J. Phys. B **26**, 783 (1993).
16. L. J. Fransinski, J. Phys. B **27**, L109 (1994).
17. P. A. Hatherly, M. Stankiewicz, K. Codling, *et al.*, J. Phys. B **28**, 2993 (1994).
18. C. Cornaggia, M. Schmidt, and D. Normand, J. Phys. B **27**, L123 (1994).
19. J. H. Posthumus, J. Plumridge, M. K. Thomas, *et al.*, J. Phys. B **31**, L553 (1998).
20. M. Schmidt, D. Normand, and C. Cornaggia, Phys. Rev. A **50**, 5037 (1994).
21. T. D. G. Walsh, F. A. Ilkov, S. L. Chin, *et al.*, Phys. Rev. A **58**, 3922 (1998).
22. Ph. Herring and C. Cornaggia, Phys. Rev. A **59**, 2836 (1999).
23. A. Giusti-Suzor and Ch. Jungen, J. Chem. Phys. **80**, 986 (1984).
24. E. Charron and A. Suzor-Weiner, J. Chem. Phys. **108**, 3922 (1998).
25. T. Seideman, M. Yu. Ivanov, and P. B. Corkum, Phys. Rev. Lett. **75**, 2819 (1995).
26. T. Zuo and A. Bandrauk, Phys. Rev. A **51**, R26 (1995).
27. S. Chelcowski and A. Bandrauk, J. Phys. B **28**, L723 (1995).
28. J. H. Posthumus, L. J. Fransinski, A. J. Giles, and K. Codling, J. Phys. B **28**, L349 (1995).
29. T. Yu and A. Bandrauk, Phys. Rev. A **56**, 685 (1997).

30. A. M. Popov, O. V. Tikhonova, and E. A. Volkova, Laser Phys. **7**, 108 (1997).
31. I. Last and J. Jortner, Phys. Rev. A **58**, 3826 (1998).
32. Y. Fyodorov and Y. Alhassid, Phys. Rev. A **58**, R3375 (1998).
33. J. H. Posthumus, J. Plumridge, L. J. Fransinski, *et al.*, J. Phys. B **31**, L985 (1998).
34. J. Grochmalicki, M. Lewenstein, and K. Rzazewski, Phys. Rev. Lett. **66**, 1038 (1991).
35. R. Grobe and C. K. Law, Phys. Rev. A **44**, R4114 (1991).
36. B. Sundaram and R. V. Jensen, Phys. Rev. A **47**, 1415 (1993).
37. F. Benvenuto, G. Casati, and D. L. Shepelyansky, Phys. Rev. A **47**, R786 (1993).
38. G. Casati, I. Guarneri, and G. Mantica, Phys. Rev. A **50**, 5018 (1994).
39. R. V. Karapetyan, Laser Phys. **10**, 160 (2000).
40. M. Gavrila and J. Z. Kaminski, Phys. Rev. Lett. **52**, 613 (1984).
41. O. V. Smirnova, Zh. Éksp. Teor. Fiz. **117**, 702 (2000) [JETP **90**, 609 (2000)].

*Translated by V. Astakhov*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 301–312.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 340–352.*
*Original English Text Copyright © 2001 by Tskhakaya, Shukla, Tsintsadze.*

**PLASMA,
GASES**

# Electrodynamics and Dispersion Properties of a Magnetoplasma Containing Elongated and Rotating Dust Grains[¶]

**D. D. Tskhakaya[a], P. K. Shukla[b, c, *], and N. L. Tsintsadze[a]**
[a]*Institute of Physics, Georgian Academy of Sciences, Tbilisi, 380077 Georgia*
[b]*Institut für Theoretische Physik IV, Fakultät für Physik und Astronomie, Ruhr-Universität Bochum, D-44780, Bochum, Germany*
[c]*Department of Plasma Physics, Umeå University, S-90187 Umeå, Sweden*
**e-mail: ps@tp4.ruhr-uni-bochum.de*

Received November 28, 2000

**Abstract**—The electrodynamics and dispersion properties of a magnetized dusty plasma containing elongated and rotating charged dust grains are examined. Starting from an appropriate Lagrangian for dust grains, a kinetic equation for the dust grain and the corresponding equations of motion are derived. Expressions for the dust charge and dust current densities are obtained with the finite size (the dipole moment) of elongated and rotating dust grains taken into account. These charge and current densities are combined with the Maxwell–Vlasov system of equations to derive dispersion relations for the electromagnetic and electrostatic waves in a dusty magnetoplasma. The dispersion relations are analyzed to demonstrate that the dust grain rotation introduces new classes of instabilities involving various low-frequency waves in a dusty magnetoplasma. Examples of various unstable low-frequency waves include the electron whistler, the dust whistler, dust cyclotron waves, Alfvén waves, electromagnetic ion-cyclotron waves, as well as lower-hybrid, electrostatic ion cyclotron, modified dust ion-acoustic waves, etc. Also found is a new type of unstable waves whose frequency is close to the dust grain rotation frequency. The present results should be useful in understanding the properties of low-frequency waves in cosmic and laboratory plasmas that are embedded in an external magnetic field and contain elongated and rotating charged dust grains. *© 2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

About a decade ago, Shukla and collaborators [1, 2] introduced the idea of considering the dynamics of charged dust grains, which formed the foundation for the dust acoustic waves (DAWs) [1]. In the latter, the restoring force comes from the pressures of the inertialess electrons and ions, while the dust mass provides the inertia to maintain the wave. The phase velocity (the frequency) of DAWs is much smaller than the electron and ion thermal velocities (the dust plasma frequency). On the other hand, when the wave frequency is much higher (lower) than the dust (ion) plasma frequency, we have the dust ion-acoustic waves (DIAWs) [3] whose phase velocity is much lower (higher) than the electron (ion) thermal velocity. In DIAWs, the restoring force comes from the pressure of the inertialess electrons, while the ion mass provides the inertia because the massive dust grains remain immobile at the time scale of the DIAWs. Both the dust acoustic and dust ion-acoustic waves are spectacularly verified in several laboratory experiments [4–8]. We note that the previous theories of DAWs and DIAWs and the corresponding laboratory experiments have dealt with spherical dust grains. Comprehensive reviews of waves and instabilities in a weakly coupled unmagnetized dusty plasma with spherical dust grains were given in [9, 10].

However, elongated charged dust grains are ubiquitous in cosmic and laboratory plasmas [11–14]. The formation of elongated charged dust grains is attributed to the coagulation of particulates in partially or fully ionized gases due to some attractive forces. Elongated charged grains can acquire a rotational motion due to their interaction with photons and particles of the surrounding gas, or due to the presence of an oscillating electric field in a plasma [11, 15]. In astrophysical objects, the angular frequency of the dust grain rotation can reach a rather large value, viz. between tens of kHz to MHz for thermal dust grains and hundreds and thousands of MHz for super thermal grains [11, 12, 16]. There is an orientation of a different kind involving preferred direction (relative to the galactic disk) of the dust grain angular momentum vector.

In general, elongated charged dust grains have a nonzero dipole moment due to a finite grain size. Accordingly, Mahmoodi *et al.* [17] investigated the dispersion properties of an unmagnetized dusty plasma in the presence of rotating and elongated dust grains. It was found that the dust rotational energy can be coupled to both the electromagnetic and electrostatic waves. However, cosmic and laboratory plasmas are usually embedded in an external magnetic field that can

---

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

have substantial effects on the dusty plasma wave spectra when elongated and rotating dust grains are present in a dusty plasma system.

In this paper, we present the electrodynamics and dispersion properties of a dusty magnetoplasma whose constituents are electrons, ions, and finite-sized elongated dust grains. In Section 2, we find expressions for the charge and current densities of dust grains by including the effect of the dust dipole moment and the dust grain rotation. The forces acting on the dust grains as well as the corresponding dust kinetic equation and the equations of motion are presented in Section 3. In Section 4, we derive dispersion relations for both the electromagnetic and electrostatic waves. Specific instability results are discussed in Section 5. Finally, Section 6 contains a brief summary and possible applications of our work to cosmic and laboratory plasmas.

## 2. DERIVATION OF THE CHARGE AND CURRENT DENSITIES FOR DUST GRAINS

We consider a multicomponent dusty plasma in the external magnetic field $\hat{\mathbf{z}}B_0$, where $\hat{\mathbf{z}}$ is the unit vector along the $z$ axis and $B_0$ is the strength of the external magnetic field. The dusty plasma constituents are electrons, ions, and negatively charged nonspherical rotating dust grains. The dust sizes are much smaller than the characteristic scale sizes of the inhomogeneities (wavelength of disturbances in our system). To construct the electrodynamics of charged dust grains in a magnetized dusty plasma, we must obtain appropriate expressions for the charge and current densities of dust grains through the dust grain distribution function, taking the size of the dust grain into account. On the other hand, expressions for the charge and current densities of electrons and ions assume the standard form.

For our purposes, we assume that the charged dust grains are a system of discrete parts [18]. The charge microdensity of the grains is represented as

$$\rho_m = \sum_i \left[\sum_j dq(\mathbf{r}_j)\delta(\mathbf{r} - \mathbf{r}_j)\right], \tag{1}$$

where the summation over $i$ is taken over different grains and the one over $j$ is taken over different parts of the $i$th grain. Here, $dq_i(\mathbf{r}_j)$ is the charge of the $j$th part of the $i$th grain and $\delta(\mathbf{r} - \mathbf{r}_j)$ is the standard Dirac function. If there is a continuous charge distribution onto the grain, the summation over $j$ can be replaced with the integral over the grain volume, and the charge density on the grain can be introduced. Hence, we have

$$\rho_m = \sum_i \int\limits_{V_i(\mathbf{R}_i)} \bar{\rho}_i(\mathbf{r}' - \mathbf{R}_i, \mathbf{R}_i)\delta(\mathbf{r} - \mathbf{r}'), \tag{2}$$

where $\mathbf{R}_i$ is the radius vector of the center of mass of the grain and the integral is taken over the grain volume $V_i(\mathbf{R}_i)$. In (2), we introduced the density of the charge distribution onto the grain

$$dq_i(\mathbf{r}) = \frac{dq(\mathbf{r})}{d\mathbf{r}}d\mathbf{r} \equiv \bar{\rho}_i(\mathbf{r} - \mathbf{R}_i, \mathbf{R}_i)d\mathbf{r}. \tag{3}$$

For a point grain charge, we have

$$\bar{\rho}_i(\mathbf{r} - \mathbf{R}_i, \mathbf{R}_i) = q_i\delta(\mathbf{r} - \mathbf{R}_i), \tag{4}$$

which leads to the usual expression for the charge microdensity of the grain

$$\rho_m = \sum_i q_i\delta(\mathbf{r} - \mathbf{R}_i), \tag{5}$$

where $q_i$ is the total charge of the $i$th grain.

For the statistical description of a dust grain gas, we must introduce the probability density $D$ for the grain gas state [19, 20]. If all grains are identical, we have

$$\begin{gathered} D = D(\mathbf{R}_1, \mathbf{v}_1, \mathbf{\Omega}_1, \theta_1, \psi_1, \varphi_1; \\ \mathbf{R}_2, \mathbf{v}_2, \mathbf{\Omega}_2, \theta_2, \psi_2, \varphi_2; \ldots, t), \end{gathered} \tag{6}$$

where $\mathbf{v}_i$ is the velocity of the center of mass, $\mathbf{\Omega}_i$ is the angular velocity of the $i$th grain, and $\theta_i$, $\psi_i$, and $\varphi_i$ (the Euler angles) describe the orientation of elongated grains. For the averaged charge density of the grain, we can then write

$$\rho(\mathbf{r}, t) = \int d\Gamma_1, d\Gamma_2, \ldots, d\Gamma_N D\rho_m, \tag{7}$$

where $N$ is the total number of grains and

$$d\Gamma_i = d\mathbf{R}_i d\mathbf{v}_i d\mathbf{\Omega}_i d\theta_i d\psi_i d\varphi_i.$$

Introducing the one-particle distribution function for the dust grain

$$\begin{aligned} & f_d(\mathbf{R}_1, \mathbf{v}_1, \mathbf{\Omega}_1, \theta_1, \psi_1, \varphi_1) \\ & \quad = N\int d\Gamma_2 d\Gamma_3, \ldots, d\Gamma_N D, \end{aligned} \tag{8}$$

we can write the charge density of the grains as

$$\begin{aligned} \rho(\mathbf{r}, t) &= \int d\Gamma_1 \int\limits_{V_1} \bar{\rho}_1(\mathbf{r}'')\delta(\mathbf{r} - \mathbf{R}_1 - \mathbf{r}'') \\ &\times f_d(\mathbf{R}_1, \mathbf{v}_1, \mathbf{\Omega}_1, \theta_1, \psi_1, \varphi_1, t)d\mathbf{r}''. \end{aligned} \tag{9}$$

In what follows, we omit the subscript 1 and consider the one-dimensional grain rotation such that the angular velocity is oriented along the external magnetic field direction, $\mathbf{\Omega} = (0, 0, \Omega)$. Equation (9) can then be written as

$$\rho(\mathbf{r}, t) = \int d\Gamma\hat{\rho}(\mathbf{r} - \mathbf{R}, \varphi)f_d(\mathbf{R}, \mathbf{v}, \Omega, \varphi, t), \tag{10}$$

where the integrand

$$\hat{\rho}(\mathbf{r} - \mathbf{R}, \varphi) = \int\limits_V d\mathbf{r}'\bar{\rho}(\mathbf{r}')\delta(\mathbf{r} - \mathbf{R} - \mathbf{r}'), \tag{11}$$

describing the charge distribution onto a single grain, depends on the shape of the grain and the azimuthal orientation of the grain elongation axis. Outside the grain volume, we have $\hat{\rho} = 0$. For identical grains, we can partly determine the dependence of $\hat{\rho}$ on the azimuthal angle $\varphi$. Every given direction of the grain elongation axis, determined by the angle $\varphi$, can be considered as the final position of the axis (and simultaneously the entire grain) rotation from the direction where $\varphi = 0$. This allows us to write

$$\begin{aligned}\hat{\rho}(\mathbf{r}-\mathbf{R}, \varphi) &= \hat{\rho}[\overset{\leftrightarrow}{\mathbf{F}}(\varphi)(\mathbf{r}-\mathbf{R}), 0)] \\ &\equiv \hat{\rho}[\overset{\leftrightarrow}{\mathbf{F}}(\varphi)(\mathbf{r}-\mathbf{R})],\end{aligned} \tag{12}$$

where $\overset{\leftrightarrow}{\mathbf{F}}(\varphi)$ is the rotation matrix for the angle $\varphi$,

$$\overset{\leftrightarrow}{\mathbf{F}} = F_{ij}(j) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{13}$$

In the dipole approximation, when the dust grain size $a$ is much smaller than the scale length of the plasma inhomogeneity $\lambda$,

$$a \ll \lambda, \tag{14}$$

we insert (12) in (10) and expand the distribution function $f_d$ around the point $\mathbf{r}$. This gives the grain charge density

$$\rho_d(\mathbf{r}, t) = \int (q - \mathbf{d}\cdot\nabla) f_d(\mathbf{r}, \mathbf{v}, \Omega, \varphi, t) d\Lambda, \tag{15}$$

where $d\Lambda = d\mathbf{v}\, d\Omega\, d\varphi$,

$$q = \int d\mathbf{r}\hat{\rho}(\mathbf{r}) \tag{16}$$

is the total charge of the dust grain, and

$$\mathbf{d} = \overset{\leftrightarrow}{\mathbf{F}}{}^{-1}(\varphi)\int d\mathbf{r}\,\mathbf{r}\hat{\rho}(\mathbf{r}) \tag{17}$$

is the dipole moment of the grain. Here, $\overset{\leftrightarrow}{\mathbf{F}}{}^{-1}$ is the inverse matrix of $\overset{\leftrightarrow}{\mathbf{F}}(\varphi)$.

Similar calculations lead to the following expression for the dust current density:

$$\begin{aligned}\mathbf{J}_d(\mathbf{r}, t) &= \int d\Lambda[\mathbf{v}(q - \mathbf{d}\cdot\nabla) + \mathbf{\Omega}\times\mathbf{d}] \\ &\times f_d(\mathbf{r}, \mathbf{v}, \Omega, \varphi, t).\end{aligned} \tag{18}$$

The first term in the right-hand side of (18) describes the transfer of charge (15) and the second term describes the current arising from the dust grain rotation. In the next section, we show that Eq. (15) and (18) are related to the continuity equation.

## 3. FORCES ACTING ON GRAINS AND THE GRAIN KINETIC EQUATION

To construct the kinetic equation for dust grains, we must completely know the forces that act on dust grains in the presence of electromagnetic fields. Assuming that charged dust grains constitute a discrete system of particles [18], we have the Lagrangian

$$\begin{aligned}\mathcal{L} &= \sum_i \frac{\Delta m_i u_i^2}{2} \\ &+ \frac{1}{c}\sum_i \Delta q_i[\mathbf{v}_i\cdot\mathbf{A}(\mathbf{r}, t)] - \sum_i \Delta q_i \phi(\mathbf{r}, t),\end{aligned} \tag{19}$$

where $\Delta m_i$ and $\Delta q_i$ are the mass and the charge of the $i$th part of the grain, respectively, $\mathbf{r}_i$ and $\mathbf{u}_i$ are its coordinate and velocity, $\mathbf{A}$ and $\phi$ are the vector and scalar potentials, respectively, and $c$ is the speed of light in vacuum. Separating the center-of-mass motion and the rotation around the center of mass, we can write

$$\mathbf{u}_i = \mathbf{v} + \mathbf{\Omega}\times\Delta\mathbf{r}_i \quad \text{and} \quad \mathbf{r}_i = \mathbf{r} + \Delta\mathbf{r}_i,$$

where $\mathbf{v}$ and $\mathbf{r}$ are the velocity and the position of the center of mass, $\Delta\mathbf{r}_i$ is the coordinate of the $i$th part of the grain relative to the center of mass, and $\mathbf{\Omega}$ is the angular velocity of the dust grain. Assuming that the inhomogeneity scale $\lambda$ of the electromagnetic field is much larger than the grain size $a$, we can use dipole approximation (14) up to the third order in the small parameter $a/\lambda$ and expand the potentials as

$$\begin{aligned}\mathbf{A}(\mathbf{r}_i, t) &= \mathbf{A}(\mathbf{r}, t) + (\Delta\mathbf{r}_i\cdot\nabla)\mathbf{A}(\mathbf{r}, t) \\ &+ \frac{1}{2}(\Delta\mathbf{r}_i\cdot\nabla)^2\mathbf{A}(\mathbf{r}, t) + \ldots,\end{aligned} \tag{20}$$

and

$$\begin{aligned}\phi(\mathbf{r}_i, t) &= \phi(\mathbf{r}, t) + (\Delta\mathbf{r}_i\cdot\nabla)\phi(\mathbf{r}, t) \\ &+ \frac{1}{2}(\Delta\mathbf{r}_i\cdot\nabla)^2\phi(\mathbf{r}, t) + \ldots .\end{aligned} \tag{21}$$

Accordingly, Lagrangian (19) becomes

$$\begin{aligned}\mathcal{L} &= \frac{m_d\mathbf{v}^2}{2} + \frac{1}{2}I_{\alpha\beta}\Omega_\alpha\Omega_\beta + \frac{q}{c}\mathbf{v}\cdot\mathbf{A}(\mathbf{r}, t) - q\phi(\mathbf{r}, t) \\ &+ \mathbf{m}\cdot\mathbf{B} + \left[\mathbf{d} + \frac{1}{2}\sum_i \Delta q_i\Delta\mathbf{r}_i(\Delta\mathbf{r}_i\cdot\nabla)\right]\left(\mathbf{E} + \frac{\mathbf{v}}{c}\times\mathbf{B}\right),\end{aligned} \tag{22}$$

where

$$m_d = \sum_i \Delta m_i, \quad q = \sum_i \Delta q_i$$

are the total mass and charge of the grain,

$$I_{\alpha,\beta} = \sum_i \Delta m_i[(\Delta\mathbf{r}_i)^2\delta_{\alpha\beta} - (\Delta\mathbf{r}_i)_\alpha\cdot(\Delta\mathbf{r}_i)_\beta]$$

is the inertia moment tensor,

$$\mathbf{d} = \sum_i \Delta q_i \Delta\mathbf{r}_i$$

is the dipole moment of the elongated grain, and

$$\mathbf{m} = (1/2c)\sum_i \Delta q_i(\Delta\mathbf{r}\times\mathbf{U}_i)$$

(with $\mathbf{U}_i = \mathbf{\Omega}\times\Delta\mathbf{r}_i$ being the rotation velocity) is the magnetic moment of the grain. The electric and magnetic fields are

$$\mathbf{E} = -\nabla\phi - c^{-1}\partial_t\mathbf{A}(\mathbf{r}, t), \quad \mathbf{B} = \nabla\times\mathbf{A}(\mathbf{r}, t),$$

respectively. In deriving (22), we used the relation

$$d\mathbf{d}/dt = \mathbf{\Omega}\times\mathbf{d}.$$

In the presence of the gravity field $\mathbf{g}$, we must add the term $m_d\mathbf{g}\cdot\mathbf{r}$ to the right-hand side of (22). In what follows, we neglect the second term in the square bracket in the right-hand side of (22), which is associated with the multidipole effect.

The equations of motion for the charged dust grains can be readily deduced from (22) as

$$\begin{gathered}\frac{d\mathbf{p}}{dt} = (q + \mathbf{d}\cdot\nabla)\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right) \\ + \frac{1}{c}(\mathbf{\Omega}\times\mathbf{d})\times\mathbf{B} + (\mathbf{m}\times\nabla)\times\mathbf{B}\end{gathered} \tag{23}$$

and

$$\begin{gathered}\frac{dM_\alpha}{dt} = -\frac{1}{2}S_{\alpha\beta}\left[\frac{\partial B_\beta}{\partial t} + (\mathbf{v}\cdot\nabla)B_\beta\right] \\ + \left[\mathbf{d}\times\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right)\right]_\alpha + (\mathbf{m}\times\mathbf{B})_\alpha,\end{gathered} \tag{24}$$

where $\mathbf{p} = m_d\mathbf{v}$ is the momentum,

$$S_{\alpha\beta} = c^{-1}\sum_i \Delta q_i[(\Delta\mathbf{r}_i)^2\delta_{\alpha\beta} - (\Delta\mathbf{r}_i)_\alpha(\Delta\mathbf{r}_i)_\beta],$$

and $M_\alpha = I_{\alpha\beta}\Omega_\beta$ is the angular momentum of the grain. If we choose the principal axis of the moment of inertia, then

$$M_x = I_x\Omega_x, \quad M_y = I_y\Omega_y, \quad M_z = I_z\Omega_z.$$

The kinetic equation for the dust grains can now be written as

$$\begin{gathered}\frac{\partial f_d}{\partial t} + \mathbf{v}\cdot\frac{\partial f_d}{\partial\mathbf{r}} + \mathbf{\Omega}\cdot\frac{\partial f_d}{\partial\varphi} \\ + \frac{d\mathbf{p}}{dt}\cdot\frac{\partial f_d}{\partial\mathbf{p}} + \frac{d\mathbf{M}}{dt}\cdot\frac{\partial f_d}{\partial\mathbf{M}} = 0,\end{gathered} \tag{25}$$

where the respective forces $d\mathbf{p}/dt$ and $d\mathbf{M}/dt$ are defined by Eqs. (23) and (24). Kinetic equation (25) and definitions (15) and (18) imply that the dust grain charge and the current densities satisfy the continuity equation

$$\frac{\partial\rho_d}{\partial t} + \frac{\partial\mathbf{J}_d}{\partial\mathbf{r}} = 0. \tag{26}$$

Using the expressions for $\rho_d$ and $\mathbf{J}_d$, we can construct the kinetics and electrodynamics of a dusty plasma with elongated and rotating dust grains. In what follows, we consider the wave dynamics of such a magnetized dusty plasma.

## 4. DIELECTRIC PERMITTIVITY

We assume that the dust grain size is much smaller than the grain gyroradius and that the dust grain thermal velocity is smaller than the characteristic velocity of our problem. Under these conditions, taken together with (14), equations of motion (23) and (24) can be simplified. For simplicity, we furthermore consider the one-dimensional case of the dust grain rotation; we then have $\mathbf{M} = (0, 0, M)$, where $M = I\Omega$ and $I$ is the $z$ component of the principal moment of inertia. The kinetic equation for the dust grain (25) then assumes the form

$$\begin{gathered}\frac{\partial f_d}{\partial t} + \mathbf{v}\cdot\frac{\partial f_d}{\partial\mathbf{r}} + \Omega\frac{\partial f_d}{\partial\varphi} \\ + (\mathbf{d}\times\mathbf{E})_z\frac{\partial f_d}{\partial M} + q\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}_0\right)\cdot\frac{\partial f_d}{\partial\mathbf{p}} = 0.\end{gathered} \tag{27}$$

For electrons and ions, we have the well-known kinetic equation

$$\begin{gathered}\frac{\partial f_\alpha}{\partial t} + \mathbf{v}\cdot\frac{\partial f_\alpha}{\partial\mathbf{r}} \\ + e_\alpha\left[\mathbf{E} + \frac{1}{c}\mathbf{v}\times(\mathbf{B}_0 + \mathbf{B})\right]\cdot\frac{\partial f_\alpha}{\partial\mathbf{p}} = 0,\end{gathered} \tag{28}$$

where $\alpha$ equals $e$ for electrons and $i$ for ions, and $e_\alpha$ is the charge of the species $\alpha$.

Assuming that the wave electric and magnetic field perturbations are small, we can express the perturbed distribution function as

$$\delta f_d = f_d - f_{d0} \ll f_{d0} \quad \text{and} \quad \delta f_\alpha = f_\alpha - f_{\alpha 0} \ll f_\alpha.$$

The equilibrium distribution functions are [21]

$$\begin{gathered}f_{d0} = \frac{n_{d0}}{2\pi(2\pi m_d T_d)^{3/2}}\frac{1}{(2\pi I T_d)^{1/2}} \\ \times\exp\left[-\frac{p^2}{2m_d T_d} - \frac{(M - M_0)^2}{2IT_d}\right],\end{gathered} \tag{29}$$

and

$$f_{\alpha 0} = \frac{n_{\alpha 0}}{(2\pi m_\alpha T_\alpha)^{3/2}}\exp\left(-\frac{p^2}{2m_\alpha T_\alpha}\right), \qquad (30)$$

where $n_{\beta 0}$ and $T_\beta$ ($\beta = e, i, d$) are the unperturbed number density and the temperature of the species $\beta$. We assumed that the dust grains rotate with a preferred angular velocity $\Omega_0$, and therefore, $M_0 = I\Omega_0$.

The components of the dust dipole moment are

$$d_x = d\cos\varphi, \quad d_y = d\sin\varphi. \qquad (31)$$

Thus, the perturbed dust grain distribution function is represented as

$$\delta f_d = \sum_{n=-\infty}^{\infty}\delta f_n \exp(in\varphi), \qquad (32)$$

and therefore, Eqs. (27) and (28) give [22]

$$\begin{aligned}&\frac{\partial\delta f_n}{\partial t} + \mathbf{v}\cdot\frac{\partial\delta f_n}{\partial\mathbf{r}} + in\Omega\delta f_n - \omega_{cd}\frac{\partial\delta f_n}{\partial\psi}\\ &\quad = -q\mathbf{E}\cdot\frac{\partial f_{d0}}{\partial\mathbf{p}}\Delta(n) - \frac{i}{2}\frac{\partial f_{d0}}{\partial M}\\ &\times d[(E_x - iE_y)\Delta(n-1) + (E_x + iE_y)\Delta(n+1)]\end{aligned} \qquad (33)$$

and

$$\frac{\partial\delta f_\alpha}{\partial t} + \mathbf{v}\cdot\frac{\partial\delta f_\alpha}{\partial\mathbf{r}} - \omega_{c\alpha}\frac{\partial\delta f_\alpha}{\partial\psi} = -e_\alpha\mathbf{E}\cdot\frac{\partial f_{\alpha 0}}{\partial\mathbf{p}}, \qquad (34)$$

where

$$\omega_{cd} = qB_0/m_d c, \quad \omega_{c\alpha} = e_\alpha B_0/m_\alpha c$$

are the cyclotron frequencies of the dust grain and the species $\alpha$, respectively. Furthermore, $\Delta(n)$ equals 1 for $n = 0$ and 0 for $n \neq 0$. The symbol $\psi$is the azimuthal angle in the momentum space [22],

$$p_x = p_\perp\cos\psi, \quad p_y = p_\perp\sin\psi.$$

In accordance with (31), only $n = 0, \pm 1$ give a contribution to the summation in (32).

Assuming that the perturbed quantities are proportional to $\exp(-i\omega t + i\mathbf{k}\cdot\mathbf{r})$, where $\omega$ and $\mathbf{k}$ are the frequency and the wave vector, respectively, we obtain [22] the following solutions of Eqs. (33) and (34):

$$\begin{aligned}\delta f_0 &= \frac{q\mathbf{E}}{\omega_{cd}}\int_{\pm\infty}^{\psi} d\psi'\frac{\partial f_{d0}}{\partial\mathbf{p}}\\ &\times\exp\left\{-i\int_{\psi'}^{\psi}\left[\frac{\omega - (\mathbf{k}\cdot\mathbf{v}(\psi''))}{\omega_{cd}}\right]d\psi''\right\},\end{aligned} \qquad (35)$$

$$\begin{aligned}\delta f_{\pm 1} &= \pm\frac{i}{2}\frac{d(E_x \mp iE_y)}{\omega_{cd}}\int_{\pm\infty}^{\psi} d\psi'\frac{\partial f_{d0}}{\partial M}\\ &\times\exp\left\{-\int_{\psi'}^{\psi}\left[\frac{\omega \mp \Omega - \mathbf{k}\cdot\mathbf{v}(\psi'')}{\omega_{cd}}\right]d\psi''\right\},\end{aligned} \qquad (36)$$

$$\begin{aligned}\delta f_\alpha &= \frac{e_\alpha\mathbf{E}}{\omega_{c\alpha}}\int_{\pm\infty}^{\psi} d\psi'\frac{\partial f_{\alpha 0}}{\partial\mathbf{p}}\\ &\times\exp\left\{-i\int_{\psi'}^{\psi}\left[\frac{\omega - (\mathbf{k}\cdot\mathbf{v}(\psi''))}{\omega_{c\alpha}}\right]d\psi''\right\}.\end{aligned} \qquad (37)$$

Inserting Eqs. (35) and (36) in (18) and also inserting (37) in the expression for the electron and ion current densities

$$\mathbf{J}_\alpha = e_\alpha\int d\mathbf{p}\mathbf{v}f_\alpha, \qquad (38)$$

we obtain the total current density

$$J_i = \left[\sigma^r_{ij}(\omega, \mathbf{k}) + \sum_{\beta = e, i, d}\sigma^\beta_{ij}(\omega, \mathbf{k})\right]E_j, \qquad (39)$$

where the first term in the right-hand side is related to the rotational motion of the dust grain and the second term represents the contributions of the electrons and ions including the center of mass motion of the grains. The various components [25] of $\sigma^r_{ij}$ and the dielectric permittivity are given in the Appendix. For $k_\perp^2 V_{td}^2 \ll \omega_{cd}^2$ and $|\omega \pm \Omega_0| \gg KV_{dt}$, the dust grains are assumed to be cold and the rotational part of the dielectric tensor (cf. Eq. (A.15) in the Appendix) is given by

$$\epsilon^r_{ij} = \begin{pmatrix}\epsilon^r_\perp & ig^r & 0\\ -ig^r & \epsilon^r_\perp & 0\\ 0 & 0 & 0\end{pmatrix}, \qquad (40)$$

where

$$\epsilon^r_\perp = -\frac{\Omega_r^2}{(\omega - \Omega_0)^2} - \frac{\Omega_r^2}{(\omega + \Omega_0)^2} \qquad (41)$$

and

$$g^r = \frac{\Omega_r^2}{(\omega - \Omega_0)^2} - \frac{\Omega_r^2}{(\omega + \Omega_0)^2}. \qquad (42)$$

We note that this involves a new characteristic frequency

$$\Omega_r = (4\pi n_{d0}d^2/4I)^{1/2}$$

for dust grains that have a nonzero dipole moment. This frequency is of the same order as the dust plasma frequency $\omega_{pd}$.

## 5. DISPERSION PROPERTIES

The general analysis of the dispersion relation

$$\left|k^2\delta_{ij} - k_i k_j - \frac{\omega^2}{c^2}\epsilon_{ij}(\omega, \mathbf{k})\right| = 0 \tag{43}$$

for waves in a magnetized dusty plasma is rather complicated, because the number of wave branches is large. Here, we present the dispersion properties of some most interesting modes and describe the underlying approximations required for the existence of these modes. We first consider waves that are propagating along $\hat{\mathbf{z}}B_0$. For waves in a cold dusty plasma with

$$\begin{gathered} k_\perp V_{t\alpha} \ll \omega_{c\alpha}, \quad |k_z|V_{t\alpha} \ll \omega \\ \text{and} \quad |\omega \pm n\omega_{c\alpha}| \gg |k_z|V_{t\alpha}, \end{gathered} \tag{44}$$

we have

$$\begin{aligned} \epsilon_{xx} &= \epsilon_{yy} = \epsilon_\perp \\ &= 1 - \sum_\beta \frac{\omega_{p\beta}^2}{\omega^2 - \omega_{c\beta}^2} - \frac{\Omega_r^2}{(\omega - \Omega_0)^2} - \frac{\Omega_r^2}{(\omega + \Omega_0)^2}, \end{aligned} \tag{45}$$

$$\begin{aligned} \epsilon_{xy} &= -\epsilon_{yx} = ig \\ &= -i\sum_\beta \frac{\omega_{p\beta}^2\omega_{c\beta}}{\omega(\omega^2 - \omega_{c\beta}^2)} + i\frac{\Omega_r^2}{(\omega - \Omega_0)^2} - i\frac{\Omega_r^2}{(\omega + \Omega_0)^2}, \end{aligned} \tag{46}$$

$$\epsilon_{zz} = \epsilon_\parallel = 1 - \sum_\beta \frac{\omega_{p\beta}^2}{\omega^2}, \tag{47}$$

and

$$\epsilon_{xz} = \epsilon_{zx} = \epsilon_{yz} = \epsilon_{zy} = 0. \tag{48}$$

The electric field components are determined by the set of equations

$$\left(k_z^2 - \frac{\omega^2}{c^2}\epsilon_\perp\right)E_x - i\frac{\omega^2}{c^2}gE_y - k_\perp k_z E_z = 0, \tag{49}$$

$$\frac{\omega^2}{c^2}gE_x + \left(k^2 - \frac{\omega^2}{c^2}\epsilon_\perp\right)E_y = 0, \tag{50}$$

$$-k_\perp k_z E_x + \left(k_\perp^2 - \frac{\omega^2}{c^2}\epsilon_\parallel\right)E_z = 0. \tag{51}$$

We note that for $k_\perp = 0$ (i.e., for $\mathbf{k} = \hat{\mathbf{z}}k_z$), we have $\epsilon_\parallel = 0$ if $E_z \neq 0$, which shows that the dust grain rotation does not affect the longitudinal waves. Obviously, the dust grain rotation can act on the waves when the electric field is in the rotation plane. The energy exchange between the dust grain rotation and such a wave is most efficient when the rotation frequency is close to the wave frequency.

For the circularly polarized electromagnetic waves, we have

$$\frac{k^2c^2}{\omega^2} = 1 - \sum_\beta \frac{\omega_{p\beta}^2}{\omega(\omega \mp \omega_{c\beta})} - \frac{2\Omega_r^2}{(\omega \pm \Omega_0)^2}, \tag{52}$$

where ± in the denominators corresponds to the left/right-hand circularly polarized waves. By replacing $\Omega_0$ with $-\Omega_0$, we can make the dust grain rotation direction coincide with the wave polarization direction.

Dispersion relation (52) can be written as

$$\frac{k^2c^2}{\omega^2} = \epsilon(\omega) - \frac{2\Omega_r^2}{(\omega - \Omega_0)^2}, \tag{53}$$

where

$$\epsilon(\omega) = 1 - \sum_\beta \frac{\omega_{p\beta}^2}{\omega(\omega + \omega_{c\beta})}. \tag{54}$$

Introducing a small frequency shift $\Delta$ around $\Omega_0$, we set $\omega = \Omega_0 + \Delta$, where $\Delta \ll \Omega_0$, and express (53) as

$$\frac{k^2c^2}{\Omega_0^2} - \epsilon(\Omega_0) + \Delta\frac{\partial}{\partial\Omega_0}\left[\frac{k^2c^2}{\Omega_0^2} - \epsilon(\Omega_0)\right] = -\frac{2\Omega_r^2}{\Delta^2}. \tag{55}$$

We now assume that $\Omega_0$ is far from the characteristic frequency $\omega_0$ of the magnetized dusty plasma, which satisfies

$$H(\omega_0) = \frac{k^2c^2}{\omega^2} - \epsilon(\omega_0) = 0. \tag{56}$$

The condition

$$\left|\frac{H(\Omega_0)}{\Omega_0(dH(\Omega_0)/d\Omega_0)}\right| \gg \frac{\Delta}{\Omega_0}, \tag{57}$$

is then satisfied (this case is referred to as the nonresonance case) and we obtain

$$\Delta = \pm i\sqrt{2}\frac{\Omega_r}{kc}\Omega_0\left[1 + \frac{\Omega_0^2}{k^2c^2}\epsilon(\Omega_0)\right], \tag{58}$$

where we also assumed that $\Omega_0^2 \ll k^2c^2$. Equation (58) describes a new type of unstable transversal waves whose frequency is close to the rotation frequency $\Omega_0$. In the resonance case, when inequality (57) is reversed, $\Omega_0$ is close to some characteristic frequency of the magnetized dusty plasma,

$$H(\Omega_0) = 0, \tag{59}$$

and we obtain the frequency shift

$$\Delta = \left[-\frac{2\Omega_r^2}{\Omega_0^3 \partial H(\Omega_0)/\partial\Omega_0}\right]^{1/3} \Omega_0 \left(\frac{1}{-\frac{1 \pm i\sqrt{3}}{2}}\right). \qquad (60)$$

Equation (60) exhibits an unstable root with a substantial growth rate that is proportional to $\Omega_r^{2/3}$. This was expected because dispersion relation (53) is formally similar to the dispersion relation for a two-stream instability discussed in [23].

We now present several examples of the magnetized dusty plasma wave spectra for the resonance case. Because $\Omega_0$ is small in most of the astrophysical and terrestrial environments, we consider low-frequency regimes of the plasma oscillations.

For $|\omega_{cd}|, \omega_{ci} \ll \omega \ll |\omega_{ce}|$, we have

$$H(\omega) = \frac{k^2c^2}{\omega^2} - \frac{\omega_{pe}^2}{\omega|\omega_{ce}|}. \qquad (61)$$

Setting

$$\omega = \Omega_0 + i\gamma,$$

where

$$\Omega_0 \approx \omega_0 = k^2c^2|\omega_{ce}|/\omega_{pe}^2$$

(the electron whistler waves), we obtain the growth rate

$$\gamma \approx \Omega_0\left[2\frac{\Omega_r^2}{k^2c^2}\right]^{1/3}. \qquad (62)$$

In the frequency regime where $|\omega_{cd}| \ll \omega \ll \omega_{ci}$, we have

$$H(\omega) = \frac{k^2c^2}{\omega^2} - \frac{\omega_{pd}^2}{\omega|\omega_{cd}|}. \qquad (63)$$

In deriving (63), we used the dusty plasma quasi-neutrality condition at equilibrium

$$|e|n_{e0} + |q|n_{d0} = e_i n_{i0}. \qquad (64)$$

Setting

$$\omega = \Omega_0 + i\gamma,$$

where

$$\Omega_0 \approx \omega_0 = k^2c^2|\omega_{cd}|/\omega_{pd}^2$$

(the dust whistler wave [2, 24–26]), we obtain the growth rate

$$\gamma \approx \left(2\frac{\Omega_r^2}{k^2c^2}\right)^{1/3}\Omega_0. \qquad (65)$$

On the other hand, for $\omega \sim \omega_{cd}$, we have

$$H(\omega) = \frac{k^2c^2}{\omega^2} - \frac{\omega_{pd}^2}{\omega(|\omega_{cd}| - \omega)}. \qquad (66)$$

In this case, setting

$$\omega = \Omega_0 + i\gamma,$$

where

$$\Omega_0 = \omega_0 = |\omega_{cd}|(1 - \omega_{pd}^2/k^2c^2)$$

(the electromagnetic dust cyclotron wave), we obtain the growth rate

$$\gamma \approx \left(2\frac{\Omega_r^2}{k^2c^2}\frac{\omega_{pd}^2}{k^2c^2}\right)^{1/3}|\omega_{cd}|. \qquad (67)$$

For the frequency range $\omega \approx \omega_{ci}$ (ion cyclotron waves), the growth rate is given by

$$\gamma \approx \left(2\frac{\Omega_r^2}{k^2c^2}\frac{\omega_{pi}^2}{k^2c^2}\right)^{1/3}\omega_{ci}. \qquad (68)$$

We now take the thermal motion of the electrons into account assuming that

$$V_{td}, V_{ti} \ll \frac{\omega}{|k_z|} \ll V_{te}. \qquad (69)$$

We restrict ourself to the wavelengths longer than the Larmor radii

$$k_\perp^2 V_{t\alpha}^2, \quad |k_z|^2 V_{t\alpha}^2 \ll \omega_{c\alpha}^2.$$

From (A.9)–(A.11), we then obtain the dielectric permittivity tensor components

$$\begin{aligned}\epsilon_{xx} = \epsilon_{yy} &= 1 + \frac{c^2}{V_A^2} - \frac{\omega_{pd}^2}{\omega^2 - \omega_{cd}^2} \\ &- \frac{\Omega_r^2}{(\omega - \Omega_0)^2} - \frac{\Omega_r^2}{(\omega + \Omega_0)^2},\end{aligned} \qquad (70)$$

$$\begin{aligned}\epsilon_{xy} = -\epsilon_{yx} &= -i\frac{\omega_{pd}^2}{\omega^2 - \omega_{cd}^2}\frac{\omega}{\omega_{cd}} \\ &+ i\frac{\Omega_r^2}{(\omega - \Omega_0)^2} - i\frac{\Omega_r^2}{(\omega + \Omega_0)^2},\end{aligned} \qquad (71)$$

$$\epsilon_{zz} = 1 - \frac{\omega_{pi}^2}{\omega^2} - \frac{\omega_{pd}^2}{\omega^2} + \frac{\omega_{pe}^2}{k_z^2 V_{te}^2}, \qquad (72)$$

$$\epsilon_{xz} = \epsilon_{zx} = \epsilon_{yz} = \epsilon_{zy} = 0, \qquad (73)$$

where we used (64) and set

$$V_A = \frac{B_0}{\sqrt{4\pi m_i n_{i0}}}.$$

We also ignored the Landau damping on electrons. Dispersion relation (43) separates into two equations:

$$\epsilon_{zz}(\omega, k_z) = 0, \tag{74}$$

which is not influenced by the rotation of the grain, and

$$\begin{aligned} H(\omega) = \frac{k^2c^2}{\omega^2} - \frac{c^2}{V_A^2} \\ + \frac{\omega_{pd}^2}{\omega_{cd}(\omega \pm \omega_{cd})} = -\frac{2\Omega_r^2}{(\omega \mp \Omega_0)^2}. \end{aligned} \tag{75}$$

We now assume that $\omega \ll |\omega_{cd}|$. Setting

$$\omega = \Omega_0 + i\gamma,$$

where

$$\Omega_0 = kV_A$$

(Alfvén waves), and using (58), we obtain the growth rate

$$\gamma \approx \left(\frac{\Omega_r^2}{k^2c^2}\right)^{1/3} kv_A. \tag{76}$$

We next consider the longitudinal waves for which the dispersion relation assumes the form

$$\frac{k_\perp^2}{k^2}\epsilon_{xx}(\omega, \mathbf{k}) + \frac{k_z^2}{k^2}\epsilon_{zz}(\omega, \mathbf{k}) = 0, \tag{77}$$

where the components $\epsilon_{xx}$ and $\epsilon_{zz}$ for the cold plasma are defined by (70) and (72). Inserting the latter equation in (77), we obtain

$$\begin{aligned} 1 - \frac{k_\perp^2}{k^2}\sum_\beta \frac{\omega_{p\beta}^2}{\omega^2 - \omega_{c\beta}^2} - \frac{k_z^2}{k^2}\sum_\beta \frac{\omega_{p\beta}^2}{\omega^2} \\ = \frac{k_\perp^2}{k^2}\left[\frac{\Omega_r^2}{(\omega - \Omega_0)^2} + \frac{\Omega_r^2}{(\omega + \Omega_0)^2}\right]. \end{aligned} \tag{78}$$

It follows from (78) that the dust grain rotation contributes only for waves with $k_\perp \neq 0$, because the electric field of the longitudinal waves then has a component that lies in the dust grain rotation plane.

To obtain the growth rates for longitudinal waves, we use the same procedure as was used to deduce Eqs. (58) and (60).

We now consider the lower hybrid waves with $|k_z|V_{te}$, $|k_z|V_{ti}$, $\omega_{ci} \ll \omega \ll |\omega_{ce}|$. Setting

$$\omega = \Omega_0 + i\gamma,$$

where

$$\Omega_0 = \omega_0 = \frac{\omega_{pi}\omega_{ce}}{\sqrt{\omega_{pe}^2 + \omega_{ce}^2}}, \tag{79}$$

we find the growth rate

$$\gamma = \left(\frac{1}{2}\frac{\Omega_r^2}{\omega_{pi}^2}\right)^{1/3}\Omega_0. \tag{80}$$

Next, we consider the frequency regime where

$$|k_z|V_{td}, \quad |k_z|V_{ti} \ll \omega \ll |k_z|V_{te}.$$

The dielectric permittivity components in Eq. (77) are now defined by Eqs. (72) and (78). Using these expressions, we obtain the dispersion relation

$$\begin{aligned} 1 + \frac{1}{k^2r_{De}^2} - \frac{k_\perp^2}{k^2}\left(\frac{\omega_{pi}^2}{\omega^2 - \omega_{ci}^2} + \frac{\omega_{pd}^2}{\omega^2 - \omega_{cd}^2}\right) \\ - \frac{k_z^2}{k^2}\frac{\omega_{pi}^2}{\omega^2} = \frac{k_\perp^2}{k^2}\left[\frac{\Omega_r^2}{(\omega - \Omega_0)^2} + \frac{\Omega_r^2}{(\omega + \Omega_0)^2}\right]. \end{aligned} \tag{81}$$

This equation can be analyzed in two limiting cases. First, we consider the ion-cyclotron waves with $\omega_{cd} \ll \omega \approx \Omega_{ci}$ and $k_z \ll k_\perp$. Setting

$$\omega \approx \omega_0 + i\gamma,$$

we then obtain the growth rate

$$\gamma \approx \left(\frac{\Omega_r^2}{2\omega_{pi}^2}\frac{k_\perp^4c_s^4}{\Omega_0^4}\right)^{1/3}\Omega_0, \tag{82}$$

where

$$\Omega_0 = \omega_0 = (\omega_{ci}^2 + k_\perp^2c_s^2)^{1/2}, \quad c_s = \left(\frac{n_{i0}T_e}{n_{e0}m_i}\right)^{1/2}$$

is the ion acoustic speed.

Second, we consider the modified dust ion-acoustic waves (MDIAWs) characterized by $\omega_{pd}$, $\omega_{cd} \ll \omega \ll \omega_{ci}$. In this case, Eq. (81) gives

$$\begin{aligned} 1 + k^2r_{De}^2 + k_\perp^2\rho_s^2 - \frac{k_z^2c_{ss}^2}{\omega^2} \\ = k_\perp^2r_{De}^2\left[\frac{\Omega_r^2}{(\omega - \Omega_0)^2} + \frac{\Omega_r^2}{(\omega + \Omega_0)^2}\right], \end{aligned} \tag{83}$$

where

$$\rho_s = \frac{r_{De}\omega_{pi}}{\omega_{ci}} \equiv \frac{c_s}{\omega_{ci}}.$$

Equation (83) admits an instability of the MDIAWs with the frequency

$$\Omega_0 = \omega_0 = \frac{k_z c_s}{(1 + k^2 r_{De}^2 + k_\perp^2 \rho_s^2)^{1/2}},$$

and the growth rate is

$$\gamma \approx \left(\frac{k_\perp^2}{2k^2}\frac{\Omega_r^2}{\omega_{pi}^2}\right)^{1/3}\Omega_0. \qquad (84)$$

Finally, we consider coupled dust acoustic-dust cyclotron waves in a dust-electron plasma (without ions) with positive dust grains [27]. For $kV_{td} \ll \omega \ll k_z V_{te}$, we then have

$$1 + \frac{1}{k^2 r_{De}^2} - \frac{k_\perp^2}{k^2}\frac{\omega_{pd}^2}{\omega^2 - \omega_{cd}^2} - \frac{k_z^2}{k^2}\frac{\omega_{pd}^2}{\omega^2} = \frac{k_\perp^2}{k^2}\left[\frac{\Omega_r^2}{(\omega - \Omega_0)^2} + \frac{\Omega_r^2}{(\omega + \Omega_0)^2}\right]. \qquad (85)$$

For $\omega \ll |\omega_{cd}|$, Eq. (85) admits an instability of short wavelength DIWs when

$$\Omega_0 = \omega_0 = \frac{k r_{De}\omega_{pd}}{(1 + k^2 r_{De}^2 + k_\perp^2 \rho_{sd}^2)^{1/2}},$$

where

$$\rho_{sd} = \lambda_{De}\omega_{pd}/\omega_{cd}.$$

The growth rate of this instability is

$$\gamma \approx \left(\frac{1}{2}\frac{k_\perp^2}{k^2}\frac{\Omega_r^2}{\omega_{pd}^2}\right)^{1/3}\Omega_0. \qquad (86)$$

On the other hand, for $\omega \sim |\omega_{cd}|$, $k_\perp \gg k_z$, and $k^2 r_{De}^2 \ll 1$, an instability of the dust cyclotron waves occurs when

$$\Omega_0 = \omega_0(\omega_{cd}^2 + k^2 r_{De}^2 \omega_{pd}^2)^{1/2}.$$

The growth rate of the instability is

$$\gamma \approx \left(\frac{1}{2}k_\perp^4 r_{De}^4 \frac{\Omega_r^2}{\Omega_0^2}\frac{\omega_{pd}^2}{\Omega_0^2}\right)^{1/3}\Omega_0. \qquad (87)$$

It is interesting to note that a dust-electron plasma with positively charged grains can occur in the Earth's polar mesosphere [28, 29], where the grains are irradiated by the sun light, in which case the grains act as a source of electrons and collect ions from the ambient plasma to become positively charged. There also is the prediction [30] that positively charged dust grains in retrograde orbits are most likely to be observed by the Cosmic Dust Analyzer aboard the Cassini Orbiter mission to Saturn. Furthermore, the dust electron plasma can also be created in a laboratory discharge when the dust grains are irradiated by ultraviolet (UV) radiation [31–34].

## 6. SUMMARY AND CONCLUSIONS

In this paper, we have developed the electrodynamics of a magnetized dusty plasma taking the finite size of elongated and rotating charged dust grains into account. Starting from an appropriate Lagrangian for charged dust grains, we have derived the dust charge and dust current densities, as well as a kinetic equation for charged dust grains and the corresponding equations of motion in the external magnetic field. The effects of the dipole moment and the principal moment of inertia of the elongated and rotating dust grains are self-consistently incorporated. The newly derived dust charge and dust current densities, together with the corresponding quantities for electrons and ions, are combined with the Maxwell–Vlasov system of equations to obtain dielectric response functions for a magnetized dusty plasma.

For a cold dust gas, we have obtained explicit expressions for the permittivities associated with the dust grain rotation and for those of the ambient plasma species. The dispersion relations for transverse and longitudinal waves were then derived. Our analytical results exhibit the instabilities of the electron whistler, the dust whistler, the Alfvén waves, electromagnetic ion and dust cyclotron waves, as well as lower-hybrid, electrostatic ion-cyclotron, and coupled dust acoustic and dust cyclotron waves. The instability arises due to the resonance interaction between waves and elongated rotating dust grains. The free energy stored in the dust grain rotational motion is basically coupled to both the electromagnetic and electrostatic waves, driving them at nonthermal levels. The presence of nonthermal fluctuations can be used for diagnostic purposes. For example, coherent or incoherent scatterings of star light and/or electromagnetic waves off nonthermal fluctuations in cosmic plasmas may yield valuable information regarding the light polarization, the dust number density and the dust charge in situ, and other plasma parameters including the external magnetic field strength. We stress that the oscillating electric fields of electromagnetic waves may produce dust grain rotation, the energy of which is required for driving waves at nonthermal levels.

In conclusion, we emphasize that the present investigation should be useful for understanding waves and instabilities in astrophysical and laboratory plasmas that contain elongated and rotating charged dust grains. Finally, we suggest that new laboratory experiments in a weakly coupled dusty magnetoplasma must be designed to test the ideas described in this paper. A recent experimental work by Molotkov *et al.* [14] has conclusively demonstrated the Coulomb crystallization of 300-μm highly charged elongated cylindrical grains (with $|q| \sim 7.7 \times 10^5$ and with the length-to-diameter ratio 20–40) of the mass density 1.1 g/cm$^3$ and the diameters 15 and 7.5 μm in a low-pressure gas discharge plasma, where the electron energy ranges between 1–10 eV. Thus, a sheath electric field of the order 30 V/cm can levitate the grain.

Molotkov *et al.* [14] have discussed the role of the induced dipole moment that can influence the grain orientation. At small pressures (0.1 Torr), they also observed oscillations with the wavelength ~1 mm and the frequency 20–50 Hz. The latter can be associated with the dust acoustic waves that are deduced from our Eq. (83). Furthermore, by applying the external magnetic field 1–6 kG and with the plasma ($n_{i0} \sim 10^9$ cm$^{-3}$ and $n_{d0} \sim 10^3$ cm$^{-3}$) and dust parameters similar to those in Molotkov *et al.* [14], one should be able to observe the magnetization of ions and the electrostatic ion-cyclotron wave instability described by (81).

Finally, we mention that several authors [35–38] have experimentally observed rotation of spherical dust grains by magnetic fields. The rotation is attributed to the azimuthal $\mathbf{E} \times \mathbf{B}_0$ ion drift, which also drags the dust grain along due to the space charge electric field that is set up between ions and grains.

## ACKNOWLEDGMENTS

This work was partially supported by the Deutsche Forschungsgemeinschaft (Bonn) through the Sonderforschungsbereich 191 “Physikalische Grundlagen der Niedertemperatur Plasmen”, by the European Commission (Brussels) through Contract no. HPRN-CT-2000-00140 for the Human Potential and Training Network entitled “Complex Plasmas: The Science of Laboratory Colloidal Plasmas and Mesospheric Charged Aerosols”, and by the NATO project “Studies of Collective Processes in Dusty Plasmas” (Grant no. SA (PST.CGL974733)5066). One of the authors (P.K. Shukla) also acknowledges the support of the International Space Science Institute (Bern, Switzerland) for the international team “Complex Plasmas in Space: Applications to the Solar System and Near Earth Space.”

## APPENDIX: VARIOUS COMPONENTS OF $\sigma^r_{ij}$

The components of $\sigma^r_{ij}$ are given by

$$\sigma^r_{ij}(\omega, \mathbf{k}) = \begin{pmatrix} \sigma^r_{xx} & \sigma^r_{xy} & 0 \\ -\sigma^r_{xy} & \sigma^r_{yy} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{A.1}$$

where

$$\sigma^r_{xx} = \sigma^r_{yy} = i\frac{n_{d0}d^2}{4I}\frac{1}{KV_{td}} \times \sum_{n=-\infty}^{\infty} \exp(-z_d) I_n(z_d)(\Phi^n_- + \Phi^n_+), \tag{A.2}$$

$$\sigma^r_{xy} = \frac{n_{d0}d^2}{4I}\frac{1}{KV_{td}} \times \sum_{n=-\infty}^{\infty} \exp(-z_d) I_n(z_d)(\Phi^n_- - \Phi^n_+), \tag{A.3}$$

with

$$\begin{aligned} \Phi^n_\pm = {} & \frac{k_z^2}{K^2}\frac{KV_{td}}{\omega \pm \Omega_0 - n\omega_{cd}} J_+\left(\frac{\omega \pm \Omega_0 - n\omega_{cd}}{KV_{td}}\right) \\ & - \left(\frac{\kappa^2}{K^2}\frac{\omega - n\omega_{cd}}{KV_{td}} \mp \frac{k_z^2}{K^2}\frac{\Omega_0}{KV_{td}}\right) \\ & \times \left[1 + J_+\left(\frac{\omega \pm \Omega_0 - n\omega_{cd}}{KV_{td}}\right)\right]. \end{aligned} \tag{A.4}$$

Here,

$$V_{td} = (T_d/m_d)^{1/2}$$

is the grain thermal velocity,

$$z_\beta = \frac{k_\perp^2 V_{t\beta}^2}{\omega_{c\beta}^2}, \quad K = \sqrt{k^2 + \kappa^2}, \quad \kappa = \sqrt{\frac{m_d}{I}},$$

$I_n(z)$ is the Bessel function of an imaginary argument, and the function

$$J_+(x) = \frac{x}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dt \exp\left(-\frac{t^2}{2}\right)(x - t)^{-1}$$

has the asymptotic behavior

$$J_+(x) = 1 + \frac{1}{x^2} + \frac{3}{x^2} + \ldots - i\sqrt{\frac{\pi}{2}}x\exp(-x^2/2) \tag{A.5}$$

for $|x| \gg 1$, $|\mathrm{Re}\,x| \gg |\mathrm{Im}\,x|$, and $\mathrm{Im}\,x < 0$ and

$$J_+(x) \approx -i\sqrt{\frac{\pi}{2}}x \tag{A.6}$$

for $|x| \ll 1$.

For the tensor $\sigma^\beta_{ij}(\omega, \mathbf{k})$, we have

$$\begin{aligned} \sigma^\beta_{ij}(\omega, \mathbf{k}) = {} & \frac{e_\beta^2}{\omega_{c\beta}}\int d\mathbf{p}\, v_i(\psi)\int_{\pm\infty}^{\psi} d\psi'\left(\frac{\partial f_{\beta 0}}{\partial p_j}\right) \\ & \times \exp\left\{i\int_{\psi'}^{\psi}[(\omega - \mathbf{k}\cdot\mathbf{v}(\psi''))/\omega_{c\beta}]d\psi''\right\}, \end{aligned} \tag{A.7}$$

where $e_\beta$ is $q$ for $\beta = d$.

Straightforward calculations lead to the following expressions for the dielectric permittivity tensor [22]:

$$\epsilon_{ij}(\omega, \mathbf{k}) = \hat{\epsilon}_{ij}(\omega, \mathbf{k}) + \epsilon^r_{ij}(\omega, \mathbf{k}), \tag{A.8}$$

where

$$\hat{\epsilon}_{xx} = 1 - \sum_{\beta}\sum_{n}\frac{n^2\omega_{p\beta}^2}{\omega(\omega - n\omega_{cb})}\frac{A_n(z_\beta)}{z_\beta}J_{+}(\xi_n), \tag{A.9}$$

$$\hat{\epsilon}_{yy} = \hat{\epsilon}_{xx}$$
$$+ 2\sum_{\beta}\sum_{n}\frac{\omega_{p\beta}^2 z_\beta}{\omega(\omega - n\omega_{c\beta})}A'_n(z_\beta)J_{+}(\xi_n), \tag{A.10}$$

$$\hat{\epsilon}_{xy} = -\hat{\epsilon}_{yx}$$
$$= -i\sum_{\beta}\sum_{n}\frac{n^2\omega_{p\beta}^2}{\omega(\omega - n\omega_{c\beta})}A'_n(z_\beta)J_{+}(\xi_n), \tag{A.11}$$

$$\hat{\epsilon}_{xz} = \hat{\epsilon}_{zx}$$
$$= \sum_{\beta}\sum_{n}\frac{n^2\omega_{p\beta}^2 k_\perp}{\omega\omega_{c\beta}k_z}\frac{A_n(z_\beta)}{z_\beta}[1 - J_{+}(\xi_n)], \tag{A.12}$$

$$\hat{\epsilon}_{yz} = -\hat{\epsilon}_{zy}$$
$$= -i\sum_{\beta}\sum_{n}\frac{\omega_{p\beta}^2 k_\perp}{\omega\omega_{c\beta}k_z}A'_n(z_\beta)[1 - J_{+}(\xi_n)], \tag{A.13}$$

$$\hat{\epsilon}_{zz} = 1$$
$$+ \sum_{\beta}\sum_{n}\frac{\omega_{p\beta}^2(\omega - n\omega_{c\beta})}{\omega k_z^2 V_{t\beta}^2}A_n(z_\beta)[1 - J_{+}(\xi_n)], \tag{A.14}$$

and

$$\epsilon_{ij}^r(\omega, \mathbf{k}) = \frac{4\pi i}{\omega}\sigma_{ij}^r(\omega, \mathbf{k}). \tag{A.15}$$

Here,

$$A_n(z) = \exp(-z)I_n(z), \quad \xi_n = \frac{\omega - n\omega_{c\beta}}{|k_z|V_{t\beta}},$$

$$\omega_{p\beta} = \sqrt{\frac{4\pi e_\beta^2 n_{\beta 0}}{m_\beta}}$$

is the plasma frequency of the species $\beta$, and $\mathbf{k} \equiv (k_\perp, 0, k_z)$.

## REFERENCES

1. N. N. Rao, P. K. Shukla, and M. Y. Yu, Planet. Space Sci. **38**, 345 (1990).
2. P. K. Shukla, Phys. Scr. **45**, 504 (1992).
3. P. K. Shukla and V. P. Silin, Phys. Scr. **45**, 508 (1992).
4. A. Barkan, R. L. Merlino, and N. D’Angelo, Phys. Plasmas **2**, 3563 (1995); C. Thomson, A. Barkan, N. D’Angelo, and R. L. Merlino, Phys. Plasmas **4**, 2331 (1997); R. L. Merlino, A. Barkan, C. Thompson, and N. D’Angelo, Phys. Plasmas **5**, 1607 (1998).
5. H. R. Prabhakara and V. L. Tanna, Phys. Plasmas **3**, 3176 (1996).
6. E. Thomas, Jr. and M. Watson, Phys. Plasmas **6**, 4111 (1999); V. E. Fortov, A. G. Kharpak, S. A. Kharpak, *et al.*, Phys. Plasmas **7**, 1374 (2000); V. I. Molotkov, A. P. Nefedov, V. M. Torchinskii, *et al.*, Zh. Éksp. Teor. Fiz. **116**, 902 (1999) [JETP **89**, 477 (1999)].
7. A. Barkan, N. D’Angelo, and R. L. Merlino, Planet. Space Sci. **44**, 239 (1996).
8. Y. Nakamura, H. Bailung, and P. K. Shukla, Phys. Rev. Lett. **83**, 1602 (1999).
9. P. K. Shukla, in *Frontiers in Dusty Plasmas*, Ed. by Y. Nakamura, T. Yokota, and P. K. Shukla (Elsevier, Amsterdam, 2000), p. 3.
10. F. Verheest, *Waves in Dusty Plasmas* (Kluwer, Dordrecht, 2000).
11. M. Harwit, *Astrophysical Concepts* (Springer-Verlag, New York, 1988), p. 405.
12. L. Spitzer, Jr., *Physical Processes in Interstellar Medium* (Wiley, New York, 1977; Mir, Moscow, 1981), p. 182.
13. U. Mohideen, H. U. Rahman, M. A. Smith, *et al.*, Phys. Rev. Lett. **81**, 349 (1998); H. U. Rahman, U. Mohideen, M. A. Smith, *et al.*, Phys. Scr. T **89**, 186 (2000).
14. V. I. Molotkov, A. P. Nefedov, M. Yu. Putsyl’ink, *et al.*, Pis’ma Zh. Éksp. Teor. Fiz. **71**, 152 (2000) [JETP Lett. **71**, 102 (2000)].
15. D. Tskhakaya and P. K. Shukla, Phys. Lett. A **279**, 243 (2001).
16. N. V. Voshchinnikov and D. A. Semenov, Pis’ma Astron. Zh. **26**, 789 (2000) [Astron. Lett. **26**, 679 (2000)].
17. J. Mahmoodi, P. K. Shukla, N. L. Tsintsadze, and D. D. Tskhakaya, Phys. Rev. Lett. **84**, 2626 (2000). We have noted an error in Eq. (27), which must contain an addition term $(k_\perp/k^2)\Omega_r^2/(\omega + \Omega_0)^2$, respectively, on its right-hand side. However, this term does not affect our final finding [viz. our Eq. (31) remains intact], as we have dealt with the resonance case $\omega \simeq \Omega_0$.
18. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields* (Nauka, Moscow, 1973; Pergamon, Oxford, 1975).
19. N. N. Bogolyubov, *Dynamic Problems of Statistical Physics* (Gostekhizdat, Moscow, 1946).
20. Yu. L. Klimontovich, *Kinetic Theory of Nonideal Gases and Nonideal Plasmas* (Nauka, Moscow, 1975; Pergamon, Oxford, 1982).
21. L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1989).
22. A. F. Alexandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Principles of Plasma Electrodynamics* (Vysshaya Shkola, Moscow, 1988; Springer-Verlag, Berlin, 1984).
23. A. B. Mikhailovskii, *Theory of Plasma Instabilities: Instabilities of a Homogenous Plasma* (Atomizdat, Moscow, 1971; Consultants Bureau, New York, 1974), Vol. 1.
24. M. Rosenberg and D. A. Mendis, IEEE Trans. Plasma Sci. **20**, 929 (1992); D. A. Mendis and M. Rosenberg, Annu. Rev. Astron. Astrophys. **32**, 410 (1994).
25. P. K. Shukla and H. U. Rahman, Phys. Plasmas **3**, 430 (1996).
26. P. K. Shukla, Phys. Lett. A **252**, 340 (1999).
27. P. K. Shukla, Phys. Rev. E **61**, 7249 (2000).

28. O. Havnes, J. Troim, T. Blix, *et al.*, J. Geophys. Res. **101**, 10839 (1996).
29. M. Horányi, B. Walch, S. Robertson, *et al.*, J. Geophys. Res. **103**, 8575 (1998).
30. J. E. Howard, H. R. Dullin, and M. Horányi, Phys. Rev. Lett. **84**, 3244 (2000).
31. V. W. Chow, D. A. Mendis, and M. Rosenberg, IEEE Trans. Plasma Sci. **22**, 179 (1994); M. Rosenberg and D. A. Mendis, IEEE Trans. Plasma Sci. **23**, 177 (1995); M. Rosenberg, D. A. Mendis, and D. P. Sheehan, IEEE Trans. Plasma Sci. **27**, 239 (1999).
32. V. E. Fortov, A. P. Nefedov, O. F. Petrov, *et al.*, Phys. Lett. A **219**, 89 (1996).
33. V. E. Fortov, A. P. Nefedov, and O. S. Vaulina, Zh. Éksp. Teor. Fiz. **114**, 2004 (1998) [JETP **87**, 1087 (1998)]; V. Fortov, V. I. Molotkov, A. P. Nefedov, and O. F. Petrov, Phys. Plasmas **6**, 1759 (1999).
34. A. A. Samaryan, O. S. Vaulina, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **118**, 119 (2000) [JETP **91**, 106 (2000)]; A. A. Sickafoose, J. W. Colwell, M. Horányi, and S. Robertson, Phys. Rev. Lett. **84**, 6034 (2000).
35. H. Fujiyama, H. Kawasaki, S. C. Yang, and Y. Matsuda, Jpn. J. Appl. Phys. **33**, 4216 (1994); S. Nunomura, N. Ohno, and S. Takamura, Jpn. J. Appl. Phys. **36**, 877 (1997).
36. G. Uchida, R. Ozaki, S. Iizuka, and N. Sato, in *Proceedings of the International Congress on Plasma Physics, European Physical Society, Prague, 1998,* Ed. by P. Pavlo, p. 2557; N. Sato, G. Uchida, R. Ozaki, S. Iizuka, and T. Kamimura, in *Frontiers in Dusty Plasmas,* Ed. by Y. Nakamura, T. Yokota, and P. K. Shukla (Elsevier, Amsterdam, 2000), p. 329.
37. D. W. Law, H. Steel, B. M. Annaratone, and J. E. Allen, Phys. Rev. Lett. **80**, 4189 (1998).
38. U. Konopka, D. Samsonov, A. V. Ivlev, *et al.*, Phys. Rev. E **61**, 1890 (2000).

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 313–323.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 353–365.*
*Original Russian Text Copyright © 2001 by Vladimirov, Deputatova, Nefedov, Fortov, Rykov, Khudyakov.*

**PLASMA,
GASES**

# Dust Vortices, Clouds, and Jets in Nuclear-Induced Plasmas

**V. I. Vladimirov[a], L. V. Deputatova[a, *], A. P. Nefedov[a],
V. E. Fortov[a], V. A. Rykov[b, **], and A. V. Khudyakov[b]**

[a]*Institute of High Energy Densities, Institute of High Temperatures Scientific Association (IVTAN),
Russian Academy of Sciences, Izhorskaya ul. 13/19, Moscow, 127412 Russia*
[b]*Leipunskii State Research Center "Institute of Physics and Power Engineering,"
pl. Bondarenko 1, Obninsk, Kaluzhskaya oblast, 249020 Russia*
**e-mail: dlv@hedric.msk.su*
***e-mail: rykov@ippe.obninsk.ru*

Received January 24, 2001

**Abstract**—The collective movement of dust particles in a plasma formed during deceleration of decay products of californium nuclei in neon is investigated experimentally. For the first time, compact vortex structures containing a large number of coagulating dust particles and dense dust clouds evolving in time are observed. Dust formations have clearly defined boundaries and particles in them form ordered liquid-type structures. Under steady-state conditions, dust structures exist from several minutes to hours. An increase in the voltage applied to the high-voltage electrode leads to the formation of dust particle jets. A change in the electric field configuration transforms the structures from one type to another. A strong recombination of electrons and ions at dust particles is observed. The momentum transfer from ions drifting in an external field to gas molecules is studied using the Monte Carlo method. It is shown that the transferred momentum is so large that it may cause a gas flow. The characteristic features of vortex flow in neon and in air are explained. *© 2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

Nuclear-induced dust plasma is created by ionizing particles appearing in nuclear reactions during nuclear fission as well as during the α and β decay [1, 2]. A typical feature of such plasma is its space–time inhomogeneity associated with its track structure [3]. The accumulation of charge by particles in this type of plasma is stochastic [4] and the time of charge variation is determined by the intensity of decay in a radioactive source and by the distance from the source. A dust particle in a nuclear-induced plasma in an external electric field experiences the action of electron bunches and ion clusters drifting towards different electrodes and having (in the case of a uniform field) a cylindrical shape with a symmetry axis parallel to the trajectory of the ionizing particle. As a result of diffusion, the electron bunches and ion clusters spread in the radial direction, the difference in the diffusion coefficients resulting in a considerable increase in the radii of electron bunches. These formations encounter dust particles more frequently and transfer a negative charge to them. The less frequent action of ion clusters effectively discharges dust particles. Alternating action of electron bunches and ion clusters leads to strong fluctuations of the electric charge of a dust particle [4]. In an external electric field under a considerable gas pressure, the drift of electrons and ions to the electrodes causes a momentum transfer to neutral components of the medium. In view of the large velocity and short time of the electron drift to the anode, the momentum transferred by electrons to neutral particles is negligibly small but the momentum transfer by ions plays the decisive role. As a result, the gas performs a motion whose type depends not only on the geometry of the volume occupied by the gas, but also on the concentration of dust particles, their mass and charge. Such a movement complicates the formation of stationary dust structures and leads to the evolution of vortices, streamlined clouds, and jets of dust particles.

The present work aims at experimental investigating the behavior of dust particles of various diameters in a nuclear-induced plasma in the presence of electric fields with a varying spatial configuration.

## 2. EXPERIMENTAL SETUP

The experiments were made in a setup similar to that described in [1]. As a source of ionizing radiation, we used a thin layer of $^{252}$Cf whose nuclei experience alpha-decay and spontaneous fission in a ratio approximately equal to 32 : 1. The intensity of the source was $4 \times 10^6$ divisions/s. For such an intensity of the radioactive source, the concentration of Ne ions near its surface is $N_i \sim 3 \times 10^9$ cm$^{-3}$. The source was mounted on an earthed metallic electrode (Fig. 1) made in the form of a disk of diameter 44 mm. A high-voltage electrode of the same diameter was arranged at a distance of 3.5 cm from the source. In the vicinity of this electrode, the ion

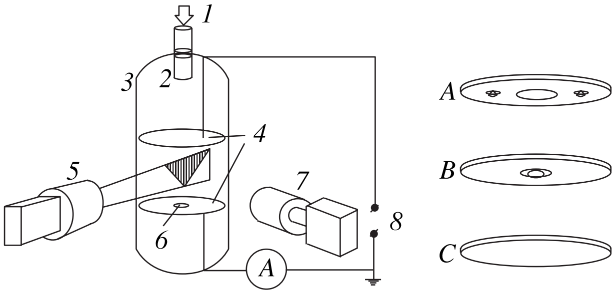

**Fig. 1.** Experimental setup: injection of gas–dust mixture from the evacuation and gas-filling system (*1*), netlike container of dust particles (*2*), glass walls (*3*), metallic electrodes (*4*), laser with a cylindrical lens (*5*), 2D radioactive source (*6*), video camera (*7*), dc source (*8*), and various types of high-voltage electrode (*A*, *B*, *C*).

concentration was $N_i \sim 1 \times 10^8$ cm$^{-3}$. The ion concentration was determined from the calculated coordinate dependence of the power density for the energy contribution of fission fragments and α-particles to the gas in analogy with [1]. Additional electrodes were mounted on the planar high-voltage electrode to create a nonuniform field. Figure 1 shows electrode *A* with a hole of diameter 15 mm and two copper tubes of diameter 2 mm inserted through an insulator. The auxiliary electrode *B* has a hemispherical shape.The electrodes were placed in a sealed cylindrical glass tube with an inner diameter of 50 mm. Different voltages were applied to the high-voltage and auxiliary electrodes.

A gas–dust mixture was produced by a pulsed action of the neon flow supplied from a dispenser with a fixed volume. This flow was directed to a container with a netlike bottom containing the particles under investigation, and the formed gas–dust mixture uniformly filled the entire volume of the glass cell. The glass cell was preliminarily evacuated to a pressure of about 1.3 Pa to prevent the loss of electrons to oxygen molecules. The gas pressure was varied from $10^4$ to $10^5$ Pa. We used Zn particles with a mean diameter of 2.4 μm and a mass of $5 \times 10^{-11}$ g as dust particles. The initial concentration of particles varied from $10^5$ to $10^6$ cm$^{-3}$. The cell was exposed to a 2D laser beam with a constriction formed by a cylindrical lens; the constriction diameter could be varied from 100 to 200 μm with the help of a diaphragm. Laser radiation scattered by particles was registered by a video camera with a CCD matrix as a radiation detector.

## 3. DUST PARTICLE VORTICES

In order to analyze the effect of a nonuniform electric field on the motion of dust particles, we used a 2D electrode with auxiliary high-voltage electrodes *A*. The central hole in the electrode, which is intended for the injection of particles from the container, did not significantly affect the motion of particles. The potential *U* of the main electrode as well as the potentials $U^*$ of the auxiliary electrodes had positive values such that the following inequality holds:

$$U^* \geq U.$$

After the injection of the gas–dust mixture into the cell, the particles were separated by weight. The remaining particles were gradually accumulated into a rotational dust structure (Fig. 2a) whose center was located under an auxiliary electrode and was displaced towards the center of the cell. The particles were accelerated towards the auxiliary electrode (in the direction shown by the black arrow) on segment *1–2*, which can be explained by the existence of a charge on the particles. Then the particles were turned away from the upper electrode on segment *2–3*, decelerated, and ultimately involved in the rotational motion. The velocity $v$ of the particles was determined by the formula

$$v = \frac{N}{k}\frac{s}{F}, \qquad (1)$$

where *N* is the number of frames per second transmitted by the video camera for recording ($N = 25$ in our experiment), *k* is the number of frames on which a dust particle displaced over a distance *s* was observed, and *F* is the magnification of the optical system ($F = 10.5$ in our experiment). In the vicinity of point *2*, the velocities of particles exceed 10 cm/s, while the velocity on the opposite side (near point *4*) is half as small (approximately equal to 5 cm/s).

In a few minutes, the majority of the particles were concentrated at the center of the structure, while the remaining volume was almost completely free of particles (Fig. 2b). The center of the structure was separated from the center of the auxiliary electrode by 6 mm (Fig. 3). With the passage of time, agglomeration of small particles into coarser fragments could be seen, the finer fraction remaining in the central part (Fig. 2c). Upon an abrupt increase in the potential at the auxiliary electrode and subsequent restoration of the previous value of the potential, the vortex flow is at first violated and particles strive to fly apart (Fig. 2d), and then the motion of particles is restored (Figs. 2e and 2f). After the restoration, the vortex motion at this stage may continue for a long time under constant external conditions. (The structure depicted in Fig. 2g was observed for more than 20 min until the power supply was switched off.) It should be noted that the shape of the structure is close to a circle of diameter 5 mm in the plane of the laser “knife” (vertical cross section), while along the axis of observation it is spindle-shaped with a linear size of approximately 1.5 cm.

By placing the second auxiliary electrode symmetrically relative to the cell axis, we could observe two structures instead of one, which rotate in opposite directions (Fig. 2h). The type of motion of the particles

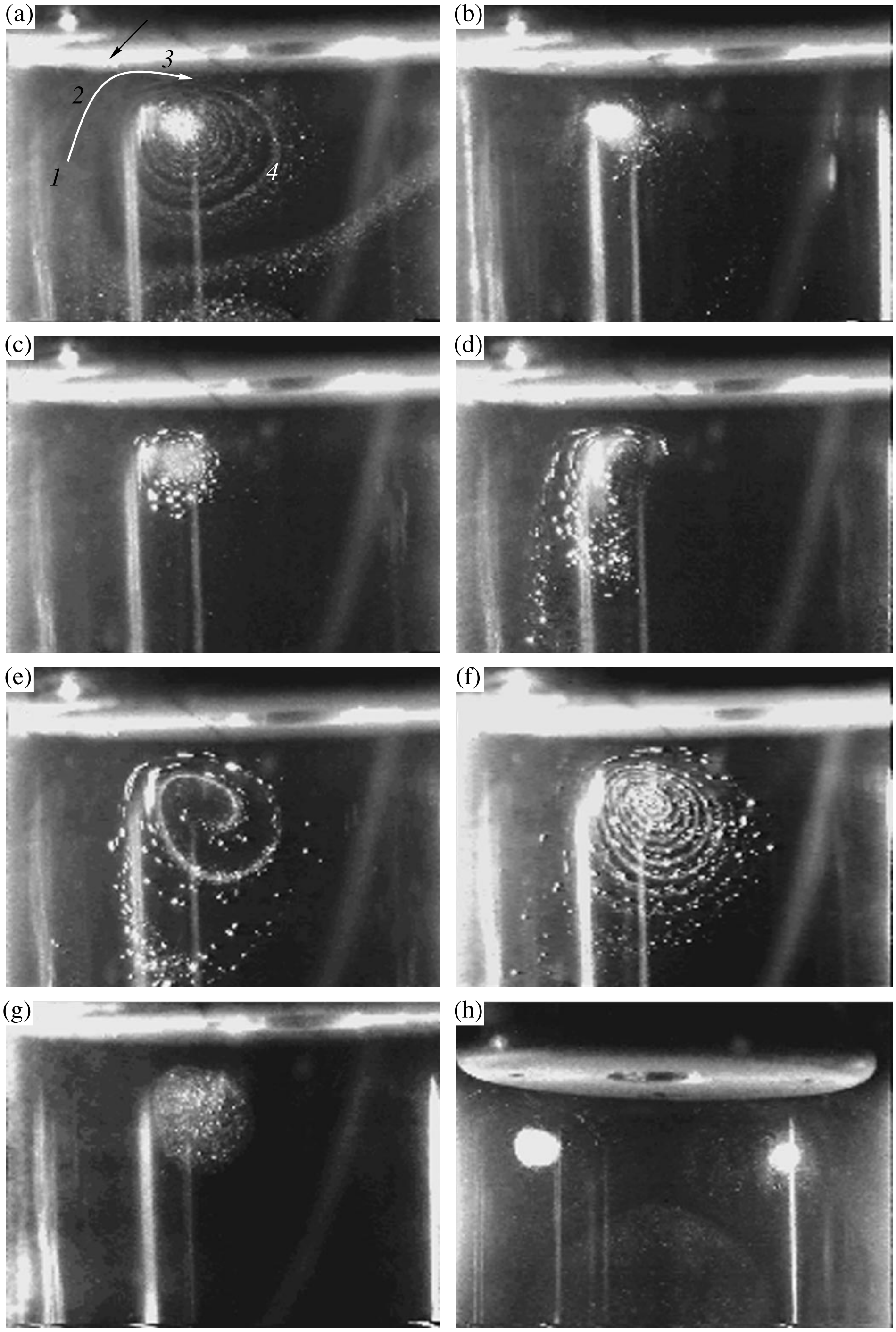

**Fig. 2.** Vortex flow of Zn dust particles for $U$ = 187 V, $U^*$ = 442 V, and neon pressure $0.4 \times 10^5$ Pa; frame size 3.2 × 2.4 cm: (a) 1.5 min, (b) 3 min, and (c) 4.5 min after the injection of the gas–dust mixture; (d–f) variation of the auxiliary electrode potential from 400 to 500 V during the eighth minute; (g) 10 min; (h) general view of two vortex structures during the fourth minute (frame size 4.2 × 3.1 cm). (Light strips in the frames are due to reflections at glass walls of the cell and at electrodes.)

does not change if the central hole in the main electrode is absent. For a low initial concentration of the particles, no dust structures are formed, indicating a clearly manifested collective type of the behavior of dust particles.

Since we used in our experiments polydisperse dust particles, their radii $r_d$ were determined experimentally from the steady-state velocity $v$ of their descent in the gas after the removal of the electric field (the resistance

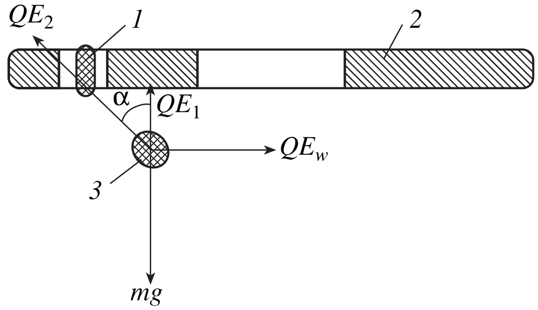

**Fig. 3.** Schematic diagram of the arrangement of dust structure: auxiliary electrode (*1*), main electrode (*2*), and dust structure (*3*).

to the motion of particles was taken into account using the Stokes formula):

$$r_d = \sqrt{\frac{9\eta v}{2\rho g}}, \tag{2}$$

where η is the gas viscosity, ρ is the density of dust particles, and *g* is the acceleration due to gravity. The rate of fall for the particles was found to be $v = 1.7 \pm 0.2$ cm/s. The radii of Zn particles constituting a rotating structure calculated by formula (2) varied from 5.5–6.2 μm. The average mass of particles was $3.5 \times 10^{-9}$ g. The fivefold increase in the radius and the nearly two-orders-of-magnitude increase in the particle mass point towards a coagulation of the particles in the rotating structure. It was mentioned above that this can be observed visually.

Let us consider the equilibrium conditions of a compact rotating dust formation (Fig. 3). It is acted upon by the field of two electrodes and the wall,

$$\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_w,$$

as well as the force of gravity $M\mathbf{g}$ and the force of pressure appearing due to different values of the velocity of the gas flow streamlining the dust structure. This pressure may be calculated by the formula

$$p = \frac{\rho_g}{2}(v_2^2 - v_4^2), \tag{3}$$

where $v_2$ and $v_4$ are the velocity of the gas in the vicinity of points *2* and *4* in Fig. 2a. Pressure (3) creates the force of pressure $\mathbf{F}_p$. The dust structure will be in equilibrium if (Fig. 3)

$$Q\mathbf{E} + M\mathbf{g} + \mathbf{F}_p = 0, \tag{4}$$

where *Q* is the charge of the structure. In view of the indeterminacy in the mass of the entire structure, we estimate the charge *q* of a particle constituting it. The equilibrium conditions lead to the following system of equations:

$$\begin{aligned} qE_2\sin\alpha + f_p\sin\beta &= qE_w, \\ q(E_1 + E_2\cos\alpha) &= mg - f_p\cos\beta. \end{aligned} \tag{5}$$

Here, α and β are the angles formed by vectors $\mathbf{E}_2$ and $\mathbf{F}_p$ with the vertical, $m = 4\pi\rho r^3/3$ is the particle mass, and $f_p$ is the force of pressure exerted on the particle. In our experiments, $f_p = \pi r^2 p \approx 3.6 \times 10^{-13}$ N. This value is negligibly small as compared to the force of gravity $mg \approx 3.5 \times 10^{-8}$ N. System (5) makes it possible to determine unknowns *q* and $E_w$ if the force of pressure $f_p$ is negligibly small and the remaining quantities appearing in it are known. The field strength created by the upper electrode near the center of the rotating structure is $E_1 \approx 45$ V/cm, while the strength of the field created by the auxiliary electrode is $E_2 \approx 100$ V/cm. The average angle $\alpha \approx 45°$. The strength of the field of the walls determined from Eqs. (5) is $E_w \approx 70$ V/cm. The charge of dust particles determined from the solution of system (5) was $2.7 \times 10^4$ in the electron charge units. On the average, there are approximately 270 electrons for each hundred smaller particles.

In order to obtain the exact value of the charge of the particles, one must solve the self-consistent problem in which the Poisson equation is written taking into account the external fields, the fields of the walls and space charges (electrons and ions), as well as the fields of particles whose charges themselves are functions of the local resultant electric field. However, such a self-consistent problem can hardly be solved at the moment; for this reason, effective potentials of interaction of dust particles are used in mathematical simulation of a dust plasma [5].

As the potential of the auxiliary electrode increases to 500 V, the rotating structure moves away from it to a distance of 9 mm. This can be explained by an increase in the electron flux to the wall, which leads to an increase in the force $qE_w$ pushing away the structure. For new values of $E_2$ and $E_w$, the structure finds its new equilibrium position in which conditions (5) are satisfied for the new values of the angles.

Coagulation of dust particles in the plasma created by a radioactive source was considered for the first time by Belov *et al.* [6]. These authors do not refute the coagulation of charged particles. However, the charge per small particle prior to coagulation may turn out to be too high in our case. On the other hand, it should be noted that the coagulation time in the described experiment is of the order of minutes. Coagulation may be facilitated by the nonsphericity of particles and their high conductivity. The particle charge may be slightly exaggerated in view of the entrainment of dust particles by a moving gas, which will be considered below.

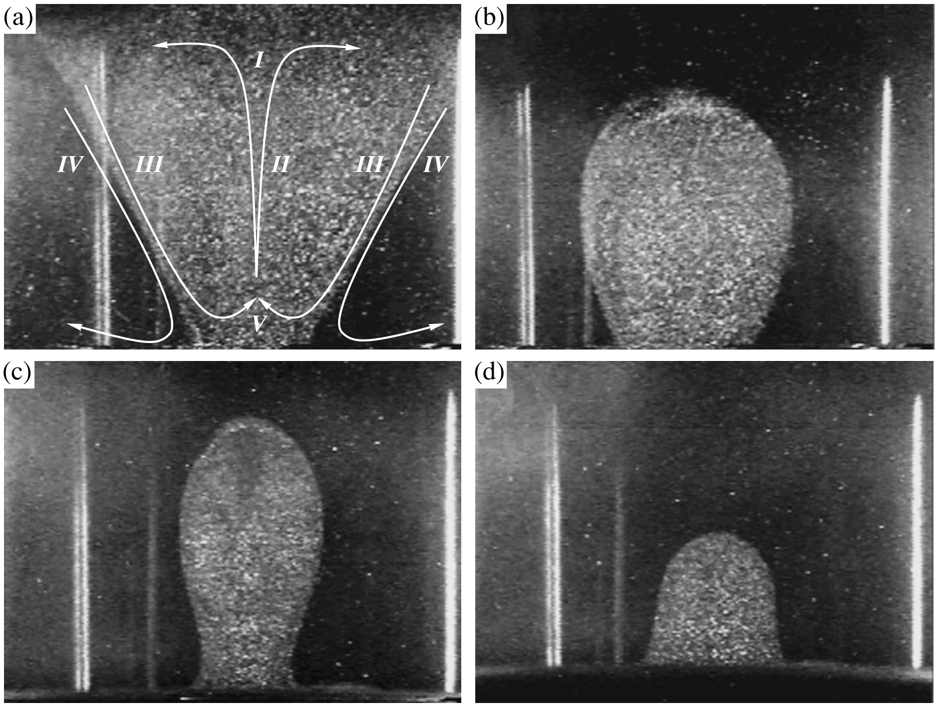

**Fig. 4.** Evolution of the dust cloud formed by Zn particles: (a) 2 min, (b) 4 min, (c) 4.5 min, (d) 4 min 45 s after the injection of the dust component. The upper electrode has the shape *C* (Fig. 1). The upper electrode potential is 152 V, the separation between the upper and lower electrodes is 3.5 cm, the neon pressure is $0.76 \times 10^5$ Pa; frame size is 4.2 × 3.1 cm.

## 4. LIQUID CLOUDS OF DUST PARTICLES

In the case when potentials $U$ and $U^*$ are identical and electrode $C$ is used (see Fig. 1), Zn dust particles, after the injection of the gas–dust mixture, coagulate in a few minutes into a cloud with well-defined boundaries (Figs. 4a–4d). The cloud has the shape of a truncated cone with the base lying in the plane of the upper electrode and the top near the radioactive source. The entire volume of the experimental cell can be divided into five regions in which the particles behave in different manners. In region I, particles are almost stationary (levitation). In region II, particles move slowly, creating an ascending flow. The velocity of particles in the flow decreases as they approach the boundaries of the structure and its upper part. In the middle part of the flow, the velocity has the maximum value equal to 0.6 cm/s. In region III, particles move downwards along the generator of the cone at a velocity of 1.2 cm/s; some of the particles are deflected to the axis of the structure in the vicinity of the lower electrode and then move upwards. Another fraction of particles get into region V. Region IV contains a very small number of particles whose velocity at the boundaries of the structure is equal to the velocity of particles in region III. When dust particles approach the lower electrode, they are deflected from it to the walls, where they ascend to form closed trajectories. In region V, all particles fall down at a rate approximately equal to 1 cm/s. The velocities of particles were determined by formula (1).

Thus, there are two types of vortex flows of particles. One of them is the motion of particles with a low number density at the periphery, while the other is the motion within the structure at approximately half the velocity. These flows are separated by a sharp boundary near which the velocities of particles are identical on both sides. It will be proven below that the momentum transfer from ions to neutral particles is large enough for initiating the motion of the gas. In our opinion, this is one of the main reasons behind the emergence of vortices. Another reason may be the intense recombination of ions and electrons at dust particles, which may give rise to large concentration gradients and auxiliary flows of plasma particles. The ionization and recombination of the gas (including that at dust particles) occur at the highest rate in the vicinity of the source of ionizing radiation. The ion flow reducing the diameter of the structure near this source is directed precisely at this region.

At the lowest part of the structure (region V), the charge of particles becomes positive [4], and they fall to the lower earthed electrode.

The radius $r_d$ of the particles was determined from the rate of fall of the particles of the structure after the connection of a power source between the electrodes [formulas (1) and (2)] and was found to lie in the interval 0.9–1.4 μm. The particle mass varied from $3 \times 10^{-11}$ to $8 \times 10^{-11}$ g.

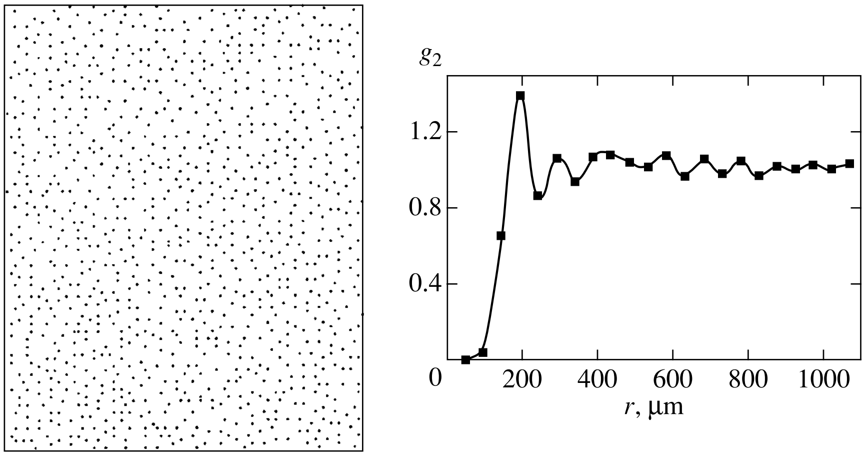

**Fig. 5.** Digital image of the central part of the structure and the corresponding pair distribution function.

The value of the electric charge $q$ of particles was calculated from the condition of equilibrium between levitating particles in region I and slowly moving particles in region II:

$$q = \frac{mg}{E}, \tag{6}$$

where $E$ is the strength of the external electric field. The charges of particles varied, depending on the radius, from 400 to 1000 units of the electron charge.

The mathematical processing of the central part of the structure give the average distance $b \approx 190 \pm 30$ μm between particles and the concentration $n_d \approx (3–4) \times 10^4$ cm$^{-3}$. The electric field of particles with such a concentration and with the maximum charge creates a field of strength $E \approx 4$ V/cm at a distance 1 cm from the center of the sphere of radius 1 cm (characteristic size of the structure), which is an order of magnitude smaller than the applied field and may be disregarded in the first approximation.

Digital processing of the frames makes it possible to obtain a pair distribution function for dust particles (pair correlation function [7]) presented in Fig. 5. It has a clearly manifested peak typical of liquid structures. The nonideality parameter for the dust component is given by

$$\Gamma = \frac{1}{4\pi\varepsilon_0}\frac{q^2}{bT}, \tag{7}$$

where $T$ is the gas temperature in energy units, which varies from 45 to 340 for particles of different sizes, which also speaks in favor of the liquid type of the obtained structures. For particles whose radius is equal to the average radius (1.2 μm), the nonideality parameter is 130. The spread in the size of particles, the fluctuations of their charge [4] and the motion of the medium considerably hinder the formation of crystalline dust structures.

Under a constant pressure and for constant potentials of electrodes, the cloud forms in a few minutes with the streamlined spherelike upper part (Fig. 4b). Then it gradually changes the contour of its boundaries and smoothly falls on the lower electrode (Figs. 4c and 4d). At the same time, the vortex motion of periphery particles, as if grinding the structure, creates a constriction at its lower part. In the situation depicted in Fig. 4d, the vortex motion also considerably affects the behavior of the upper part, exerting an additional pressure on it. After the structure falls on the lower electrode, the motion which was formerly typical of periphery particles embraces the entire volume. The velocity of particles in such a vortex near the cell axis is equal to 1.2 cm/s.

Recombination of plasma particles on dust particles is reflected in the time dependence of the electric current between the electrodes (Fig. 6). The figure shows the results of measurements of current for three different injections of the gas–dust mixture to the volume of the experimental cell. In the initial instant after the injection of the gas–dust mixture, the density and the

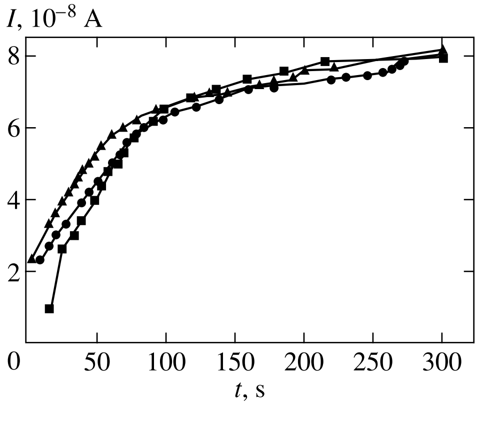

**Fig. 6.** Time dependence of current.

number of dust particles are large and the current is small. During the formation of the structure and its evolution in time, the volume occupied by dust particles decreases. The electric current increases and attains saturation after the falling of the dust structure on the lower electrode.

Another reason behind the decrease in the current in the presence of dust particles could be the collisions of dust particles with ionizing particles, after which an ionizing particle is stuck in a dust particle and stops moving. As a result of such a process, further ionization is terminated due to the interruption of the track. If we assume that dust particles are spheres of radius $r_d$, the mean free path of an ionizing particle relative to collisions with dust particles is calculated by the formula

$$\lambda = (n_d \pi r_d^2)^{-1}. \tag{8}$$

The calculations based on this formula gives a mean free path of 16 cm for the largest particles of radius 1.4 μm with the maximum concentration $10^6$ cm$^{-3}$. This value is much larger than the actual mean free path of fission fragments and α particles in neon near atmospheric pressure. Consequently, the loss of ionizing particles on dust particles can be neglected.

The clouds described above are formed under pressures from $0.1 \times 10^5$ to $1 \times 10^5$ Pa and their behavior is always the same. The gas pressure was never higher than the atmospheric pressure. In the absence of an electric field or a radioactive source, no clouds are formed, and the particles injected into the volume gradually fall down along the trajectories parallel to the acceleration due to gravity.

## 5. DUST JETS

If the potential of the upper electrode is increased after the formation of a dust particle structure, dust particles rush towards it with a velocity proportional to the potential. It is interesting to note that the entire structure does not move upwards. One or a few jets of moving particles are formed in its upper part. In front of the electrode, the velocity of particles in a jet decreases and the jet expands to form a funnel. The continuity condition typical of liquid media is satisfied. The jet may change its shape; the location of its bases may also change (Fig. 7). In the regions of space outside the structure and jets, the number density of dust particles is close to zero.

The use of an auxiliary electrode having a hemispherical shape and insulated from the main electrode by a dielectric makes it possible to remove dust particles from the cloud (Figs. 8a–8c). In this case, a single dust jet directed to this electrode is formed. As a rule, the potential at the auxiliary electrode is twice as high as at the main electrode. Dust particles attracted to the electrode remain on its surface. This effect can be used for designing a device for the removal of dust from technological volumes in nuclear reactors.

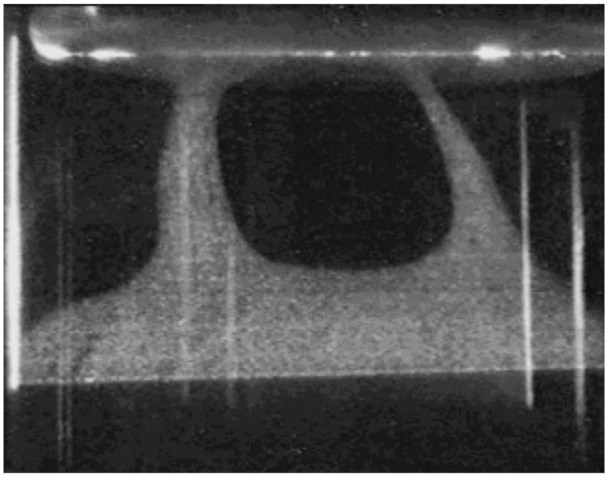

**Fig. 7.** Motion of CeO$_2$ dust particles in the form of jets directed to the upper electrode. The electrode is of the *C* shape (Fig. 1). The upper electrode potential is 200 V; the separation between the upper and lower electrodes is 3.5 cm, the neon pressure is $0.79 \times 10^5$ Pa; the frame size is $4.2 \times 3.1$ cm. The lower part of the cloud is not illuminated.

If a high voltage is supplied to the auxiliary electrode mounted at a certain distance from the center of the high-voltage electrode after the formation of a cloud of particles (as in Figs. 2a–2g), the dust cloud is transformed into a rotating structure similar to that considered in Section 3 (Fig. 8h).

## 6. CALCULATION OF MOMENTUM TRANSFER FROM DRIFTING FLOWS OF IONS AND ELECTRONS TO GAS MOLECULES

In this section we will demonstrate the important role of momentum transfer from drifting ions to neutral components of the medium in the formation of vortex motion of dust particles. Drifting ions acquire an additional momentum between collisions as a result of their interaction with an electric field. During the time of motion $t$, $N_i$ electrons acquire (and hence transfer to gas molecules) a momentum equal to $m_i w_i N_i t/\tau$ in the linear approximation, where $\tau$ is the mean time between collisions and $w_i$ is the drift. For the momentum $\Delta p$ transferred to a gas of mass $\Delta m$ during the drift time $T_d$, we can easily obtain the following relation:

$$\begin{aligned} \frac{\Delta p}{\Delta m} &\approx \frac{m_i w_i N_i T_d/\tau}{m_m n V} \\ &\approx \frac{j\langle v_T \sigma\rangle T_d}{e} \approx \frac{qET_d}{m_m n V}, \end{aligned} \tag{9}$$

where $j$ is the current density, $v_T$ is the thermal velocity of an ion, $\sigma$ is the collision cross section of ions with atoms, $V$ is the gas volume, $n$ the density of gas atoms, and $E$ is the field strength. Substituting into expression (9) the quantities typical of our experiment leads to a

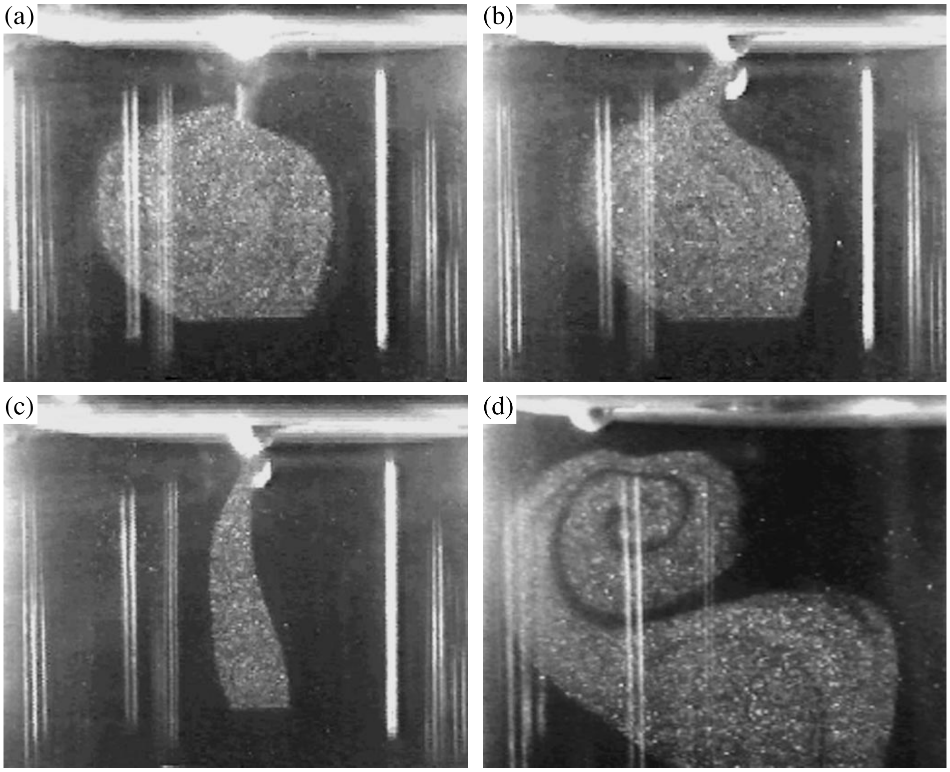

**Fig. 8.** (a–c) Removal of dust cloud of Zn particles using a spherical auxiliary electrode (Fig. 1, *B*). The auxiliary electrode diameter is 3 mm, the potential of the main electrode is +153 V, the potential of the auxiliary electrode is +273 V, the separation between the upper and lower electrodes is 4 cm; the neon pressure is $0.76 \times 10^5$ Pa, and the frame size is $4.2 \times 3.1$ cm. The lower part of the structure is not seen due to the finite size of the laser "knife". (d) Formation of a vortex from the cloud.

value of several centimeters per second for the momentum per unit mass of the gas.

Let us consider the same process on macroscopic level. Figure 9 shows the forces acting on ions during their motion through the gas. The external electric field exerts a force $\mathbf{F}_E$ on a drifting ion cloud. Since the drift velocity is constant, the gas exerts the resistive force $\mathbf{F}_C$ on the cloud. In turn, ions act on the gas with an equal and opposite force $\mathbf{F}$ which is precisely the driving force for the gas. The reason behind the emergence of this force on molecular level lies in the transfer of the additional momentum acquired by ions due to the field to molecules of the medium.

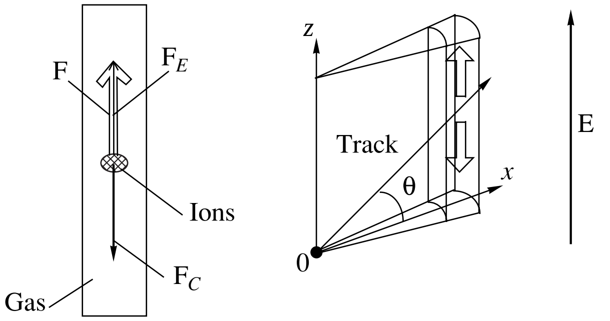

**Fig. 9.** Forces emerging during the motion of ions through the gas.

The transferred momentum of particles per unit mass, averaged over the time interval *T*, was calculated by the formula

$$\left\langle \frac{\Delta p}{\Delta m} \right\rangle = \left\{ \frac{1}{\Delta m T} \int F(t) dt \right\} \Delta t, \tag{10}$$

where $\Delta p$ is the momentum transferred to the gas of mass $\Delta m$ during the time $\Delta t$ under the action of the time-averaged force *F* on the gas. This force is caused by the interaction of the charge of the ion or electron cloud with an external electric field *E*. Under our experimental conditions, the time dependence of force *F* is of the pulsed type in view of the limited intensity of the source of ionizing particles:

$$F(t) = \begin{cases} Q_{e,i} E, & t \in [t_s, t_e], \\ 0, & t \notin [t_s, t_e]. \end{cases} \tag{11}$$

Here, $Q_{e,i}$ is the charge of an electron bunch or ion cluster and $t_s$ and $t_e$ are the initial and final time of its motion to the electrode. The averaging time *T* may be arbitrary, but it must correspond to a sufficiently large number of events. The time interval $\Delta t$ is equal to the characteristic time of variation of the vortex velocity; in our experiments, it was approximately equal to 1 s. In order to simplify the calculations (which are of approximate nature), we assume that the field is uniform.

Figure 10 shows dependence (11) schematically. By way of an example, three typical situations each are depicted for an inert gas and an electrically neutral gas (e.g., air). The first pulses (ionization events) occurring in the vicinity of the source produce large impulses of force. Their contribution to the value of $\Delta p/\Delta m$ is decisive. In the second case, when ionization occurs in the middle part of the electrode gap, impulses of force for positive and negative ions are approximately equal. In the third case, the impulse of force created by negative ions prevails. The contribution of electrons to $\Delta p/\Delta m$ is negligibly small since their drift velocity is approximately four orders of magnitude larger than the velocity of ions in all cases. Upon an increase in the distance from the radioactive source, the number of electron–ion pairs decreases, leading to a decrease in the amplitude of the impulses of force.

The time-averaged value of the momentum per unit mass transferred from ions to neutral components was calculated using the Monte Carlo method. The track nature of the processes was taken into consideration. Figure 9 shows that an ion cluster and an electron bunch drifting in opposite directions appear upon the passage of an ionizing particle through a gas in an electric field directed along the *z* axis. If we consider the entire track, these clusters have a cylindrical shape, the axis of the cylinder being parallel to the initial track. We took into account the contribution of fission fragments only. For a californium source, the contribution of 16α particles (the other half from 32 goes to the substrate) to the process under investigation is approximately equal to the contribution from a fission fragment; hence, the result is simply doubled. The cone angle in the direction of the *x* axis was chosen so small that the time interval between the tracks was larger than the time of ion drift to the electrodes. In this case, there is no need to process the effects associated with accidental coincidence of overlapping tracks. The exit angle θ is selected randomly. The result is averaged over the time corresponding to the passage of 2000 tracks. The length of the cell along the *z* axis is assumed to be equal to the mean free path of fission fragments, while the size of each cell along the *x* axis is equal to one hundredth of the mean free path of a fission fragment. We calculated the number of ions produced by a fragment in the cells crossed by it. The contributions of ions and electrons (or oppositely charged ions) in each unit cell cut with a step along the *x* axis are taken with opposite signs.

The calculations were made for neon under a pressure of $5 \times 10^4$ Pa with a Cf source of intensity $4 \times 10^6$ divisions/s, and the electric field strength was chosen equal to 100 V/cm (the values typical of the experiment). The field strength vector was directed downwards since the upper electrode in the experiments was maintained under a positive voltage. The drift velocities required for calculations were taken from [8–10]. The results are presented in Fig. 11 (curve *1*). Electrons are the carriers of negative charge in neon. If the field is antiparallel to the *z* axis, the ascending electrons cannot make a large contribution to the transferred momentum in view of their very short drift time. On the other hand, the ion contribution leads to the momentum transfer towards the source, which precisely affects the motion of dust particles. Since the energy losses for ionization along the track and the density of tracks decrease, the ions moving in the vicinity of the *z* axis and generated in the vicinity of the source make a larger contribution to the transferred momentum.

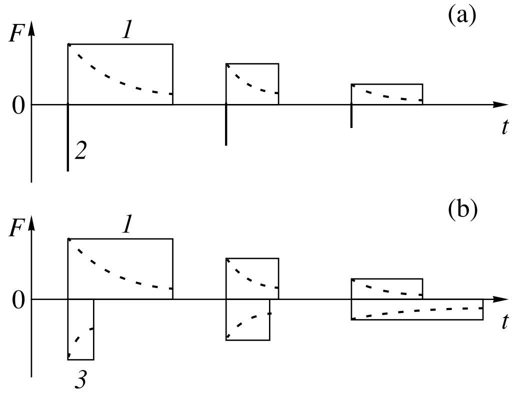

**Fig. 10.** Time dependence of force: pulses of positive ions (*1*), pulses of electrons (*2*), pulses of negative ions (*3*); dashed curves correspond to strong recombination on the path of pulses to the electrodes: (a) inert gas, (b) electronegative gas.

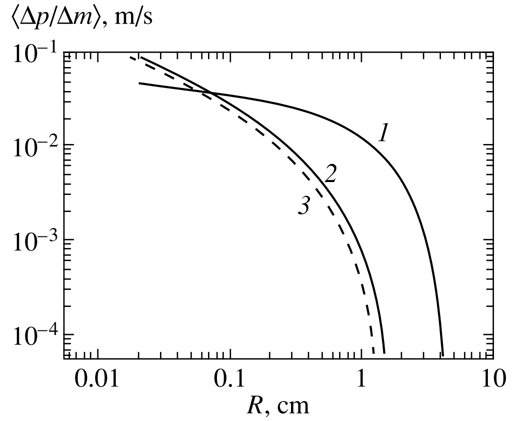

**Fig. 11.** Results of calculation of the transferred momentum.

In our previous publication [1], we studied the behavior of $CeO_2$ dust particles in the nuclear-induced plasma formed in atmospheric air by fission fragments and α particles of $^{252}$Cf with an intensity of $1 \times 10^5$ fissions/s. As the external electric field strength increases, levitation of particles is replaced by rotational motion,

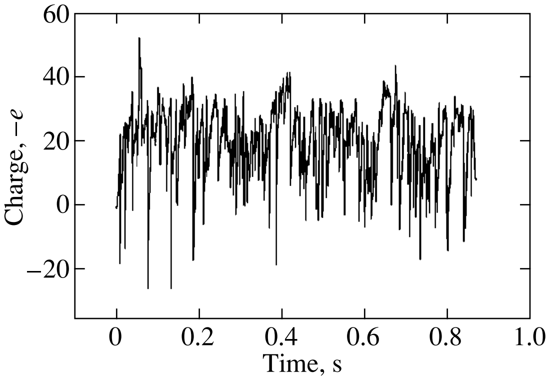

**Fig. 12.** Time dependence of the inverted charge of a dust particle at 1 cm from the source. The average charge is 20.6 units.

their trajectories under the steady-state conditions are closed curves, and the motion of the entire ensemble of particles in the plane of observation has the form of two stable vortex formations. The rotation motion was also observed for solitary particles. The type of motion of macroparticles did not change upon the sign reversal of the applied voltage and under rotations of the experimental cell. In all cases, the vortices moved so that particles near the axis of the setup had a velocity directed away from the radioactive source. Another feature of the experimental results [1] is that the dependence of the angular velocity for particles in a vortex on the electric field strength resembles the current–voltage characteristic.

Let us first consider the electric charge transfer to a dust particle of diameter 1 μm subjected in air to the action of drifting ions of different polarity. In analogy with [4], the time dependence of the charge was calculated by the Monte Carlo method for an electric field strength of 100 V/cm. The results of calculations are presented in Fig. 12. The dust particle was placed at a point with coordinate $z = 1$ cm (the middle part of the experimental cell in [1]). The $R$ coordinate of the cylindrical reference frame varied from 2 cm to 0. Although the charge increases with decreasing distance to the source, is still remains small (from 9 charge units at the periphery to 20 units at the axis). Charge fluctuations are caused by the random nature of interaction between nuclear particles and matter [4]. The magnitude of the average charge is too small to cause a fast motion. Earlier [2], the flow of the gas was not taken into account in determining the charge, which led to an exaggerated value of the charge.

If we assume that the main reason behind the vortex flow of particles is the motion of air, and that particles visualize this flow, the features of the experiments [1] carried out in air can be interpreted. We calculated the transferred momentum under these experimental conditions (air under atmospheric pressure and the electric field strength 100 V/cm). The results are presented in Fig. 11 (curve *2* corresponds to the upper electrode at a positive potential and curve *3*, at a negative potential). A decrease in energy losses for ionization along the track and a decrease in the density of tracks, as with neon, lead to a larger contribution to the transferred momentum of the ions moving in the vicinity of the $z$ axis. The number of ion pairs in each unit cell increases upon a decrease in the distance from the source. The ions moving to the upper electrode traverse a longer path in the gas and transfer a larger momentum. Negative oxygen ions $O_2^-$ are formed in air in a time equal to a fraction of a microsecond during the interaction with oxygen and having a mobility which differs insignificantly from that of the positive $N_2^+$ ions; the contribution to the momentum transfer of these ions upon a change in the direction of the field is almost the same. The momentum transfer in air is virtually independent of the direction of the electric field, which is precisely observed in experiments. This leads to the emergence of an air flow directed away from the source of ionizing particles for any polarity. Consequently, the direction of rotation of an electrically negative gas in a limited volume coincides with the normal to the surface of the source. A vortex flow of the gas in the space confined by walls and electrodes can be visualized with the help of dust particles.

The calculations were made for the ionizing chamber operating under the saturation of the current–voltage characteristic. For lower values of the electric field strength, recombination leads to charge losses in the track itself as well as on the path of ions to the electrodes, which is shown schematically by dashed curves in Fig. 10. As a result, the electric current and the transferred momentum decrease. The proposed model of the emergence of a vortex flow explains another important experimental result, viz., the saturation of the velocity of rotation corresponding to the plateau on the current–voltage characteristic. The reason behind these two effects is the same. Before attaining the plateau, both quantities increase since recombination in tracks decreases, while after the attainment of the plateau, all the ions produced by the source reach the electrodes. For this reason, the velocity of the vortex flow exhibits the same dependence on the applied voltage as the current–voltage characteristic. Indeed, in accordance with (9), the value of the transferred momentum is proportional to the drift time (inversely proportional to the drift velocity) and to the field strength. In turn, the drift velocity is directly proportional to the field. In the case of saturation, i.e., in the absence of charge losses, the dependence on the field strength vanishes:

$$\frac{\Delta p}{\Delta m} \approx \frac{qET_d}{m_m nV} = \frac{qE}{m_m nV}\frac{L}{w_i} = \frac{qL}{m_m nVK}$$

(here $L$ is the separation between the electrodes and $K$ is the ion mobility).

The dimensions and the physical meaning of the momentum per unit mass of the gas correspond to the velocity of this gas. However, it can be determined only by taking into account the force of viscous friction in the gas and by solving the hydrodynamic equations. The estimates obtained by us here prove that such an analysis must be carried out.

## 7. CONCLUSIONS

As a result of the experiments, we studied the collective motion of Zn dust particles in the plasma formed during the deceleration of the decay products of californium nuclei in neon. An electric field of various spatial configurations was created in the volume containing the gas medium, the source of ionizing radiations, and dust particles. We observed for the first time compact vortex structures with a large number of dust particles. Dust particles in these structures coagulate, after which the mass of particles increases almost by two orders of magnitude. We also observed dense dust clouds with sharp boundaries, which evolve in time. Particles in such clouds create ordered structures of the liquid type. Under steady- state conditions, dust structures of both types exist for tens of minutes. As the voltage at the high-voltage electrode increases, jets of dust particles are formed. Such jets can be directed in a desired direction by using an auxiliary electrode whose potential is higher than that of the main electrode.

It is shown that the momentum transfer from ions to the neutral component of the medium considerably affects the formation of the vortex motion of dust particles. The motion of ions is induced by the external electric field as well as by the intense recombination of ions and electrons on dust particles, which might cause the emergence of large concentration gradients and additional flows of plasma particles. Gas ionization and recombination (including that on dust particles) have the maximum intensity in the vicinity of the source of ionizing radiation. The ion flow causing a decrease in the diameter of the structures in the vicinity of the source is directed precisely at this region. The spread in the particle size, fluctuations of particle charges, and the hydrodynamic flow of the medium are considerable obstacles in the formation of crystal dust structures. The behavior of dust particles is determined to a considerable extent by the external field. Our calculations made it possible to explain different behavior of the vortex flow of dust particles in inert gases and in electrically negative gases. In the latter case, the momentum transferred by ions always induces vortex motion with a direction of rotation such that the particles lying in the vicinity of the experimental cell axis move away from the radioactive source.

Dust particles may considerably affect the processes occurring in plasmas. For this reason, the kinetic models of the nuclear-induced dust plasma should be developed, taking into account the recombination of plasma particles on dust particles.

## ACKNOWLEDGMENTS

The authors are grateful to A. I. Chusov for his help in the experiments and also to V. S. Filinov, A. P. Budnik, and V. A. Zherebtsov for fruitful discussion of the results.

This work was financed by the Russian Foundation for Basic Research (project no. 00-02-17620).

## REFERENCES

1. V. E. Fortov, A. P. Nefedov, V. M. Torchinsky, *et al.*, Phys. Lett. A **258**, 305 (1999).
2. V. E. Fortov, V. I. Vladimirov, L. V. Deputatova, *et al.*, Dokl. Akad. Nauk **366** (2), 184 (1999) [Dokl. Phys. **44**, 279 (1999)].
3. A. P. Budnik, Yu. V. Sokolov, and A. S. Vakulovskiy, Hyperfine Interact. **88**, 185 (1994).
4. V. A. Rykov, A. V. Khudyakov, V. I. Vladimirov, *et al.*, in *Proceedings of the XV International Conference Equations of State of Matter, Terskol, 2000*, p. 175.
5. O. M. Belotserkovskiĭ, I. E. Zakharov, A. P. Nefedov, *et al.*, Zh. Éksp. Teor. Fiz. **115**, 819 (1999) [JETP **88**, 449 (1999)].
6. I. A. Belov, A. S. Ivanov, D. A. Ivanov, *et al.*, Zh. Éksp. Teor. Fiz. **117**, 105 (2000) [JETP **90**, 93 (2000)].
7. I. A. Kvasnikov, *Thermodynamics and Statistical Physics. Theory of Equilibrium Systems* (Mosk. Gos. Univ., Moscow, 1991).
8. L. G. H. Huxley and R. W. Crompton, *The Diffusion and Drift of Electrons in Gases* (Wiley, New York, 1974; Mir, Moscow, 1976).
9. E. W. McDaniel and E. A. Mason, *The Mobility and Diffusion of Ions in Gases* (Wiley, New York, 1973; Mir, Moscow, 1976).
10. V. E. Golant, A. P. Zhilinskii, and I. E. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).

*Translated by N. Wadhwa*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 324–335.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 366–378.*
*Original Russian Text Copyright © 2001 by Vikharev, Gorbachev, Ivanov, Kolysko, Kuznetsov.*

**PLASMA, GASES**

# Spatial Structures of Continuous Microwave Discharge

**A. L. Vikharev*, A. M. Gorbachev, O. A. Ivanov, A. L. Kolysko†, and O. A. Kuznetsov**
*Institute of Applied Physics, Russian Academy of Sciences, Nizhni Novgorod, 603600 Russia*
**e-mail: val@appl.sci-nnov.ru*

Received April 28, 2000

**Abstract**—The paper deals with the results of investigations of spatial structures of continuous microwave discharge in a quasi-optical resonator. The results are given of experimental observations and measurements of the parameters of plasma in discharges of different forms, and the reasons are analyzed for the formation of spatial discharge structures. It is demonstrated that, as a result of the plasma-resonance amplification of the field, the discharge makes a transition to the contracted state with a size that is much less than the microwave-frequency wavelength and with an electron concentration in excess of the critical. It is found that the stratification of the contracted state across the electric field vector, which arises in some gases, is caused by the development of thermoelectric-current instability that was not previously observed in microwave discharges. *© 2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

The study of plasma instabilities has one of the central places in the physics of microwave gas discharge. Such a discharge demonstrates a wide variety of spatial structures forming as a result of the development of various instabilities. Fairly well studied at the present are the stratification of plasma across the field vector, which arises as a result of plasma-resonance instability [1] at low values of gas pressure, and the formation of thin plasma filaments extending along the electric field vector, which form at high values of pressure during the development of ionization-overheating instability [2]. In addition, in pulsed discharges, the field amplification on plasma inhomogeneities (that are initial or form under the effect of instabilities) may bring about the formation of a filamentary discharge structure (high-frequency streamers) [3]. The above-identified instabilities were observed experimentally in pulsed discharges of microsecond duration [4–6]. In continuous discharges, instabilities of a different type come to the fore, which are associated both with the kinetic processes in a discharge plasma [6–10] and with the singularities of interaction between the field and plasma at the nonlinear stage of the development of the instabilities identified above.

The urgency of the investigations of continuous microwave discharge is caused by the potentialities of its various practical uses. Such a discharge is widely used for gas laser pumping, in processes of dry etching in microelectronics, for diamond film deposition, and in various plasma-chemical reactors. The efficiency of operation of most of those devices is defined by the concentration of chemically active particles in the discharge plasma. In its turn, the rate of generation of active particles depends on the concentration of electrons in the discharge plasma and on the constants of kinetic processes, which are clearly defined functions of reduced electric field. Therefore, the efficiency of the processes is defined by the self-consistent evolution of the field and density of plasma in the discharge and may vary considerably in the process of instability development. In a number of cases, when a high degree of discharge homogeneity is required (gas lasers, etching processes), the instability development is an undesirable process, while in other cases the formation of structures with a high temperature and a high electron concentration brings about an increase in the efficiency of plasma-chemical processes.

This paper deals with the investigation of spatial structures arising as a result of the development of various instabilities in a continuous microwave discharge ignited in a quasi-optical resonator. The results of measurements of the discharge plasma parameters are given, as well as the description of the experimentally observed forms of discharge. A detailed analysis is performed of the reasons for the emergence of the contraction of discharge in electropositive and electronegative gases. It is demonstrated that the mechanism of contraction and the parameters of plasma in the contracted state are considerably affected by the quasistatic amplification of the electric field on plasma formations with a high concentration of electrons. As a result, a self-localized contracted form of discharge (autosoliton) arises in a system with integral negative feedback.

A new (from the standpoint of a microwave discharge) type of small-scale stratification of plasma across the electric field vector has been revealed, which is associated with the development of thermoelectric-current instability [11]. It is demonstrated that this instability develops in a contracted discharge under

---

† Deceased.

conditions of electron density exceeding the critical value and in gases characterized by a clearly defined dependence of the electron temperature on the electric field. The development of instability leads to the formation of a structure in the form of periodic fixed striations formed in active systems with diffusion, described by two diffusion equations.

## 2. EXPERIMENTAL SETUP AND DIAGNOSTIC TECHNIQUES

A schematic view of the experimental setup is given in Fig. 1. A detailed description of the experimental apparatus may also be found in [7]. An open quasi-optical resonator *3*, *4* was excited in the $TEM_{00q}$ mode by microwave radiation from a continuous-wave magnetron *1* with a wavelength $\lambda \approx 3$ cm and power $P_0 = 10$ to 200 W. The resonator was formed by round spherical mirrors (diameter of 24 cm, curvature radius of 18 cm) and had a Q factor without plasma of $Q_0 \approx 5 \times 10^3$. The experiment involved measurements, by a calorimeter *2*, of the incident power and of the power reflected from the resonator, with the power input into the resonator calculated as their difference. The distance between the mirrors was adjustable, which enabled one to vary the resonator eigenfrequency and, with a constant magnetron frequency, to adjust the power input into the resonator. The relative variations of the field amplitude in the resonator were registered by a stub antenna *13* located in the vicinity of one of the mirrors.

The electron concentration $N_e$ in the discharge was determined by the known procedure based on measuring the variation of the Q factor and resonance frequency of the resonator upon emergence of plasma in the latter [12]. The correlation between the electron concentration and the resonator parameters was found using the perturbation theory [12, 13], and the plasma dimensions required for calculations were determined using photographs of the discharge. The condition of validity of this procedure is the smallness of the electron concentration $N_e$ compared with the critical concentration $N_c$ and the smallness of the detuning of the resonator $\Delta f$ compared with the difference between the frequencies of two adjacent types of oscillation,

$$N_e \ll N_c, \quad N_c = \frac{m_e(\omega^2 + \nu^2)}{4\pi e^2} = \left(1 + \frac{\nu^2}{\omega^2}\right)N_{c0},$$
$$N_{c0} = \frac{m_e\omega^2}{4\pi e^2}, \quad \Delta f \ll |f_1 - f_2|, \tag{1}$$

where ω is the circular frequency of electromagnetic wave, ν is the electron–neutral collision frequency, and $f_1$ and $f_2$ are the resonance frequencies of the resonator.

Used as a diagnostic resonator for measurements of low electron concentrations (~$10^{10}$–$10^{11}$ cm$^{-3}$) was the same 3-cm resonator excited on a different frequency from a low-power diagnostic oscillator *9* via bandpass filter *12* in the $TEM_{00q\pm2}$ mode [7]. Higher concentrations (up to $10^{13}$ cm$^{-3}$) were measured using a diagnostic quasi-optical resonator *5* with spherical mirrors (diameter of 10 cm, curvature radius of 11 cm) located on either side of the axis of the 3-cm resonator. In this resonator, the fundamental mode $TEM_{00q}$ was excited on a frequency $f_r = 54$ GHz with a $Q$ factor of $7 \times 10^3$, with the electric field vector of the probing wave being perpendicular to that of the high-power wave. The resonance frequency shift was determined by a microwave spectrum analyzer *11*. The accuracy of measurements of the electron concentration by this method is not high, with the error arising largely because of the indeterminacy of the form of spatial distribution of plasma and reaching 50%.

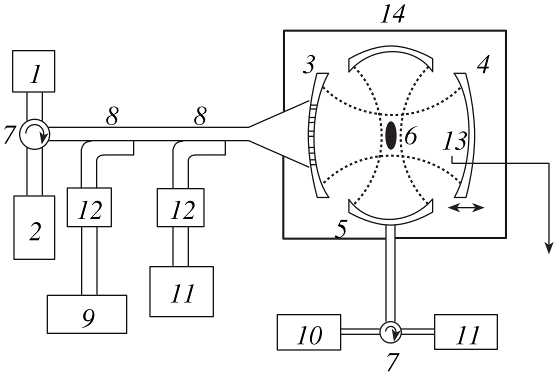

**Fig. 1.** Schematic of the experimental setup: (*1*) 3-cm magnetron, (*2*) calorimeter, (*3*) stationary semi-reflecting resonator mirror, (*4*) movable mirror, (*5*) 8-mm diagnostic open resonator, (*6*) discharge plasma, (*7*) circulator, (*8*) directional coupler, (*9*) 3-cm diagnostic oscillator, (*10*) 8-mm diagnostic oscillator, (*11*) spectrum analyzer, (*12*) bandpass microwave filter, (*13*) stub antenna, (*14*) vacuum chamber

## 3. EXPERIMENTALLY OBSERVED FORMS OF DISCHARGE

Typical forms of microwave discharge, that arise in various gases during variation of pressure and power input, are given in Fig. 2. Even the first investigations of a discharge in a resonator [8] revealed the presence of two fundamental forms of its existence, namely, diffuse and contracted. In the diffuse form (Figs. 2a, 2b, and 2c), the discharge was burning in one or several antinodes of the field, with the discharge size being close to that of the region of field localization and the boundary being smooth and diffuse. The number of antinodes, in which the discharge existed, was decreasing with increasing pressure. When some threshold pressure was reached during matching the resonator (by displacing the movable mirror) and increasing the microwave power input into the resonator, the discharge made a jumpwise transition to the contracted state, see Figs. 2d and 2e. In this form, the discharge was shaped as an ellipsoid slightly extended along the

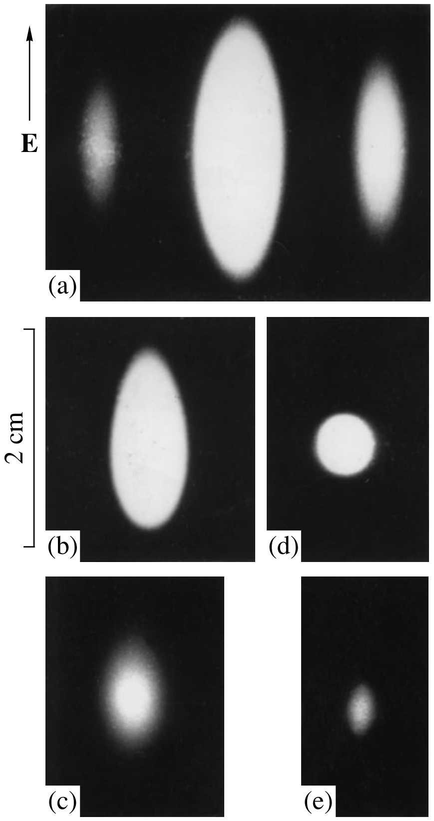

**Fig. 2.** Photographs of typical forms of continuous microwave discharge in an open resonator. Diffuse discharge (a) in air at $p$ = 7.5 torr, (b) in air at $p$ = 43 torr, (c) in oxygen at $p$ = 11 torr; contracted discharge (d) in air at $p$ = 64 torr, (e) in oxygen at $p$ = 11 torr.

electric field vector and having a size much less than that of diffuse discharge and the electromagnetic wavelength. The discharge transition to the contracted state was accompanied by a 10–20% reduction of the field in the resonator.

The threshold pressure, below which no contraction occurred at any values of the power input to the resonator, was substantially different for different gases. In highly electronegative gases (with a high rate of attachment), such as ammonia, carbon dioxide, and water vapors, the threshold pressure amounted to $p^* = 3$ to 8 torr, while in air and nitrogen it amounted to $p^* \approx$ 50 torr. It was found in [8] that the threshold pressure of the diffuse-to-contracted transition in an $N_2$:$O_2$ mixture decreased considerably with increasing fraction of the electronegative component (oxygen).

In some electronegative gases, autooscillation occurred of the amplitude of the field sustaining the discharge and of the electron concentration, that was associated with the development of attachment instability and was investigated in detail in [9].

It was noted in [8, 10] that, in the case of contraction in the region of high values of threshold pressure, the gas temperature increased by 300–400 K to reach 1500–2000 K. In a mixture with a high content of electronegative component, the gas temperature increment did not exceed 100 K.

It is known [14, 15] that the transition of a discharge into the contracted form is possible when two necessary conditions are valid, namely, the existence of the nonlinear dependence of the ionization rate on the electron concentration and the presence of volume neutralization of charged particles. It is pointed out in [16] that, in an electronegative gas, the channel of acceleration of the process of production of electrons may be provided by the decay of negative ions and the emergence of new electrons during detachment. Theoretical studies were made into discharge structures in a highly electronegative gas, arising in a homogeneous field, with due regard for the detachment processes and the effect of the dependence of the electron diffusion coefficient on the concentration of negative ions. It has been demonstrated that, for spatially inhomogeneous distributions of the electron concentration, an increase in the electron concentration in the central region of the discharge brings about a decrease in the density of negative ions, as a result of which the electron diffusion coefficient decreases. On the contrary, at the periphery, the electron diffusion is high as a result of considerable prevalence of negative ions over electrons. The combined action of these effects brings about a decrease in the characteristic scale of distribution of electrons with an increase in their concentration in the central region. However, this model did not include the field variation in the plasma in the case of fairly high (comparable to critical) values of electron concentration, that may have a considerable effect on the process of discharge contraction.

### *3.1. Discharge Contraction*

In order to find the mechanisms of contraction of a microwave discharge, one must know the plasma parameters in the vicinity of the contraction threshold and in the contracted state. Figure 3 gives the concentration of electrons in diffuse and contracted discharges in carbon dioxide as a function of power input to the resonator. An analogous dependence for oxygen was given previously in [6]. One can see that this dependence exhibits an ambiguous hysteretic behavior. The lower branch corresponds to the diffuse discharge, and the upper branch, to the contracted discharge. With an increase in the power input to the resonator, the electron concentration in the diffuse form increases, with the plasmoid size remaining unchanged; when some threshold value $N_e^*$ is reached, the discharge makes a jumpwise transition to the contracted state, with the power input somewhat decreasing and the field in the resonator dropping (by approximately 20%). In the contracted state, a variation of the power input to the resonator does not lead to a variation of the electron concentration, but brings about a corresponding increase or decrease in the size of the plasmoid whose volume is a linear function of the power input. When the size of the contracted form of discharge becomes small (of the order of the diffusion length of electron loss), the reverse, contracted-to-diffuse, transition occurs.

One can see in Fig. 3 that there exists a range of power input values, in which the existence of two forms of discharge is possible. In this range, a strong perturbation of the electron concentration is capable of changing the discharge from the diffuse to contracted form. In the experiment, this was accomplished with the aid of a pulsed breakdown between the ends of two metal wires introduced inside a diffuse discharge normally to the electric field vector.

The electron concentration in discharges of both forms was measured for different gases while varying the pressure and power input to the resonator. Some results are given in the table for different states of discharge along with other parameters that characterize the discharge (density $N$ of neutral molecules, gas pressure $p$, plasma volume $V_{pl}$, power input to the resonator $P_{in}$).

We will analyze the obtained data. First, note the closeness of the parameters in the contracted state in different gases. The average electron concentration in this form is approximately the same and exceeds the critical value given by Eq. (1),

$$N_e \approx (1\text{–}3)N_c.$$

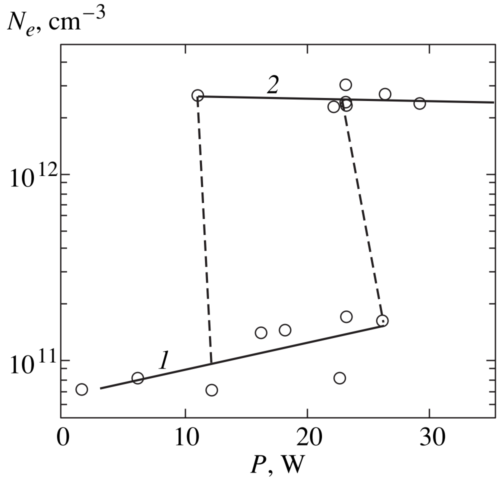

**Fig. 3.** The electron concentration as a function of the power input to the resonator for a discharge in $CO_2$, $p$ = 15 torr: (*1*) diffuse discharge, (*2*) contracted discharge. The broken lines indicate the transition from the diffuse to contracted discharge and back.

The electron concentrations at the contraction threshold in different gases likewise prove to be close in magnitude,

$$N_{eth} \sim (0.1\text{–}0.5)N_c.$$

**Table**

| Form | Gas | $p$, torr | $N_e$, cm$^{-3}$ | $N$, $10^{17}$ cm$^{-3}$ | $V_{pl}$, cm$^3$ | $P_{in}$, W |
|---|---|---|---|---|---|---|
| Diffuse | $NH_3$ | 8 | $1.3 \times 10^{10}$ | 2.6 | – | 10.5 |
| | $CO_2$ | 8 | $2 \times 10^{10}$ | 2.6 | 0.65$^{3)}$ | 1 |
| | $CO_2$ | 8 | $8 \times 10^{10}$ | 2.6 | 0.90 | 4.6 |
| | $O_2$ | 5 | $3.6 \times 10^{10}$ | 1.7 | 4.7 | 23 |
| | $O_2$ | 5 | $(1.2\text{–}4.3)^{1)} \times 10^{11}$ | 1.7 | 1.3 | 18.5 |
| | $O_2$ | 20 | $(0.5\text{–}1.1) \times 10^{11}$ | 6.6 | 3.0 | 28 |
| | $N_2$ | 5 | $2.6 \times 10^{10}$ | 1.7 | 6.8 | 14 |
| | $N_2$ | 15 | $4.5 \times 10^{10}$ | 3$^{2)}$ | 3.7 | 12 |
| | $N_2$ | 15 | $8.8 \times 10^{10}$ | 3$^{2)}$ | 7.3 | 30 |
| | $N_2$ | 40 | $2 \times 10^{11}$ | 4.5$^{2)}$ | 2.9 | 38 |
| Diffuse on the contraction threshold | $CO_2$ | 20 | $2 \times 10^{11}$ | 6.6 | 1.3 | 19 |
| | $O_2$ | 10 | $(1.6\text{–}8) \times 10^{11}$ | 3.3 | 0.62 | 17 |
| | $N_2$ | 90 | $3.3 \times 10^{11}$ | 8$^{2)}$ | 1.1 | 24 |
| Contracted | $NH_3$ | 8 | $1.5 \times 10^{12}$ | 2.6 | – | 10.5 |
| | $NH_3$ | 8 | $2 \times 10^{12}$ | 2.6 | – | 45 |
| | $CO_2$ | 20 | $3.4 \times 10^{12}$ | 6.6 | 0.053 | 16 |
| | $O_2$ | 10 | $(1.8\text{–}4.5) \times 10^{12}$ | 3.3 | 0.075 | 18 |
| | $O_2$ | 20 | $(1.5\text{–}4) \times 10^{12}$ | 6.6 | 0.075 | 25 |
| | $N_2$ | 90 | $2.5 \times 10^{12}$ | 6$^{2)}$ | 0.05 | 20 |

Note: 1) Two values of $N_e$, separated by a dash, correspond to the minimum and maximum detuning of the frequency of diagnostic resonator in the autooscillation mode of sustaining the discharge [9]. 2) The gas density was calculated proceeding from the pressure and from the results of measurements of the gas temperature [10]. 3) The plasma volume was determined by the size of the luminosity region and used to calculate the electron concentration.

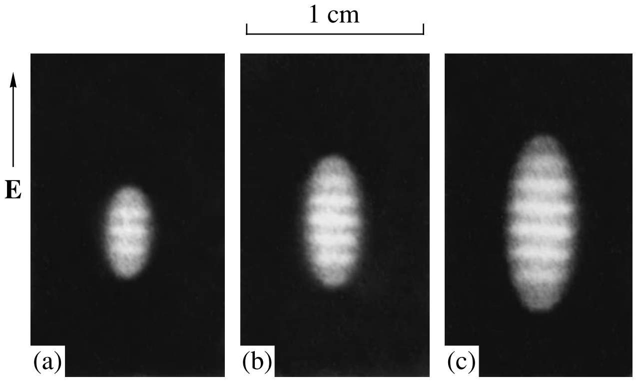

**Fig. 4.** Photographs of the stratified form of contracted discharge in $NH_3$, $p = 6$ torr, for different values of microwave power input to the resonator: (a) 6, (b) 13, (c) 21 W.

In spite of the different values of pressure, the gas density in the contracted form and at the contraction threshold and in electronegative gases, as well as in electropositive ones (as a result of heating), is such that the electron–neutral collision rate is insignificant ($\nu/\omega < 2$). That is, the imaginary part of plasma permittivity

$$\varepsilon = 1 - \frac{N_e}{N_c}\left(1 + i\frac{\nu}{\omega}\right) \quad (2)$$

is small, and effects associated with plasma resonance may show up.

A marked difference is observed between electronegative and electropositive gases in the diffuse form of discharge. For comparable values of pressure and power input, the discharge plasma in an electropositive gas takes a much larger volume than in an electronegative gas (see table and Fig. 2). As a result, the electron concentration in an electropositive gas is lower. An increase in the power input to the resonator has different effects on different gases in the low-pressure region. In an electropositive gas, this brings about a considerable increase in the plasma volume and a slight increase in the electron concentration; in an electronegative gas, on the contrary, this causes mainly an increase in the electron concentration with an insignificant increase in volume.

Therefore, high values of the electron concentration in the vicinity of the contraction threshold and in the contracted form bring about the considerable effect of the discharge plasma on the magnitude and distribution of the electric field. Therefore, in constructing an adequate model of contraction, this phenomenon must be included. At the same time, the observed difference in the behavior of the discharge between electronegative and electropositive gases is indicative of the importance of the kinetic processes.

### *3.2. Stratification of Contracted Discharge*

In addition to the contracted and diffuse forms, yet another unusual form, that of contracted discharge stratified along the electric field vector (Fig. 4), was observed in some gases such as $NH_3$, $H_2O$, and $CO_2$ in the pressure range $p = 3$ to 20 torr. The discharge had the form of an ellipsoid slightly extended along the electric field vector and consisting of alternating light- and dark-colored disks. No such stratification was observed in other investigated gases (air, $N_2$, $O_2$, He).

The distance between the disks (stratification scale) was always in the range of 1–2 mm. As the power input to the resonator was increased, the discharge size increased, as well as the distance between the disks, with the number of disks remaining the same. At a certain moment, the number of disks increased by one, and the stratification scale decreased jumpwise (see Fig. 4).

A stratified contracted discharge as a whole is characterized by the same regularities as a regular contracted discharge. The mean concentration of electrons in a stratified contracted discharge exceeds the critical value as it does in an unstratified discharge.

The emergence of stratification depends substantially on the type of gas. An addition of several percent of nitrogen to ammonia resulted in suppression of stratification. The stratification disappeared also in the case of sustaining a discharge for a long time in ammonia, which decomposed as a result of dissociation.[1] These facts point to the kinetic pattern of the revealed instability. At the same time, this stratification was only observed in a plasma with a higher-than-critical electron density.

## 4. DISCUSSION OF THE RESULTS

### *4.1. Mechanism of Discharge Contraction*

As follows from the foregoing experimental data, the process of contraction occurs upon exceeding a certain value of electron concentration in a diffuse discharge, which depends little on the type and pressure of gas. In the contracted form, the electron concentration also proves to be close in magnitude for different experimental conditions. This leads one to conclude that the mechanism of contraction in electropositive and electronegative gases is of one and the same nature. At the same time, the type of gas has a fundamental effect on the process of reaching the threshold electron concentration required for transition from one state to another.

The contracted form of discharge is represented by an ellipsoid extended slightly along the vector **E**, whose size is much less than the electromagnetic wavelength. In this case, the electric field may be found approxi-

---

[1] The increase with time of the pressure in the vacuum chamber was indicative of the dissociation of ammonia under the effect of the discharge.

mately from the solution of the electrostatic problem on dielectric ellipsoid with a clearly defined boundary in homogeneous external field (see, for example, [17]),

$$E_i^k = \frac{E_r}{|1+(\varepsilon-1)n_*|}, \qquad (3)$$

where $E_i^k$ is the value of the field within the contracted form of discharge, $n_*$ is the depolarization coefficient dependent on the ellipsoid semiaxes, and $E_r$ is the amplitude of external field, i.e., of the field in the resonator in the region of plasmoid. For the complex permittivity given by Eq. (2), this formula may be represented by

$$E_i^k = \frac{E_r}{\left[\left(1-\frac{n_* N_e}{N_c}\right)^2 + \left(\frac{\nu}{\omega}\frac{n_* N_e}{N_c}\right)^2\right]^{1/2}}. \qquad (4)$$

The value of $E_r$ depends on the incident microwave power, $Q$ factor, and the value of detuning of the resonator frequency with plasma. Note that it was a signal proportional to the quantity $E_r$ that was recorded in the experiment from the stub antenna (see Fig. 1).

For an ellipsoid corresponding to the contracted form of discharge, $n_* \le 1/3$ [17]. In the region of low values of gas density, where $\nu \le \omega$, the correlation $E_i^k(N_e)$ given by Eq. (4) exhibits a resonant behavior. In this case, the field inside the plasmoid at values of the electron concentration of less than

$$N_e^* \approx \frac{N_c}{n_*} \sim (3–4)N_c \approx (3–4)\times 10^{12}\ \text{cm}^{-3}$$

increases with $N_e$. It is in the contracted form of discharge that this effect is most pronounced. The diffusion of plasma boundaries, characteristic of real plasmoid, brings about an increase in the internal loss and causes attenuation of resonance; however, it does not suppress the resonance completely in the case when the boundary width is much less than the plasmoid size [18]. In a diffuse discharge ($N_e \ll N_c$), the value of the field is close to that of undisturbed field in the resonator, $E_i^d \approx E_r$.

Therefore, it is natural to assume that the contracted state is maintained owing to a quasistatic amplification of the field. This, in particular, explains the fact that the transition from the diffuse to contracted form is accompanied by a reduction of the field in the resonator ($E_r$ in formula (4)). Indeed, during transition to the contracted state, in spite of the decrease in the external field, the value of the field in the plasma of contracted discharge increases, whereby a high electron concentration is maintained. At the same time, a reduction of the field outside of contracted discharge brings about a considerable decrease in the electron concentration in this region.

A more detailed analysis of spatial distribution of the field and electron concentration in a contracted discharge was performed using a numerical model. For this purpose, the stationary distributions of the electron concentration and of the electric field were calculated within the framework of quasistatic equations and balance equations for particles.

The electric field distribution on scales of much less than the electromagnetic wavelength was calculated from Poisson's equation with complex permittivity $\varepsilon$ given by Eq. (2),

$$\nabla(\varepsilon\nabla\phi) = 0, \quad \mathbf{E} = -\nabla\phi. \qquad (5)$$

Equation (5) was solved, proceeding from the symmetry of the problem, in cylindrical coordinates ($\rho$, $z$), with the $z$ axis directed along the external electric field vector $\mathbf{E}_r$. The boundary conditions were preassigned at the boundary of the integration region exceeding considerably the size of plasmoid and corresponded to a uniform unperturbed field.

The stationary distribution of electron concentration was found from balance equations for particles analogous to those used in [9],

$$\begin{aligned}
\frac{\partial N_e}{\partial t} &= D_a\Delta N_e + (\nu_i - \nu_a)N_e \\
&\quad + k_d N^* N_- - \alpha_{ei}N_e N_+, \\
\frac{\partial N_-}{\partial t} &= \nu_a N_e - k_d N^* N_- - \beta_{ii}N_- N_+, \\
\frac{\partial N^*}{\partial t} &= k^* N_0 N_e - k_d N^* N_- - \frac{N^*}{\tau^*}, \\
N_+ &= N_- + N_e.
\end{aligned} \qquad (6)$$

Here, $N_e$, $N_-$, $N_+$, $N^*$, and $N_0$ denote the concentration of electrons, negative and positive ions, excited particles, and neutral molecules, respectively; $D_a$ is the coefficient of ambipolar diffusion for electrons; $\alpha_{ei}$ and $\beta_{ii}$ are the coefficients of electron–ion and ion–ion recombination, respectively; $\nu_i$ and $\nu_a$ denote the rates of ionization by electron impact and of dissociative attachment of electrons, respectively; $k_d$ and $k^*$ are the rate constants of reactions of detachment and formation of excited particles, respectively; and $\tau^*$ is the characteristic relaxation time of $N^*$. In the stationary case, one can assume that the density $N^*$ of excited particles is proportional to $N_e$, and then the term describing the detachment of electrons from negative ions in reactions with excited particles may be written in the form $k_1 N_e N_-$ ($k_1 \approx k^* k_d \tau^* N_0$), thereby eliminating the third one of Eqs. (6).

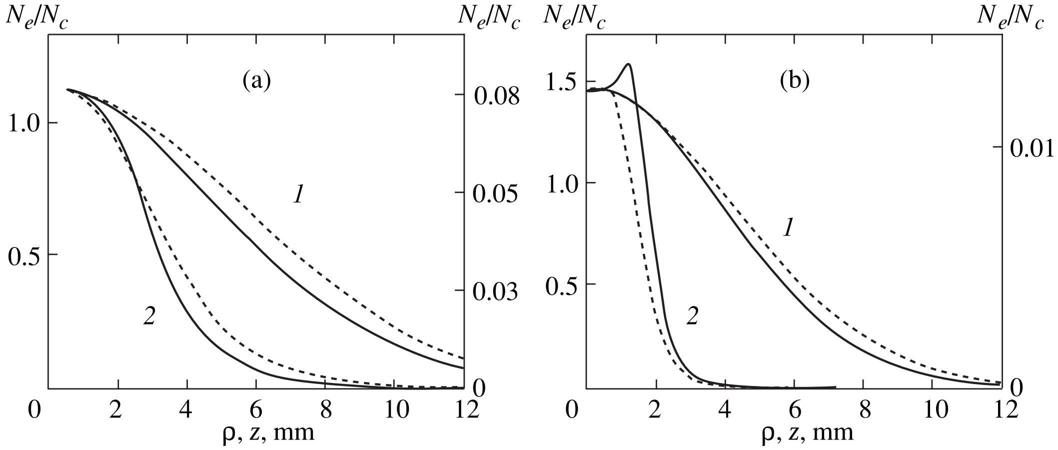

**Fig. 5.** Stationary distribution of the electron concentration in the (*1*) diffuse and (*2*) contracted forms of discharge. The solid curve indicates the distribution along the *z* axis at $\rho = 0$, and the broken line indicates the distribution on the radius ρ at $z = 0$: (a) in air ($A = 1$) for different values of the field in the resonator (*1*) $E_{r0}/E_c = 2.51$ and $E_r/E_c = 0.95$, (*2*) $E_{r0}/E_c = 4.78$ and $E_r/E_c = 0.75$; (b) in a strongly electronegative gas ($A = 10$): (*1*) $E_{r0}/E_c = 1.64$ and $E_r/E_c = 1.57$, (*2*) $E_{r0}/E_c = 1.72$ and $E_r/E_c = 1.13$. (Here, $E_{r0}$ is an electric field in the resonator without plasma, and $E_r$ is a self-consistent field in the resonator with plasma for the selected value of $E_{r0}$).

The ionization and attachment rates may be conveniently written in the model form, disregarding the concrete type of gas (in so doing, we will treat two limiting cases, namely, those of weakly and strongly electronegative gas, that differ only by the value of the attachment rate),

$$\nu_i = \nu_{a0} f(\rho, z)\left(\frac{|E(\rho, z)|}{E_c}\right)^n, \qquad (7)$$

$$\nu_a = \nu_{a0} A, \quad n \gg 1,$$

where $\nu_{a0}$ and $E_c$ denote the attachment rate and the breakdown field ($\nu_i(E_c) = \nu_{a0}$) in a weakly electronegative gas, respectively; and *A* is a coefficient that characterizes the degree to which the gas is electronegative. It is assumed that $A = 1$ for the weakly electronegative gas and $A = 10$ for the strongly electronegative gas. In calculating the frequency, the coefficients appearing in Eqs. (6) and (7) for a weakly electronegative gas approximately corresponded to those for air.

The function $f(\rho, z)$ describes the spatial inhomogeneity of the ionization rate in a uniform unperturbed electromagnetic field, that may be caused, for example, by the inhomogeneity of the gas density. It is assumed that $f = 1$ at the origin and $f \ll 1$ at the boundary of the integration region. This method of localization of the discharge region is often used in numerical calculations of the discharge dynamics in an initially uniform field [5, 19].

The magnitude of the field $E_r$ in the resonator (at the boundary of the integration region) for the known form of distribution of $N_e(\rho, z)$ and $E(\rho, z)$ was defined by the variation of the Q factor and eigenfrequency of the resonator, which were calculated by the formulas of the perturbation theory given in [12, 13].

The set of equations (6) and (5) was solved by simple iterations; in the calculations, the magnitude of the initial electric field $E_{r0}$ in the resonator without plasma and the parameter *A* (attachment frequency $\nu_a$) were varied. The results of calculation of the stationary structures of discharge for the gas density corresponding to the ratio $\nu/\omega = 0.5$ and for $N_c = 10^{12}$ cm$^{-3}$ ($\lambda \approx 3$ cm) are given in Figs. 5–7. One can see in the figures that, depending on the magnitude of the initial field, two forms of discharge may exist in the resonator, which differ substantially from one another by the characteristic width of the spatial distribution of electron concentration. The size of diffuse discharge is defined by the spatial localization of the ionization rate given by Eq. (7) (by the form of the function $f(\rho, z)$). In the contracted form, the distribution of electron concentration is formed as a result of quasistatic amplification of the field, and the plasma takes up a much smaller volume. An increase in the initial field in the contracted form brings about an increase in the plasma volume with almost unchanged values of the maximum concentration and of the electric field (see Fig. 6).

On analyzing the distribution of the field and electron concentration (see Fig. 7), one can understand the reason why no elongation of the plasmoid (characteristic of microwave streamers [3]) occurs along the electric field vector. Indeed, a considerable field increase is observed at the ellipsoid poles, especially, in the $N_e = N_c$ region, compared with the external field and with the field in the internal region of the plasmoid. However, in the region of the higher field, a region is observed, in which the field is much weaker and the plasma is

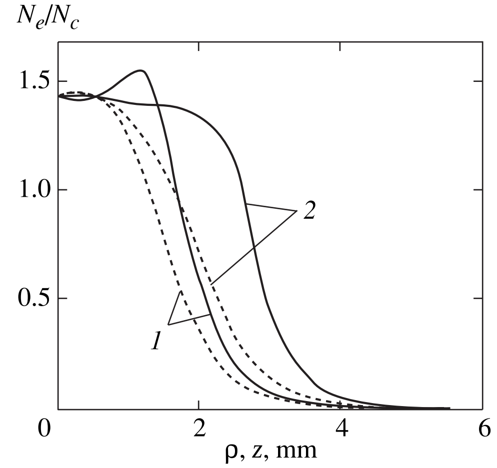

**Fig. 6.** Stationary distribution of the electron concentration in the contracted form of discharge for $A = 10$ and different values of the initial field $E_{r0}$: (*1*) $E_{r0}/E_c = 1.72$ and $E_r/E_c = 1.13$, (*2*) $E_{r0}/E_c = 2.51$ and $E_r/E_c = 1.13$. The solid curve indicates the distribution along the $z$ axis at $\rho = 0$, and the dashed curve indicates the distribution along the radius $\rho$ at $z = 0$.

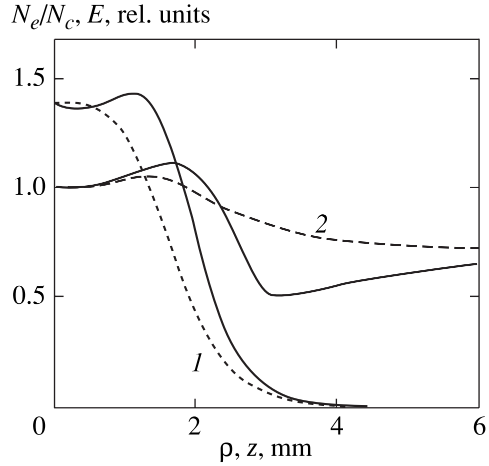

**Fig. 7.** Stationary distribution of (*1*) the electron concentration and (*2*) the modulus of electric field in the contracted form of discharge ($A = 10$, $E_{r0}/E_c = 1.82$). The solid curve indicates the distribution along the $z$ axis at $\rho = 0$, and the dashed curve indicates the distribution along the radius $\rho$ at $z = 0$.

absent, and the electrons produced in the plasma resonance zone diffuse to the region of lower field to be lost as a result of recombination and attachment. This fact prevents the discharge from propagating along the electric field vector. The presence of the region of lower field is associated with the continuity of the normal component of the electric induction vector **D** and with the change of sign of the real part of permittivity given by Eq. (2) during transition through the plasmoid boundary.

One can see in Fig. 5 that a much lower initial field is required for contraction in a strongly electronegative gas. Indeed, the width of the plasmoid boundary, at which the electron concentration decreases, depends on the diffusion length of electron loss $l_d \approx \sqrt{D_a/\nu_l}$, where $\nu_l$ is the rate of electron loss. In a strongly electronegative gas, the rate of loss is high, and the plasmoid boundary is much thinner than in an electropositive gas (compare Fig. 5a with Fig. 5b and Fig. 6). For example, for a gas density $N \approx 6 \times 10^{17}$ cm$^{-3}$, the attachment rate $\nu_a \sim 10^6$ s$^{-1}$ in oxygen exceeds considerably the recombination rate $\alpha_{ei}N_e$ in nitrogen. As was demonstrated by Gil'denburg *et al.* [18], an increase in the thickness of the plasmoid boundary brings about an attenuation of resonance and of the degree of field amplification in a plasmoid. However, at high pressures, the heating of gas causes a decrease in its density and, accordingly, a rise of the ionization rate in the region of plasma existence. This facilitates the maintenance of the contracted state in an electropositive gas at high pressures.

The experimental results demonstrate (see table) that the threshold concentration in different gases is reached under different conditions. At low pressures, a discharge in electronegative gas takes up a much smaller volume that varies little as the power input is increased. Apparently, this behavior is associated with the effect of the nonlinear diffusion mechanism described above [16] and with the difficulty of breakdown in other antinodes because of the high breakdown field. As a result, an increase in the power input to the resonator leads to a rise of the electron concentration. No clearly defined breakdown threshold is present in electropositive gases; therefore, an increase in the power input causes the emergence of plasma in new regions and an increase in the discharge volume with an insignificant rise of the electron concentration. At high pressures, the heating of gas and decrease in its density lead to localization of the region of plasma existence and make it possible to raise the electron concentration to a value required for transition to the contracted state.

Therefore, the formation and maintenance of the contracted form of microwave discharge may be attributed to the plasma-resonance amplification of the field on small-scale plasma formations.

### *4.2. Small-Scale Stratification of Contracted Discharge*

A small-scale stratification of contracted discharge is observed only in some gases and under conditions of the electron concentration $N_e$ exceeding the critical value $N_c$. According to the quasistatic equations, at $N_e > N_c$, a local increase in $N_e$ in a layer perpendicular to the vector **E** must bring about a field decrease in this region and *vice versa*. At the same time, a field increase is accompanied by a rise of the electron temperature $T_e$ and, in a stratified discharge, the quantities $T_e$ and $N_e$ change in antiphase. The diffusion equation for electrons with due regard for spatial inhomogeneity of distribution of the electron temperature may be represented as

$$\frac{\partial N_e}{\partial t} + \mathrm{div}(\mathbf{u}_a N_e) = F(N_e, T_e, \ldots), \tag{8}$$

where

$$\mathbf{u}_a = -D_a\frac{\nabla N_e}{N_e} - D_a^T\frac{\nabla T_e}{T_e} \tag{9}$$

is the ambipolar electron velocity in view of thermodiffusion [20], and $F(N_e, T_e, \ldots)$ is the term describing the processes of production and loss of electrons. The coefficients of ambipolar diffusion $D_a$ and of thermodiffusion $D_a^T$, disregarding the importance of negative ions in the case of highly nonequilibrium plasma, when $T_e \gg T_i$, are[2] [20]

$$\begin{aligned} D_a &\approx D_i\left(1 + \frac{T_e}{T_i}\right), \\ D_a^T &\approx D_i\frac{T_e}{T_i}(1-g) \approx D_a(1-g), \end{aligned} \tag{10}$$

where $D_i$ is the coefficient of eigendiffusion for positive ions, and

$$g = \frac{\partial \ln \nu}{\partial \ln T_e}$$

is the coefficient allowing for the effect of thermal force [20]. Note that these expressions for diffusion coefficients include the characteristic electron temperature $T_e = D_e/\mu_e$.

In the one-dimensional case corresponding to the stratification across the vector **E**, Eq. (8) in view of Eqs. (9) and (10) may be written as

$$\frac{\partial N_e}{\partial t} = \frac{\partial}{\partial x}\left(D_a\frac{\partial N_e}{\partial x} + D_a^T\frac{N_e}{T_e}\frac{\partial T_e}{\partial x}\right) + F. \tag{11}$$

In this case, the correlation between the electric field and electron concentration is algebraic,

$$\begin{aligned} E &= E(N_e) = \frac{E_0}{|\varepsilon|} \\ &= E_0\left[\left(1 - \frac{N_e}{N_c}\right)^2 + \left(\frac{\nu}{\omega}\frac{N_e}{N_c}\right)^2\right]^{-1/2}, \end{aligned} \tag{12}$$

where $E_0$ is the field in the plasmoid in the absence of stratification.

If the scale being treated is not too small (exceeds the electron thermal conductivity), the correlation between the electron temperature and the field, $T_e = T_e(E(N_e))$, may likewise be regarded as local. In this case, Eq. (11) may be transformed to the ordinary form

$$\frac{\partial N_e}{\partial t} = \frac{\partial}{\partial x}D_{\mathrm{eff}}\frac{\partial N_e}{\partial x} + F \tag{13}$$

with effective diffusion

$$\begin{aligned} D_{\mathrm{eff}} &= D_a\left[1 + (1-g)\frac{\partial \ln T_e}{\partial \ln N_e}\right] \\ &= D_a\left[1 + (1-g)\frac{\partial \ln T_e}{\partial \ln E}\frac{\partial \ln E}{\partial \ln N_e}\right]. \end{aligned} \tag{14}$$

For the correlation $E(N_e)$ of the form of (12) with the plasma parameters corresponding to the contracted state, the relation

$$\frac{\partial \ln E}{\partial \ln N_e} = -(1\text{–}2)$$

is usually valid. It follows from Eq. (14) that, with a fairly sharp field dependence of $T_e$, such that

$$\frac{\partial \ln T_e}{\partial \ln E} > -\left[(1-g)\frac{\partial \ln E}{\partial \ln N_e}\right]^{-1}, \tag{15}$$

the coefficient of effective diffusion becomes negative. In this case, an instability may develop, whose physical mechanism is as follows. In the case of a higher-than-critical electron concentration in the plasmoid, a decrease in $N_e$ in a layer perpendicular to the vector **E** causes a field increase. In the case of a clearly defined $T_e(E)$ dependence, a considerable heating of electrons occurs, and the thermodiffusion flow of electrons from this region, associated with the inhomogeneity of $T_e$, exceeds the return flow caused by the inhomogeneity of the electron density. In this manner, $N_e$ continues to decrease.

For a glow discharge, such an instability, referred to as thermoelectric-current instability, was first predicted by Timofeev [21], who obtained an instability criterion analogous to inequality (15). Experimental and theoretical investigations of the thermoelectric-current instability in a glow discharge are dealt with in [22–25]. Note that a decrease in the electron density caused by thermodiffusion is also observed when powerful electromagnetic radiation acts on an ionospheric plasma [26].

For an instability to arise in our conditions, it is necessary that the electron temperature increase rapidly with the field given by Eq. (15). A comparison of the dependences of the electron temperature $T_e = D_e/\mu_e$ on the parameter $E/N$ for different gases (Fig. 8), such as those given in [27–29], leads one to the following con-

---

[2] In the expression for the coefficient of ambipolar diffusion, the effect of negative ions may be ignored, because, at high electron concentrations characteristic of the contracted state, the number $N^*$ of excited particles is high, attachment is compensated by detachment, and the concentration of negative ions is low [7].

clusion. For gases such as nitrogen, air, oxygen, and helium, in which no stratification of the contracted form is observed, the dependence $T_e(E)$ is smooth,

$$\frac{\partial \ln T_e}{\partial \ln E} < \frac{1}{2},$$

in the entire range of discharge maintenance fields. On the contrary, for ammonia, water vapors, and carbon dioxide, the $T_e(E)$ curve exhibits a fairly steep region in which condition (15) may be valid.

Note that the clearly defined dependence $T_e(E)$ [29] is observed for electric field values of less than the breakdown value $E_c$ determined from the condition

$$\nu_i(E_c) = \nu_a(E_c).$$

Therefore, the instability described may be observed only in quasistationary discharges, when, as a result of the accumulation of active particles, the loss of electrons due to attachment is compensated by detachment, and the discharge maintenance field becomes markedly less than the breakdown one [7]. In short-pulse discharges, the field is fairly high, the $T_e(E)$ dependence is weak, and the thermodiffusion flow introduces an insignificant addition to the total electron flow. With the increasing function $E(N_e)$, the thermodiffusion does not bring about a variation of the electron flow; the inclusion of thermodiffusion may only bring about minor quantitative corrections rather than to new qualitative effects.

Note that the hydrodynamic description of the phenomenon, employed by us, is strictly valid only in the case of Maxwellian energy distribution of electrons. In a nonequilibrium microwave discharge, the electron distribution function may differ appreciably from the Maxwellian. Therefore, a more rigorous description calls for the calculation of the transport coefficients with due regard for the disequilibrium of the electron distribution. However, we believe that, in our case, a microwave discharge is maintained by a weak field, the disequilibrium of the electron distribution function is minor, and the hydrodynamic approach provides for a qualitatively correct description of the phenomenon. Indeed, a more rigorous calculation of the distribution function and transport coefficients for the experimental conditions produces agreement with the experimental results in observing the stratification of plasma in different gases [30].

The characteristic scale of instability being treated was determined in [11]. One can see in Eq. (13) that, with $D_{\text{eff}} < 0$, small-scale perturbations for which the diffusion term is maximal are characterized by the maximal increment. It is obvious, however, that the electron thermal conductivity brings about suppression of the rise of perturbations with a characteristic scale that is less than the thermal conductivity length of electrons. Therefore, the analysis of instability in [11] was based on the set of equations for the electron concentration and temperature. It has been revealed that the perturbations, whose characteristic scale lies between the relaxation length of electron temperature and the characteristic diffusion length of electrons, are characterized by the most increment. The obtained criterion of the emergence of instability actually coincides with that given by Eq. (15).

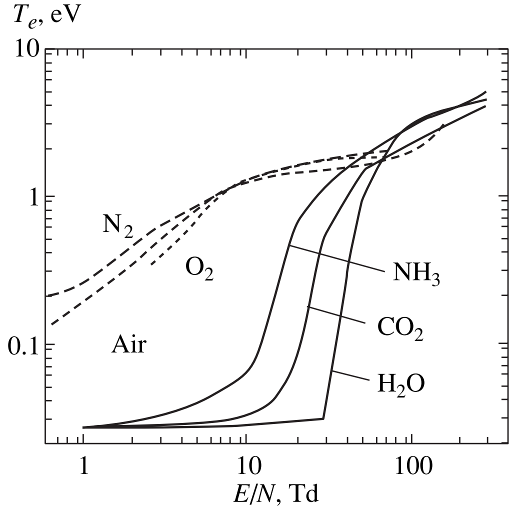

**Fig. 8.** The characteristic electron temperature $T_e = D_e/\mu_e$ as a function of reduced electric field $E/N$ for different gases according to the data of [27–29].

The theory of nonuniform steady states (autosolitons and striations) realized in nonequilibrium dissipative systems, which are described by a set of two nonlinear differential diffusion equations, is described in detail in the monograph [31]. A certain class of such systems is usually referred to as active systems with diffusion in the sense that one parameter, i.e., activator $\theta$, is used to accomplish positive feedback that is the reason for the instability of the uniform state of the system. The other parameter, inhibitor $\eta$, suppresses the process of activator rise. By its nature, the thermoelectric-current instability of microwave discharge is close to thermodiffusion autosolitons in an electron-hole plasma heated in the process of carrier photogeneration, that are treated by way of example in [31].

In [11], based on the concepts described in [31], it has been demonstrated that this situation is realized for the themoelectric-current instability being treated; in this case, it is convenient to take

$$\theta = T_e, \quad \eta = N_e T_e.$$

In such a system, it is possible to realize nonuniform steady-state solutions in the form of thermodiffusion striations. The set of equations for the electron concentration and temperature was also solved numerically in a one-dimensional approximation [11]. As a result, a nonuniform steady-state solution in the form of striations was obtained. The predicted period of stratification agrees well with that observed experimentally.

In full accordance with the theoretical concepts of [31], the distribution of inhibitor $\eta = N_eT_e$ is almost uniform. Note that $\eta$ is in fact the partial pressure of electron gas.

Therefore, the small-scale stratification of contracted state may be attributed to the development of thermoelectric-current instability, as a result of which a stationary structure in the form of fixed striations is formed. Note that the emergence of this instability in a microwave discharge calls for the development of fairly specific conditions, namely, a plasma with a supercritical electron concentration to provide for decreasing dependence $E(N_e)$, a steady-state mode of burning of discharge with low maintenance fields, and a gas of certain type with a clearly defined dependence of the electron temperature on the electric field.

## 5. CONCLUSION

The results of the investigations of a continuous microwave discharge in a quasi-optical resonator have demonstrated the possibility of existence of inhomogeneous stationary spatial structures in such a discharge. The distinguishing feature of these structures is that the form of discharge is defined by the development of instabilities rather than by the distribution of the initial electric field.

As a result of the development of a plasma-resonance instability, the diffuse form of discharge, in which the plasma distribution almost repeats the electric field distribution, changes to the contracted form. In this form, the discharge is shaped as an ellipsoid slightly extended along the electric field vector and having an electron concentration exceeding the critical value and a size that is much less than that of the diffuse discharge and the electromagnetic wavelength. The formation of such a self-localized state occurs owing to quasistatic resonance amplification of the field, arising on small-scale plasma formations at $N_e > N_c$.

In some gases such as $NH_3$, $H_2O$, and $CO_2$, the stratification of contracted discharge occurs; the discharge acquires the form of an ellipsoid consisting of alternating light- and dark-colored disks perpendicular to the electric field vector. This stratification is kinetic and due to the development of thermoelectric-current instability. The development of this instability results in a stationary structure in the form of fixed striations.

## ACKNOWLEDGMENTS

We are grateful to Prof. N.L. Aleksandrov for helpful discussions and remarks.

This study received support from the CRDF (USA, grant no. RE1-352) and from the Russian Foundation for Basic Research (project no. 00-02-16413).

## REFERENCES

1. V. B. Gil'denburg and A. V. Kim, Zh. Éksp. Teor. Fiz. **74**, 141 (1978) [Sov. Phys. JETP **47**, 72 (1978)].
2. V. B. Gil'denburg and A. V. Kim, Fiz. Plazmy **6**, 904 (1980) [Sov. J. Plasma Phys. **6**, 496 (1980)].
3. V. B. Gil'denburg, I. S. Gushchin, S. A. Dvinin, *et al.*, Zh. Éksp. Teor. Fiz. **97**, 1151 (1990) [Sov. Phys. JETP **70**, 645 (1990)].
4. A. L. Vikharev, V. B. Gil'denburg, S. V. Golubev, *et al.*, Zh. Éksp. Teor. Fiz. **94**, 136 (1988) [Sov. Phys. JETP **67**, 724 (1988)].
5. A. L. Vikharev, V. B. Gil'denburg, A. V. Kim, *et al.*, in *High-Frequency Discharge in Wave Fields*, Ed. by A. G. Litvak (Inst. Prikl. Fiz. Akad. Nauk SSSR, Gorki, 1988), p. 41.
6. A. L. Vikharev, Doctoral Dissertation in Mathematical Physics (Inst. Prikl. Fiz. Ross. Akad. Nauk, Nizhni Novgorod, 1992).
7. A. L. Vikharev, O. A. Ivanov, O. Yu. Kuznetsov, *et al.*, Fiz. Plazmy **13**, 1124 (1987) [Sov. J. Plasma Phys. **13**, 648 (1987)].
8. A. L. Vikharev, O. A. Ivanov, O. Yu. Kuznetsov, *et al.*, Dokl. Akad. Nauk SSSR **295**, 358 (1987) [Sov. Phys. Dokl. **32**, 581 (1987)].
9. A. L. Vikharev, O. A. Ivanov, L. S. Ivanova, *et al.*, Zh. Tekh. Fiz. **59** (1), 40 (1989) [Sov. Phys. Tech. Phys. **34**, 22 (1989)].
10. A. L. Vikharev, O. A. Ivanov, O. Yu. Kuznetsov, *et al.*, in *Proceedings of the IV All-Union Conference on Physics of Gas Discharge, Makhachkala, 1988,* Part 1, p. 80.
11. A. L. Vikharev, A. M. Gorbachev, O. A. Ivanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **67**, 537 (1998) [JETP Lett. **67**, 567 (1998)].
12. V. E. Golant, *Microwave Methods Plasma Research* (Nauka, Moscow, 1968).
13. A. A. Brandt, *Investigation of Dielectrics on Ultrahigh Frequencies* (Fizmatgiz, Moscow, 1963).
14. A. V. Eletskiĭ, *Chemistry of Plasma* (Énergoatomizdat, Moscow, 1982), Vol. 9, p. 151.
15. Yu. P. Raĭzer, *Modern Physical Foundations of Gas Discharge* (Nauka, Moscow, 1980).
16. A. L. Vikharev and O. Yu. Kuznetsov, Preprint No. 294, IPF AN SSSR (Institute of Applied Physics, USSR Academy of Sciences, Gorki, 1991).
17. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
18. V. B. Gil'denburg, Yu. M. Zhidko, I. G. Kondrat'ev, and M. A. Miller, Izv. Vyssh. Uchebn. Zaved., Radiofiz. **10**, 1358 (1967).
19. I. N. Inovenkov, A. V. Kim, E. I. Rakova, *et al.*, Preprint No. 271, IPF AN SSSR (Institute of Applied Physics, USSR Academy of Sciences, Gorki, 1990).
20. V. E. Golant, A. P. Zhilinskii, and I. E. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).
21. A. V. Timofeev, Zh. Tekh. Fiz. **40**, 192 (1970) [Sov. Phys. Tech. Phys. **15**, 140 (1970)].

22. N. L. Aleksandrov, A. M. Konchakov, A. P. Napartovich, and A. N. Starostin, in *Chemistry of Plasma* (Énergoatomizdat, Moscow, 1984), Vol. 11, p. 3.

23. Yu. S. Akishev, N. A. Dyatko, I. N. Lopatkin, *et al.*, in *Proceedings of the VII All-Union Conference on Physics of Low-Temperature Plasma, Tashkent, 1987*, Vol. 1, p. 261.

24. V. A. Shveigert, Fiz. Plazmy **14**, 1263 (1988) [Sov. J. Plasma Phys. **14**, 739 (1988)].

25. N. L. Aleksandrov and A. M. Okhrimovskiĭ, Fiz. Plazmy **23**, 77 (1997) [Plasma Phys. Rep. **23**, 71 (1997)].

26. A. V. Gurevich and A. B. Shvartsburg, *The Nonlinear Theory of Propagation of Radio Waves in the Ionosphere* (Nauka, Moscow, 1973).

27. Yu. P. Raizer, *Gas Discharge Physics* (Nauka, Moscow, 1987; Springer-Verlag, Berlin, 1991).

28. E. W. McDaniel, *Collision Phenomena in Ionized Gases* (Wiley, New York, 1964; Mir, Moscow, 1967).

29. M. Yousfi and M. D. Benabdessadok, J. Appl. Phys. **80**, 6619 (1996).

30. A. M. Okhrimovskiĭ, Author's Abstract of Candidate's Dissertation in Mathematical Physics (MFTI, Moscow, 1999).

31. B. S. Kerner and V. V. Osipov, *Autosolitons: Localized Strongly Nonequilibrium Regions in Homogeneous Dissipative Systems* (Nauka, Moscow, 1991).

*Translated by H. Bronstein*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 336–343.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 379–388.*
*Original Russian Text Copyright © 2001 by Likalter.*

# Structural Transition and Metallization in Liquid Selenium

**A. A. Likalter**

*Institute of Theoretical and Applied Electrodynamics, Joint Institute for High Temperatures, Russian Academy of Sciences,*
*Izhorskaya ul. 13/19, Moscow, 127412 Russia*
*e-mail:likalter@alec.msk.ru*

Received February 19, 2001

**Abstract**—The insulator–metal transitions of different kinds caused by heating above the melting temperature under pressure of tens kilobars and by compressing at the critical temperature to a pressure of about 1.1 kbar occur in liquid selenium. At tens kilobars, metallization is interpreted as the forbidden energy band vanishing due to a gradual structural transition (melting of polymer chains) described by the Clapeyron–Clausius equation. At supercritical temperatures, the insulator–metal transition is caused by percolation of overlapping electron shells (classically accessible spheres) of virtual atoms in molecules $Se_2$ remaining when polymer chains decay. The percolation threshold in such a system has been found to increase due to coupling of virtual atoms. The thermally activated conductivity in the vicinity of percolation threshold has been calculated and compared with existing experimental data. © *2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

Selenium belongs to a group of well studied low-boiling elemental metals and semiconductors which possess a considerable conductivity in the vicinity of the critical points, giving evidence of the existence of a nonideal plasma. Heated up to the critical points, the metals mercury and cesium show a more or less sharp decrease in conductivity caused by the metal–insulator transition upon expansion. In contrast, selenium demonstrates a transition from semiconductor to metallic state upon melting, heating, and compression and, as a whole, exhibits a much more complex behavior than metallic elements. As is known, there is a decaying polymer structure in melted selenium which plays an essential role in such a behavior [1]. In this connection, the insulator–metal transition in liquid selenium showing the effect of local structure is of great interest. However, available experimental information for decaying polymer structure is restricted, since it is difficult to evidence the exponential temperature dependence of conductivity which differs the insulator (semiconductor) from the metal.

The properties of semiconductors essentially connected with the forbidden energy band may radically change due to the structural transformations (note, for example, that germanium and silicon transform into metals on melting, although their amorphous phases are semiconducting [2]). Similar to germanium and silicon, selenium melts at pressures over 36 kbar to form a metallic liquid. At lower pressures, the melt is a semiconductor which transforms into a metallic liquid via a gradual phase transition at higher temperatures [3]. The insulator–metal transition can be related to the decay of a locally ordered mesoscopic structure with a forbidden band, which is responsible for the whole material properties [4] (for a review of theoretical conceptions of the insulator–metal transition, see, e.g., [5–7]). It is known that in liquid selenium near the melting point the polymer chains consist of $10^5$–$10^6$ atoms[¶], but the maximum chain length rapidly decreases with increasing temperature [8]. Assuming that the polymer structure disappears via a gradual phase transition, the insulator–metal transition can be considered a consequence of this structural transition. The significance of the structure is also evidenced by the metallization of liquid selenium expanded below an amorphous phase density, while solid selenium can only be metallized when being compressed.

At a critical temperature of the liquid–vapor phase transition ($T_c = 1888$ K, $p_c = 365$ bar, $\rho_c = 1.85$ g cm$^{-3}$) [9], a minimum metallic conductivity of about 200 $\Omega^{-1}$ cm$^{-1}$ appears at a pressure much lower than that near the melting curve, but still considerably higher than the critical pressure [10, 11]. In this domain of the phase diagram, the insulator–metal transition is due to the percolation of overlapping electron shells, which arises with increasing density. As a whole, this percolation transition is similar to that in the vicinity of the critical points of metallic elements, but there are essential peculiarities caused by the many-electron *s*-*p*-valence shell and the molecular structure of selenium. Thus, in different domains of the phase diagram, there are two different insulator–metal transitions caused by heating

---

[¶] In contrast to sulfur which polymerizes at λ point above the melting temperature.

and melting of the polymer structure and by the compression of molecular liquid. However, the transitions are mixed at a temperature of about 1500°C and a pressure of a few kbar.

Near the boiling curve, liquid selenium is a semiconductor with a predominant hole conductivity [12] determined by the energy gap width which is known from measurements of the optical absorption edge [13] (except the vicinity of the critical point where the spectral transparency window is closed, but a low conductivity shows that there is an energy gap). With increasing temperature and constant pressure, the conductivity exponentially increases and then passes through a maximum below the critical temperature where the energy gap increases due to expansion faster than temperature.[1] The conductivity maximum is characteristic of a liquid semiconductor, being in striking contrast with monotonically decreasing conductivity of expanded metals. However, the polymer melt models (see, for example, [14]) do not explain such a behavior of liquid selenium.

On the other hand, a computer simulation of the electron structure and the insulator–metal transition in such a system encounters great difficulties, since a many-electron quantum problem is linked with a great number of particles in decaying polymer chains and/or in percolation clusters. In a range of high densities and temperatures, where only short chains remain, a combined method of molecular dynamics and the electron density functional is used for the system of several tens of atoms. Simulations show, that near the critical temperature the chain distribution with respect to the number of atoms transforms into a maximum peak at $Se_2$ dimers which are also responsible for the diffusion in the system [15, 16]. However, it is obvious that the simulation of percolation transition in such a system is not possible with only a few tens of atoms.

The purpose of this paper is twofold. First of all, we interpret the structural transition in liquid selenium as the final stage of a two-step melting process in which polymer chains decay (melt). Second, we will study the insulator–metal transition in a two-atomic molecular liquid using a concept of percolation of classically accessible spheres of virtual atoms in molecules. It will be shown that the molecular structure leads to a considerable increase of the threshold density compared to the well known percolation problem of spheres without coupling and essentially influences the electronic properties of liquid. In particular, the threshold density strongly influences the activated percolation conductivity which can be compared with existing experimental data.

The outline of the paper is as follows. In Section 2, we discuss a gradual phase transition observed on heating liquid selenium. In Section 3, we consider a percolation problem for the classically accessible spheres of virtual atoms in molecules. The activated percolation conductivity of liquid selenium is analyzed in Section 4. The conclusions follow in Section 5.

A general analysis of experimental data in the context of insulator–metal transitions in different liquids has been made in our previous review [17].

## 2. MELTING OF POLYMER BUNCHES AND METALLIZATION

Selenium is a semiconductor with an energy gap of about 2 eV which occurs in amorphous, vitreous, monoclinic and hexagonal modifications with the density varying from 4.2 g cm$^{-3}$ in amorphous to 4.8 g cm$^{-3}$ in hexagonal forms. At pressures below 36 kbar, selenium melts to yield a polymer liquid which consists, as the amorphous phase, of macroscopically long deformed atomic chains packed in bunches [1]. With further heating at a high pressure, the polymer decays into $Se_2$ dimers. In view of the macroscopic size of closely packed polymer chains, this gradual phase transition (melting of chains in bunches) is extended over a relatively narrow temperature range. The melting of chains implies that the dimers can get mixed in their bunches that provides for the fluidity but, on the other hand, means the decay of polymer chains.

Except for the ends of chains, the valences of selenium atoms equal to 2 (corresponding to two unpaired spins of the electron configuration $4s^2 4p^4$) are saturated, typical of semiconductors. Due to a local order in the polymer structure, liquid (or amorphous) selenium is generally considered a semiconductor with a local forbidden band between a filled valence band (formed by nonbonding transverse *p* orbitals occupied by pairs of electrons with opposite spins) and the conduction band. The melting of polymer bunches at high pressures leads to the disappearance of the forbidden band and, therefore, to the transition of selenium into a metallic state. With relatively little heating, the transition is manifested by a sharp increase in the conductivity of liquid up to a characteristic metallic value. During the gradual transition, the length of polymer chains (characterizing the scale of ordering) continuously decreases and eventually the forbidden band disappears when this length becomes of the microscopic scale.

Thus, the metallization of liquid selenium upon heating is caused by a gradual structural phase transition (Fig. 1). There are solid and liquid semiconducting phases and a liquid metallic phase in a region where this transition branches off the melting curve ($T_{br}$ = 900 K, $p_{br}$ = 36 kbar) [3], while at higher pressures selenium, similar to other semiconductors such as silicon and germanium, melts to yield a metallic liquid. The conception of melting of polymer bunches is supported by an analysis of the experimental data based on the Clapeyron–Clausius equation of phase equilibrium

$$\frac{dp}{dT} = \frac{q}{T(v_2 - v_1)}, \tag{1}$$

---

[1] Contrary to an opinion that the gap is closed near the critical point [12], the conductivity maximum is only possible when the gap increases with increasing temperature.

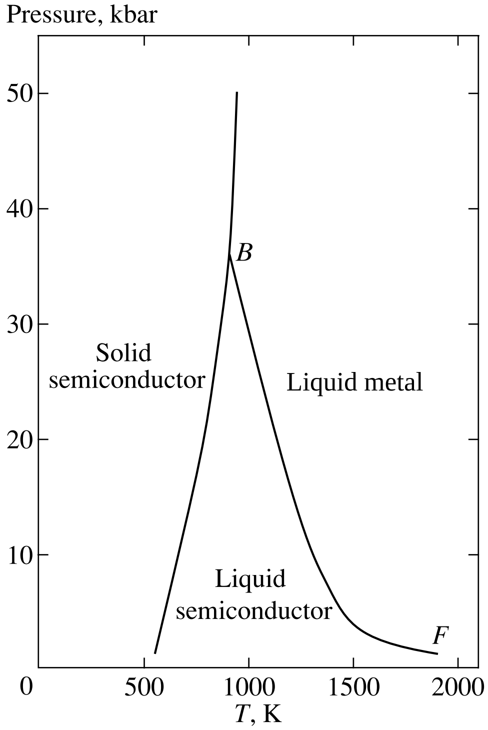

**Fig. 1.** Phase diagram of selenium: BF is the insulator–metal transition line by data of Brazhkin *et al.* [3] and of Alekseev *et al.* [25] corresponding to the minimum metallic conductivity of about 200 $\Omega^{-1}$ cm$^{-1}$.

where $p$ is the pressure, $T$ is the temperature, $q$ is the heat of melting of polymer bunches, and $v_2$ and $v_1$ are the specific volumes of metallic and semiconducting liquid phases, respectively. By existing data [3], the transition into a dimer liquid near the branching point is accompanied by a decrease in the specific volume, $v_2 - v_1 \approx -0.03/\rho$, which explains the negative slope of the phase equilibrium line, $dp/dT \approx -7$ J/cm$^3$ K. Note, that the negative slope corresponds to anomalous melting caused by heating and compression (which is also observed in usual ice). Substituting numerical values into equation (1), we find for the heat of melting of polymer bunches $q \approx 45$ J/g (i.e., 0.074 eV per Se$_2$ molecule).

On the other hand, an effective heat of melting can be expressed through the heat capacity $c_p(T)$ which is expected to have a maximum in the transition range,

$$q = \int c_p(T)dT = \langle c_p\rangle \Delta T, \tag{2}$$

where $\langle c_p\rangle$ is the average heat capacity in the transition range. Inversely, using equation (2), a latent heat calculated from the Clapeyron–Clausius equation and the transition width ($\Delta T \approx 50$ K) known from measurements, one can estimate the average capacity to be $\langle c_p\rangle \approx 0.9$ J/g K (or 8.5 $k_B$/atom, where $k_B$ is the Boltzmann constant). This value is two times greater than the heat capacity of amorphous selenium. Obviously, a considerable increase in the heat capacity is related to the dissociation of macromolecules.

Leaving studies of the semiconducting polymer melt for the future, we note that the disappearance of the ordering at a sufficiently low density leads only to the transition into a dielectric molecular liquid still possessing an energy gap. Near the transition to a metallic state, where the thermal ionization is ill-defined because of overlapping electron shells, the role playing by the energy gap goes over to a mobility gap [18]. Note that the mobility gap at the critical point reaches about 1.5 eV, i.e., differs not too much from the width of the forbidden band in liquid selenium near the melting curve. The theory of the insulator–metal transition in the vicinity of the critical point, from which, in particular, this estimate follows, is considered in more detail in the following sections.

## 3. PERCOLATION INSULATOR–METAL TRANSITION

The domain of metallic liquid selenium in the phase diagram is bounded by the line of melting of macromolecules (from the side of low temperatures) and by the metal–insulator transition due to expansion (from the side of low pressures). Along the boiling curve, selenium remains semiconducting up to the critical point [12]. The minimum metallic conductivity of about 200 $\Omega^{-1}$ cm$^{-1}$ is reached by heating the liquid up to the critical temperature with the compression up to almost three times the critical pressure. Once more we underline that the transition to a metallic state at the critical temperature occurs at a considerably lower density than the density of amorphous selenium, i.e., the absence of the ordered structure is manifested in a very radical way.

Peculiarities of the insulator–metal transition in selenium are related to features of the molecular structure. Note that the ionization potential of the Se$_2$ molecule (8.88 eV) is smaller than that of the Se atom. Thus, when the atoms form the molecule, the binding energy of electrons decreases[2] that shows evidence of a significant repulsion between electron shells. In this case, the electron states are better described by weakly overlapping Heitler–London orbitals which (in contrast to, e.g., molecular orbitals) take into account a strong interelectron correlation from the beginning. The radius of weakly overlapping classically accessible spheres of atoms, in which the residual ions are screened, is virtually unchanged, since it is determined by the potential of the residual ion and the minimum internal energy of the virtual atom $-I$ (where $I = 9.752$ eV is the ionization potential). This allows the virtual atoms in molecules to be considered as constituents of percolation clusters.

In the atom of selenium, the valence electrons with a large principal quantum number are distributed

---

[2] This also takes place for weekly bound alkali dimers and, for example, iodine dimers.

mainly near the surface of a classically accessible sphere. Taking into account that for the many-electron shell a self-consistent potential in this region can deviate from the Coulomb potential and introducing a correction, we write the classically accessible radius [17]

$$R_a = \frac{e^2}{I - \varepsilon_i}, \tag{3}$$

where $e$ is the electron charge, $\varepsilon_i \approx 0.8$ eV is the polarization affinity of the residual ion to the electron (estimated by the electron affinity of the isoelectronic As atom). Formula (3) gives classically accessible radius $R_a \approx 1.6$ Å, while the bond length in selenium molecule [19], $d = 2.17$ Å, is only by one third smaller than the sum of the radii of two atoms. Thus an overlap of the classically accessible spheres (Fig. 2) is small enough to make the concept of virtual atoms in molecules sensible.

A characteristic parameter of the percolation problem is the volume fraction of the classically accessible spheres

$$\zeta_0 = \frac{4\pi}{3} R_a^3 n_a, \tag{4}$$

where $n_a$ is the number density of virtual atoms. In atomic gas, the percolation threshold is reached when this volume fraction is about one third. However, the coupling of virtual atoms leads to a change of the percolation threshold. We define an average coordination number $B$ as the mean number of virtual atoms with centers within a coordination sphere of radius $2R_a$ surrounding the central atom. In the ideal molecular gas, the coordination number is

$$B = 1 + 4\zeta_0, \tag{5}$$

where, on the right-hand side, the unity corresponds to a permanent bond with another atom of the molecule, and the second term corresponds to bonds with atoms belonging to neighboring molecules, with one bond per molecule.[3] We use a principle of the invariable threshold coordination number of the percolation sphere problem, which is responsible for the connectivity and local structure of the infinite cluster. Substituting the threshold value $B_{pc} = 2.7$, known from Monte Carlo simulations for the ideal gas of overlapping spheres (see, for example, [20]), into equation (5), we find the threshold volume fraction of the classically accessible spheres

$$\zeta_{pc}^{(2)} = \frac{B_{pc} - 1}{4} = 0.425, \tag{6}$$

where the upper index 2 denotes the value for two-atomic fluid. In the well known percolation problem of overlapping spheres, the corresponding value is $\zeta_p^{(1)} = B_{pc}/8 = 0.3375$. Therefore, according to equation (6), the coupling increases the threshold by 25%.

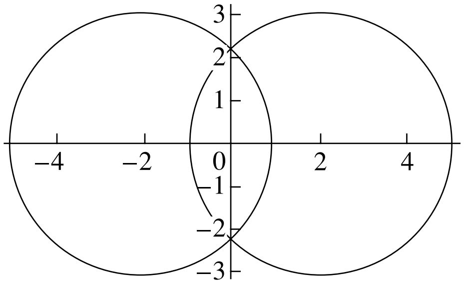

**Fig. 2.** Overlapping classically accessible spheres of virtual atoms in the molecule of selenium. Scales are given in the atomic units (Bohr radius).

The threshold for molecular chains increases still more. Although, in the case of long polymer chains, the percolation is possible at a coordination number close to 2, it is assumed that the metallic conductivity arises only when long chains decay and the coordination number increases up to the microscopic percolation threshold of the atomic system. Thus, in a broad sense, the principle of invariable threshold coordination number defines a structural percolation transition related to a definite change in the local structure. In the case of chains consisting of $k$ atoms, instead of equation (5) we have

$$B = \frac{2(k-1)}{k} + \frac{8\zeta_0}{k}, \tag{7}$$

where the first term on the right-hand side corresponds to the average coordination number within the chain and the second term corresponds to bonds with atoms belonging to neighboring molecules. Using the threshold coordination number $B_{pc}$ for the atomic system, we obtain

$$\zeta_{pc}^{(k)} = \frac{kB_{pc}}{8} - \frac{k-1}{4}. \tag{8}$$

Formula (8) with $k = 1$ and 2 yields, respectively, the thresholds for atomic and two-atomic fluids given above. With $k = 3$, we find $\zeta_{pc}^{(3)} = 0.5125$ that is comparable to the volume fraction of the classically accessible spheres at the density of amorphous selenium. Thus, the insulator–metal transition is only possible when chains are almost completely dissociated. If the mean number of atoms per chain and the volume fraction of classically accessible spheres are considered as functions of the temperature and pressure, $k(p, T)$ and $\zeta_0(p, T)$, equation (7) determines a transition line in the $p$–$T$ plane.

Now let us consider the percolation problem taking into account the repulsion between virtual atoms in the diatomic fluid. Ascribing to atoms the hard cores of diameter $d$ equal to the interatomic distance in the mol-

---

[3] We neglect the case of bonds with two atoms in one neighboring molecule, which is less probable.

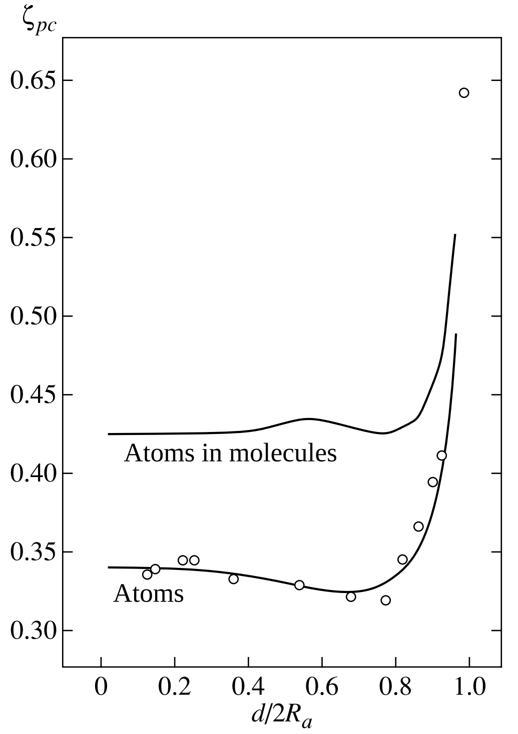

**Fig. 3.** Percolation threshold of overlapping shells as a function of the relative diameter of hard cores for virtual atoms in molecular liquid and in atomic liquid [17] (dots are by Monte Carlo simulations [22]).

ecule and taking into account the effect of an excluded volume, we write the coordination number

$$B = 1 + 4[\zeta_0 - A(c)\eta]F(\eta), \tag{9}$$

where $F(\eta)$ is the Karnahan–Starling function [21] describing the density with respect to a free volume,

$$F = \frac{1+\eta+\eta^2-\eta^3}{(1-\eta)^3},$$

$\eta = \pi d^3 n_a/6$ is the packing fraction of the hard cores. In formula (9), the subtrahend $A(c)\eta$ with the coefficient $A(c)$ depending on the relative radius of cores $c = d/2R_a$, allows for an excluded volume around the molecule with one atom placed at the center of coordination sphere. A direct calculation of the excluded volume gives at $c < 1/2$

$$A(c) = 27/16,$$

and at $1/2 < c < 1$

$$A(c) = \frac{27}{16} - \left(1 - \frac{1}{2c}\right)^2\left(1 + \frac{1}{c} + \frac{3}{4c^2}\right).$$

Substituting formula $\eta = c^3\zeta_0$ and the threshold value of the coordination number into equation (9), we obtain an algebraic equation which determines $\zeta_{pc}(c)$ as an implicit function of the relative diameter of cores (for the sake of brevity, here and below the upper index 2 is omitted). This function is plotted in Fig. 3. In a wide range of the relative size of cores $c < 0.8$ (in particular, for selenium $c \approx 2/3$) the magnitude of this function virtually coincides with the threshold given by equation (6) for the ideal gas.[4] In Fig. 3 this function is compared with the results of analogous calculations without coupling [17] and Monte Carlo simulations [22], existing for this case, which confirm the principle of the invariable threshold coordination number.

Thus, the coupling of virtual atoms leads to an essential increase of the percolation threshold compared to that of an atomic liquid. Using formulas (3), (4), and (6), it is easy to estimate that the percolation threshold corresponds to the density of about 3.2 g/cm$^3$, i.e., by approximately 25% smaller than the density of amorphous selenium. Indirectly, this estimate of the percolation threshold can be confirmed by the experiment, since the activated percolation conductivity strongly depends on the distance from the threshold, as will be shown in the next section.

The pressure at the transition point can be estimated using a scaling equation of the critical isotherm

$$\frac{p}{p_c} - 1 = A\left(\frac{\rho}{\rho_c} - 1\right)^{\delta}, \tag{10}$$

where $\delta \approx 5$ is the universal critical exponent, and $A \approx 8.71$ is the amplitude found by fitting the experimental data [9]. Substituting the estimated threshold density into equation (10), we obtain the transition pressure of about 1.1 kbar, i.e., nearly 3 times the critical pressure.

## 4. PERCOLATION CONDUCTIVITY

In a macroscopically large percolation cluster, the molecules, as well as virtual atoms, are no longer separate particles, but the constituents of an electron-ion plasma. The valence electrons are partially free, since the screening of residual ions in the percolation cluster becomes collective. Therefore, mixed states of virtual atoms have a continuous spectrum of internal energy including the energy of free motion in the screened potential of residual ions.

In the percolation problem, it is convenient to take the ground level of the remainder ion to be zero energy. Then, the minimum internal energy of the virtual atoms with a one-electron excitation [18] is

$$E_p = -I + \varepsilon_p, \tag{11}$$

where $\varepsilon_p = p^2/2m$ is the energy of asymptotically free motion of the electron, $p$ is the asymptotic momentum, $m$ is the electron mass. Depending on the proximity to

---

[4] Besides the repulsion, there is an averaged many-particle exchange attraction between virtual atoms in the percolation cluster, which only weekly depends on the configuration and, therefore, as assumed, does not influence the percolation threshold.

the transition point, the renormalized Fermi energy of such mixed states of virtual atoms is

$$\varepsilon'_F = \hbar^2 k_F^2 \vartheta^2/2m, \tag{12}$$

where

$$k_F = (3\pi^2 z n_a)^{1/3} \tag{13}$$

is the Fermi wavevector of the electron gas, $z = 6$ is the number of $s$ and $p$ valence electrons of selenium with almost the same radii of the classically accessible spheres, which are considered as equivalent, $\vartheta \le 1$ is a renormalization factor caused by partial localization of electrons, which is determined by the relative frequency of the electron transitions between virtual atoms. The renormalized Fermi energy $\varepsilon'_F$ and the wave vector

$$k'_F = k_F \vartheta \tag{14}$$

go over to the corresponding quantities for a homogeneous electron gas in the limit when the frequency of the electron transitions corresponds to a free flight between neighboring atoms, i.e., the localization factor goes to unity.[5] As will be seen, the Fermi energy decreases near the insulator–metal transition with decreasing temperature faster than the temperature; therefore, the excitation of virtual atoms is described by the Boltzmann statistics.

In the vicinity of the percolation transition, the valence electrons of overlapping $s$–$p$ shells (which can transfer between virtual atoms) provides an over-barrier hopping conductivity. Analogously to the band conductivity depending on filling of the valence band, the percolation hopping conductivity depends on filling of the $s$–$p$ shells. The frequency of hopping is proportional to the product of the number of valence electrons by the number of free places (holes) in these shells, since the transitions are only possible to such holes. We assume that due to the mixing of states the maximum number of electrons in $s$–$p$ shells is greater than that in free atoms by a small value $z$, and the number of holes is $8 + \varepsilon - z$. Then, we define an effective number of valence electrons $z_e$ proportional to the frequency of hopping and obeying the conditions $z_e(1) = z_e(8) = 1$ for the minimum and maximum numbers of $s$–$p$ electrons possible in free atoms. It is easy to see that with these conditions $\varepsilon$ must be taken equal to 1, and the effective number of valence electrons is

$$z_e = \frac{1}{8} z (9 - z). \tag{15}$$

As a function of $z$, the effective number of valence electrons defined above has a week maximum and, in the case of selenium, is close to 2, i.e., coincides with the number of holes in the valence $s$–$p$ shells of a free atom.

The percolation conductivity is described by a modified Drude formula with the localization factor [18]

$$\sigma = \frac{e^2 n_a z_e \tau \vartheta}{m}, \tag{16}$$

where $\tau$ is the relaxation time. In the case of the Boltzmann statistics, the minimum relaxation time is

$$\tau = \frac{R_s}{v_T}, \tag{17}$$

where $R_s \approx (4\pi n_a/3)^{-1/3}$ is the mean interatomic distance, $v_T = (8T/\pi m)^{1/2}$ is the mean thermal velocity of electrons. At high density, in the case of nearly free degenerate electrons, the relaxation time is

$$\tau = \frac{l}{v_F}, \tag{18}$$

where $v_F$ is the Fermi velocity, $l = R_s/\gamma$ is the free path found at the density where the localization factor becomes equal to unity, $\gamma \le 1$ is a parameter corresponding to the inverse value of the relative free path. As an interpolation, we use formula (17) if $v_T > \gamma v_F$ and formula (18) otherwise $v_T < \gamma v_F$ (however, we underline that this does not really concern the crossover between the Boltzmann case and the degenerate case, which is determined by the localization factor).

Below the percolation threshold, a partial localization of electrons, described by the localization factor, is connected with a mobility gap $\Delta_{pc}$, and in a wider range, with a soft gap $\Delta_{cp}$ of low mobility (smaller than that at the minimum free path),

$$\Delta_k = I - \varepsilon_i - e^2 \left(\frac{4\pi n_a}{3\zeta_k}\right)^{1/3} \quad (k = pc, cp), \tag{19}$$

where $\zeta_{cp} \approx 0.425$ is the threshold volume fraction of the classically accessible spheres, and $\zeta_{cp} \approx 0.74$ corresponds to close packing of these spheres. At the edge, the mobility goes to zero as the inverse of correlation length [6] being described by a power function

$$\vartheta(\varepsilon_p) = \left(\frac{\varepsilon_p - \Delta_{pc}}{\Delta_{cp} - \Delta_{pc}}\right)^{\nu}, \quad \varepsilon_p < \Delta_{cp}, \tag{20}$$

where $\nu \approx 0.9$ is the critical exponent of the correlation length of percolation clusters [23]. The localization factor is obtained by averaging the scaling function $\vartheta(\varepsilon_p)$, extended to the range $\varepsilon_p > \Delta_{cp}$ where it is equated to unity, over the Boltzmann distribution of virtual atoms (for the simplicity, the exponent $v$ can be replaced with unity). On the insulating side of the transition, the localization factor exponentially depends on the temperature

---

[5] Equations (12) and (13) replace an expression for the renormalized Fermi energy through the statistical weight of the ground level [18], which obeys this condition only for $s$ electrons.

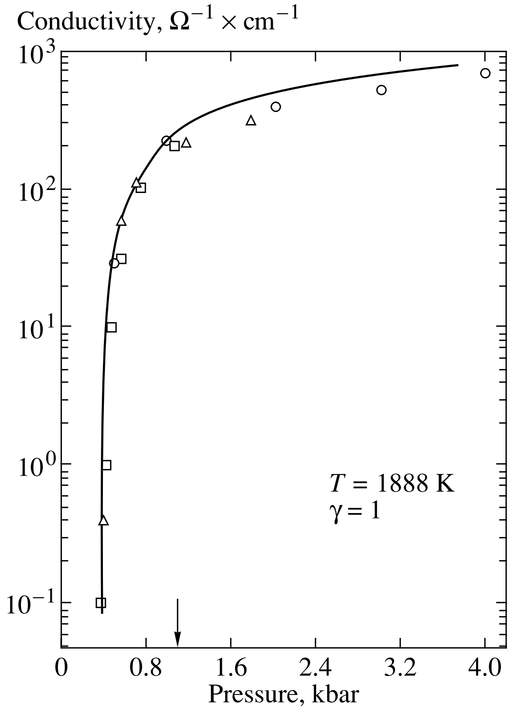

**Fig. 4.** Percolation conductivity of expanded fluid selenium as a function of the pressure along the critical isotherm. The pressure of the insulator–metal transition is about 1.1 kbar (shown by an arrow). Experimental dots: (□) Hoshino *et al.* [10, 11] (see also Fig. 1 in [9]); (○) Alekseev *et al.* [25]; (△) R. Fischer and R.W. Schmutzler [1].

$$\vartheta = \frac{T(F_{pc}-F_{cp})}{\Delta_{cp}-\Delta_{pc}} \approx \frac{2}{\sqrt{\pi}}\frac{(\Delta_{pc}T)^{1/2}}{\Delta_{cp}-\Delta_{pc}}\exp\left(-\frac{\Delta_{pc}}{T}\right), \qquad (21)$$

where functions $F_k$ are expressed by the incomplete gamma functions $\Gamma(m, x)$,

$$F_k = \frac{2}{\sqrt{\pi}}\left[\Gamma\left(\frac{5}{2}, \frac{\Delta_k}{T}\right) - \frac{\Delta_k}{T}\Gamma\left(\frac{3}{2}, \frac{\Delta_k}{T}\right)\right],$$

and the second part of formula (21) corresponds to the limit $\Delta_{cp} > \Delta_{pc} \gg T$. Thus, below the transition point, the mobility gap plays a role of the activation energy of percolation conductivity.

On the metallic side of the transition, the mobility gap is virtual, $\Delta_{pc} \leq 0$, and the localization factor retains only a weak temperature dependence

$$\vartheta = \frac{\frac{3}{2}T - \Delta_{pc} - TF_{cp}}{\Delta_{cp}-\Delta_{pc}} \approx \frac{\frac{3}{2}T - \Delta_{pc}}{\Delta_{cp}-\Delta_{pc}}. \qquad (22)$$

The second part of formula (22) corresponds to the limit $\Delta_{cp} \gg T$. This formula shows that at the transition point, where the activation energy equals zero, the localization factor is proportional to $T$. Therefore, the Fermi energy is proportional to $T^2$, i.e., as was noted above, decreases faster than temperature. According to (19) and (22), with increasing density, when the soft gap $\Delta_{cp}$ becomes of the order of temperature, the localization factor goes to unity. Then, the Fermi energy increases giving rise to the degeneracy of electrons on the metallic side of the transition (although the localization factor itself can be calculated almost always using the Boltzmann distribution [24]).

Shown in Fig. 4 is the conductivity of selenium as a function of the pressure along the critical isotherm, calculated with formulas (15)–(22) and the equation of state (10). Experimental data at high pressures [1, 25] are best fitted choosing the maximum parameter $\gamma = 1$ corresponding to the minimum free path length equal to the mean distance between virtual atoms. In view of a strong pressure dependence of the conductivity near the percolation transition, the experimental data indirectly confirm a locus of the transition point determined by formula (6). At the critical point, the calculated conductivity is about 0.1 $\Omega^{-1}$ cm$^{-1}$ in agreement with the experimental value [9]. We note, that a vertically going dependence on the pressure near the critical point is caused by a diverging compressibility, rather than by the insulator–metal transition located at nearly three times the critical pressure.

In conclusion, we discuss the thermoelectric coefficient of selenium retaining positive sign in the entire range of measurements [1]. In the case of electron–hole conductivity, the positive sign corresponds to the higher mobility of holes [12]. However, near the insulator–metal transition, the valence electrons become partially free, therefore, the thermal generation of electrons and holes is ill-defined. Thus, the reason of positive thermoelectric coefficient connected with the sign of carriers is still an open question. The problem becomes not so sharp, since the thermoelectric coefficient is proportional to the difference of the chemical potential and a weighted mean energy of electrons, [26], which can also change the sign. Additional data at higher temperatures and densities where the thermoelectric coefficient may change the sign from positive to negative, and the Hall constant measurements could shed light on the problem.

## 5. CONCLUSIONS

The transition of liquid selenium into a metallic state near the melting curve is different from that in the vicinity of the critical point. Near the melting curve, this transition is caused by the decomposition of the polymer structure responsible for a local forbidden band. Therefore, the insulator–metal transition line branching off the melting curve is described by the Clapeyron–Clausius equation for a gradual structural transition. At high pressures, a rather sharp dissociation of the polymer chains upon heating of liquid selenium has been confirmed by observations, therefore, some connection between the dissociation and the metallization can be considered an experimental fact.

On the other hand, the metallization of a molecular liquid, formed after the decomposition of polymer chains, is possible only in the case of overlapping electron shells, i.e., at high enough density. Therefore, the insulator–metal transition upon compression in the vicinity of the critical point is interpreted as percolation of weakly overlapping classically accessible spheres of coupled virtual atoms in the dimers of selenium. As is shown, the molecular structure leads to an essential increase of the percolation threshold compared to that of the atomic liquid: the longer are molecular chains, the stronger the threshold increases. Nevertheless, the activation energy of the percolation conductivity vanishes at a density smaller than that of semiconducting amorphous selenium. At the critical isotherm, the transition point is located at a pressure of about 1.1 kbar where liquid selenium transforms into a molecular metal with a conductivity greater than 200 $\Omega^{-1}$ cm$^{-1}$.

The activated percolation conductivity is caused by over-barrier electron hopping with an effective number of electrons per atom equal to the number of holes in *s*–*p* shells of selenium. The conductivity is described by a modified Drude formula allowing for a partial localization of electrons. Well above the transition point, where the localization is not important, experimental data correspond to the free path length of electrons equal to the mean interatomic distance. The percolation conductivity describes experimental data along the critical isotherm in a wide pressure range, where the conductivity varies by four orders of magnitude.

The insulator–metal transition is only possible when polymer chains are almost completely dissociated, and the local structure corresponds to a critical value of the coordination number. In the region where the percolation transition line goes over to the melting line of macromolecules, two kinds of the insulator–metal transition become identical.

## ACKNOWLEDGMENTS

This work was partly supported by the Russian Foundation for Basic Research, project nos. 01-02-16254 and 00-15-96529.

## REFERENCES

1. R. Fischer and R. W. Schmutzler, in *The Physics of Selenium and Tellurium*, Ed. by E. Gerlach and P. Grosse (Springer-Verlag, Berlin, 1979), Springer Ser. Solid-State Sci. **13**, 225 (1979).
2. N. F. Mott and E. A. Davis, *Electron Processes in Noncrystalline Materials* (Clarendon Press, Oxford, 1979; Mir, Moscow, 1992).
3. V. V. Brazhkin, R. N. Voloshin, and S. V. Popova, Pis'ma Zh. Éksp. Teor. Fiz. **50**, 392 (1989) [JETP Lett. **50**, 424 (1989)]; V. V. Brazhkin, private communication.
4. J. M. Ziman, *Models of Disorder: the Theoretical Physics of Homogeneously Disordered Systems* (Cambridge Univ. Press, Cambridge, 1979; Mir, Moscow, 1982).
5. N. F. Mott, *Metal–Insulator Transition* (Taylor & Francis, London, 1991).
6. P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
7. M. J. Hirsch, Comments Condens. Matter Phys. **13**, 249 (1987).
8. W. W. Warren, Jr. and R. Dupree, Phys. Rev. B **22**, 2257 (1980).
9. S. Hosokawa, T. Kuboi, and K. Tamura, Ber. Bunsenges. Phys. Chem. **101**, 120 (1997).
10. H. Hoshino, R. W. Schmutzler, and F. Hensel, Ber. Bunsenges. Phys. Chem. **80**, 27 (1976).
11. H. Hoshino, R. W. Schmutzler, W. W. Warren, Jr., and F. Hensel, Philos. Mag. **33**, 255 (1976).
12. W. W. Warren, Jr., J. Non-Cryst. Solids **250–252**, 506 (1999).
13. S. Hosokawa and K. Tamura, J. Non-Cryst. Solids **205–207**, 261 (1996).
14. F. Yonezawa, H. Ohtani, and T. Yamaguchi, J. Non-Cryst. Solids **250–252**, 510 (1999).
15. F. Shimojo, K. Hoshino, M. Watabe, and Y. Zempo, J. Phys.: Condens. Matter **10**, 1199 (1998).
16. F. Shimojo, K. Hoshino, Y. Zempo, and M. Watabe, J. Non-Cryst. Solids **250–252**, 542 (1999).
17. A. A. Likal'ter, Usp. Fiz. Nauk **170**, 831 (2000).
18. A. A. Likal'ter, Zh. Éksp. Teor. Fiz. **94** (12), 157 (1988) [Sov. Phys. JETP **67**, 2478 (1988)].
19. K.-P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure* (Van Nostrand, New York, 1979; Mir, Moscow, 1984).
20. B. I. Shklovskiĭ and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Nauka, Moscow, 1979; Springer-Verlag, New York, 1984).
21. N. F. Carnahan and K. E. Starling, J. Chem. Phys. **51**, 635 (1969).
22. A. L. Bug, S. A. Safran, G. S. Grest, and I. Webman, Phys. Rev. Lett. **55**, 1896 (1985).
23. D. Staufer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, Philadelphia, 1991, 2nd ed.).
24. A. A. Likal'ter, Zh. Éksp. Teor. Fiz. **107**, 1996 (1995) [JETP **80**, 1105 (1995)].
25. V. A. Alekseev, V. G. Ovcharenko, and Yu. F. Ryzhkov, J. Phys. (Paris) **41**, C8-89 (1980); missed Figures, private communication by V. A. Alekseev.
26. A. A. Likalter, in *Transport and Optical Properties of Nonideal Plasma*, Ed. by G. A. Kobzev, I. T. Iakubov, and M. M. Popovich (Plenum, New York, 1995).

*Translated by A. Likalter*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 344–362.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 389–408.*
*Original Russian Text Copyright © 2001 by Val'kov, V. Romanov, M. Romanov.*

**FLUIDS**

# Fluctuations in Bounded Cells of Liquid Crystals in an External Field

**A. Yu. Val'kov[b], V. P. Romanov[a,*], and M. V. Romanov[a]**
[a]*St. Petersburg State University (Petrodvorets Branch), Universitetskiĭ pr. 2, Petrodvorets, 198904 Russia*
[b]*St. Petersburg Institute of Foreign Economic Relations, Economics, and Law, St. Petersburg, 198104 Russia*
**e-mail: vadim.romanov@pobox.spbu.ru*

Received April 3, 2001

**Abstract**—Spatial correlation functions of orientation fluctuations in bounded cells of smectic and nematic liquid crystals are calculated taking into account the effect of external fields and finiteness of the energy of anchoring to the surface. The cases of positive and negative anisotropies of magnetic susceptibility or permittivity are considered. The calculations are based on the division of degrees of freedom into bulk and surface ones and on the reduction of the computation of the continual integral determining the correlation function to the solution of the Euler equation with corresponding boundary conditions of the first or third kind. The obtained correlation functions are used for describing the intensity of light scattered in nematics for the planar and homeotropic orientations. It is shown, in particular, that the measurements of the angular dependence of the scattered light intensity may serve as a reliable method for determining the energy of anchoring of a liquid crystal to a substrate for different values of the external field. *© 2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

An important feature of liquid crystals (LC) is the very small value of orientational melting energy. As a result, the orientation of a system in the ordered phase may change significantly due to very weak effects; in other words, the susceptibility of the system is extremely high. For this reason, most types of LC exhibit strong fluctuations of orientation [1]. Fluctuations strongly affect many properties of LC. This refers to the behavior of the system in the vicinity of phase transitions [1–5], optical properties (above all, light scattering} [1, 6], the formation of viscoelastic coefficients [1], the fluctuational contribution to attraction between walls [7–10], etc.

A distinguishing feature of orientation fluctuations in LC is their infinitely large correlation radius. In such a situation, an important role is played by factors which may be disregarded in most other systems. This primarily refers to the effect of external fields [1, 11, 12], the interaction of LC with the confining surface [5, 8, 11, 13], and finite dimensions and shape of the sample [3, 8, 14]. This problem has become quite important in connection with numerous applications of LC in information display systems. The description of the structure and spectrum of thermal noise for LC in encapsulated and twisted cells is essential in this case.

The statistical properties of fluctuations are described with the help of correlation functions. First calculations of spatial correlation functions of fluctuations in LC were made for unbounded nematic and smectic-*A* liquid crystals (NLC and SLC-*A*) [15]. The evolution of these investigations was aimed at extending the classes of LC [1], types of fluctuations [16], inclusion of dynamic processes [1], and taking into account the finiteness of the system and interaction of LC molecules with the surface [10, 17– 21].

The most serious difficulties are encountered when correlation functions are calculated for finite systems. In most cases, fluctuations in a planar cell are considered. The simplest case of stringent boundary conditions in NLC was considered in [17–19]. The correlation functions obtained in these works had the form of an infinite series in the eigenfunctions of the system. A more realistic model of mild boundary conditions described by a Rapini type potential [22] was considered for NLC in [10, 20] and for SLC in [21, 23]. The authors of [10, 20, 23] succeeded in presenting the result in closed form, while in [21] it was obtained in the form of a series.

It should be noted here that no universal approach has been developed for calculating the correlation function for finite LC. The methods of expansion in eigenfunctions [10, 17] and continual integration [10], as well as the theory of self-conjugate operators [20, 23] were used. It may appear at first glance that in each publication, the computational method was based on the specific features of the given system.

In the present work, we propose an algorithm for calculating the correlation functions for bounded LC, which combines the approaches used in [10, 17, 18, 20, 21] and makes it possible to obtain the result in closed form in the general case of a multicomponent order parameter.

This approach can be used for calculating the correlation functions for plane-parallel NLC and SLC-*A* cells in the presence of an external field. We consider the case when the directions of orientation of the director by the external field and confining surfaces coincide and analyze the systems with positive and negative anisotropy of magnetic susceptibility $\chi_a$ or permittivity. In the most commonly encountered case, with $\chi_a > 0$, the equilibrium position of the director is aligned with the external field. In the case of negative anisotropy, the director is perpendicular to the magnetic field in equilibrium. For such systems, the magnetic field suppresses only one of two fluctuational modes of the director, the other mode remaining singular [24].

We derived an explicit expression for the correlation functions for smectics *A* and nematics in planar geometry taking into account the external field and the effect of the surface simultaneously. The behavior of the correlation function near the surface is analyzed in detail. It is shown that fluctuations at the surface may be stronger or weaker than in the bulk depending on the system parameters. The fluctuations in a bounded cell with the planar geometry is calculated in detail for the first time without using the single-constant approximation.

In Section 2, basic equations describing the energy of a bounded liquid crystal in an external field are introduced. A general approach to calculating the spatial correlation function of the vector order parameter fluctuations in a bounded system is developed in Section 3. A general expression for the correlation function is derived in Section 4. In Section 5, the obtained formulas are used for calculating the correlation function of displacement fluctuations in SLC-*A* and for fluctuations of the director in NLC. The dependence of the spatial correlation function on the parameters of the system (above all, on the anchoring energy and the applied external field) is analyzed in detail. The intensity of scattered light in a liquid crystal cell is calculated in Section 6. It is shown that the parameters of a liquid crystal (above all, the anchoring energy) can be determined by measuring the angular dependence of the intensity of scattered light in external fields.

## 2. FREE ENERGY OF A LIQUID-CRYSTAL CELL IN AN EXTERNAL FIELD

In order to describe liquid crystals, we will use the standard continual model [1]. In this model, the elastic properties of liquid crystals in bounded cells are usually described taking into account the three types of contributions to the elastic energy:

$$F_{\text{tot}} = F_e + F_f + F_{sf}. \tag{2.1}$$

Here, $F_e$ is the elastic energy of a liquid crystal, $F_f$ is the contribution of the external field, and $F_{sf}$ is the surface energy. The approach which will be described below can be used for studying various types of liquid crystals. In the present work, we consider nematics and smectics *A*.

For a nematic liquid crystal, the elastic energy is equal to the Frank energy

$$\begin{aligned} F_e = {} & \frac{1}{2}\int d\mathbf{r}[K_{11}(\operatorname{div}\mathbf{n})^2 \\ & + K_{22}(\mathbf{n}\cdot\operatorname{curl}\mathbf{n})^2 + K_{33}(\mathbf{n}\times\operatorname{curl}\mathbf{n})^2], \end{aligned} \tag{2.2}$$

where $\mathbf{n}(\mathbf{r})$ is the unit vector of the director and $K_{jj}$ ($j$ = 1–3) are the Frank moduli. The minimum value of $F_e$ corresponds to the homogeneous equilibrium state of the director $\mathbf{n}^0 = \text{const}$.

In order to calculate the volume energy of SLC-*A*, we confine our analysis to the standard model [1]:

$$F_e = \frac{1}{2}\int d\mathbf{r}\{B[\partial_z u(\mathbf{r}_\perp, z)]^2 + K[\boldsymbol{\nabla}_\perp u(\mathbf{r}_\perp, z)]^2\}. \tag{2.3}$$

Here, $u(\mathbf{r})$ is the component of the displacement vector along the $z$ axis, which is directed across the smectic layers; $B$ is the smectic elastic constant associated with the compression of smectic layers; $K$ is the elastic constant associated with the distortion of the layer shape;

$$\boldsymbol{\nabla}_\perp = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right),$$

where the subscript "⊥" corresponds to vector components perpendicular to the $z$ axis. Model (2.3) corresponds to the director vector

$$\mathbf{n}(\mathbf{r}) \parallel (-\boldsymbol{\nabla}_\perp u, 1) \equiv \left(-\frac{\partial u}{\partial x}, -\frac{\partial u}{\partial y}, 1\right),$$

which is normal to smectic layers.

The term $F_f$ describing the magnetic or electric field in SLC-*A* and NLC has the form

$$\begin{aligned} F_f &= -\frac{1}{2}\chi_a\int d\mathbf{r}(\mathbf{n}\cdot\mathbf{H})^2, \\ F_f &= -\frac{1}{8\pi}\varepsilon_a\int d\mathbf{r}(\mathbf{n}\cdot\mathbf{E})^2, \end{aligned} \tag{2.4}$$

respectively, where $\chi_a$ and $\varepsilon_a$ are the anisotropies of the magnetic susceptibility and permittivity, and $\mathbf{H}$ and $\mathbf{E}$ are the magnetic and electric field strengths. For the sake of definiteness, we will confine our subsequent analysis to magnetic fields only. For a transition to electric fields, it is sufficient to make the substitution $\mathbf{H} \longrightarrow \mathbf{E}$ and $\chi_a \longrightarrow \varepsilon_a/4\pi$ in the final formulas. It follows from Eq. (2.4) that the term $F_f$ has a minimum for $\chi_a > 0$ if $\mathbf{n}^0 \parallel \mathbf{H}$ and for $\chi_a < 0$ if $\mathbf{n}^0 \perp \mathbf{H}$.

The term $F_{sf}$ describes the surface energy. The following two cases of the orientation of the director on the surface are considered most frequently: the planar case when the director is parallel to the surface and the homeotropic case when it is perpendicular to the surface. Let a liquid crystal be enclosed in a plane-parallel cell of thickness $L$. We introduce the Cartesian system of coordinates $\{\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z\}$ with the origin at the center of the cell and with the $\mathbf{e}_z$ axis perpendicular to the planes $z = l_1$ and $z = l_2$ ($L = l_2 - l_1$) confining the cell.

For nematic liquid crystals, we will use the surface energy $F_{sf}$ in the form of the Rapini potential [22]:

$$F_{sf} = \int d\mathbf{r}_\perp \sum_{j=1,2} W_j n_\perp^2(\mathbf{r}_\perp, l_j), \tag{2.5}$$

in the case of the homeotropic orientation and

$$F_{sf} = \frac{1}{2}\int d\mathbf{r}_\perp \times \sum_{j=1,2} [W_j^{(1)} n_x^2(\mathbf{r}_\perp, l_j) + W_j^{(2)} n_z^2(\mathbf{r}_\perp, l_j)] \tag{2.6}$$

in the case of the planar orientation with the easy orientation axis $y$. Here, the quantities $W_j$ and $W_j^{(1,2)}$ ($j = 1, 2$) are the anchoring energies.

For SLC-$A$, we confine ourselves to the case when smectic layers are perpendicular to the $z$ axis. In this case,

$$F_{sf} = \frac{1}{2}\int d\mathbf{r}_\perp \times \sum_{j=1,2} [\gamma_1^{(j)} u^2(\mathbf{r}_\perp, l_j) + \gamma_2^{(j)} (\nabla_\perp u(\mathbf{r}_\perp, l_j))^2], \tag{2.7}$$

where $\gamma_1^{(j)}$ are the anchoring coefficients and $\gamma_2^{(j)}$ are surface tensions ($j = 1, 2$). In the case of a freely suspended smectic film, we have $\gamma_1^{(j)} = 0$.

In the present work, we will be interested in the case when the confining surfaces as well as the external field stabilize the orientation of the director. For this reason, we will disregard the competition between the field and the surfaces (Freedericksz effect) or between two surfaces (twisted cell).

In the case of NLC, we will be interested in deviations $\delta\mathbf{n} = \mathbf{n} - \mathbf{n}^0$ of the director $\mathbf{n}$ from its equilibrium value $\mathbf{n}^0$. In the principal order, we have $\delta\mathbf{n} \perp \mathbf{n}^0$. In the case of SLC-$A$, we take for the fluctuating parameter the quantity $u(\mathbf{r})$ describing the deviation of smectic layers from the equilibrium state.

The structure of formulas (2.2)–(2.7) implies that the fluctuation contribution to the total energy (2.1) in the Gaussian approximation is the quadratic form of the fluctuating parameter $u(\mathbf{r})$ or $\delta\mathbf{n}$ containing gradients of an order not higher than the second.

The conventional method for solving the problem of fluctuations in an unbounded NLC is based on a transition to the 3D Fourier spectrum. For a plane-parallel cell, it is natural to carry out the 2D Fourier transformation. We will use the Fourier transformation in the form

$$f(\mathbf{r}) = \frac{1}{(2\pi)^2}\int f(\mathbf{q}, z) e^{i\mathbf{q}\mathbf{r}_\perp} d\mathbf{q}.$$

Thus, the problem formally involves an analysis of the correlation function of the scalar parameter $\phi(\mathbf{q}, z) = u(\mathbf{q}, z)$ or the two-component vector $\boldsymbol{\phi}(\mathbf{q}, z) = \delta\mathbf{n}(\mathbf{q}, z)$. It makes the contribution to the bulk energy $F_{bk} = F_e + F_f$ of the form

$$\delta F_{bk}(\boldsymbol{\phi}) = \int \frac{d\mathbf{q}}{(2\pi)^2} \Phi_{bk}(\boldsymbol{\phi}), \tag{2.8}$$

where $\Phi_{bk}$ is the positive-definite quadratic form of the type

$$\Phi_{bk} = \frac{1}{2} \times \int_{l_1}^{l_2} (\boldsymbol{\phi}'^{+} \hat{a} \boldsymbol{\phi}' + \boldsymbol{\phi}'^{+} \hat{b} \boldsymbol{\phi} + \boldsymbol{\phi}^{+} \hat{b}^{+} \boldsymbol{\phi}' + \boldsymbol{\phi}^{+} \hat{c} \boldsymbol{\phi}) dz. \tag{2.9}$$

Here, $\boldsymbol{\phi} = \boldsymbol{\phi}(\mathbf{q}, z)$ and $\boldsymbol{\phi}' = \partial\boldsymbol{\phi}(\mathbf{q}, z)/\partial z$ are the $n$-component vectors, $\hat{a} = \hat{a}(\mathbf{q})$, $\hat{b} = \hat{b}(\mathbf{q})$, and $\hat{c} = \hat{c}(\mathbf{q})$ are $n \times n$ square matrices, $\hat{a}$ and $\hat{c}$ being Hermitian matrices, and the superscript + denotes Hermitian conjugation. The dimension $n = 1$ for SLC-$A$ and $n = 2$ for NLC.

The contribution of $\boldsymbol{\phi}$ to the surface energy (2.5), (2.6), and (2.7) of the system has the form

$$\delta F_{sf}(\boldsymbol{\phi}) = \int \frac{d\mathbf{q}}{(2\pi)^2} \Phi_{sf}(\boldsymbol{\phi}), \tag{2.10}$$

where

$$\Phi_{sf} = \frac{1}{2}(\boldsymbol{\phi}_1^{+} \hat{w}_2 \boldsymbol{\phi}_1 + \boldsymbol{\phi}_2^{+} \hat{w}_2 \boldsymbol{\phi}_2), \tag{2.11}$$

$\boldsymbol{\phi}_1 = \boldsymbol{\phi}(\mathbf{q}, l_1)$, $\boldsymbol{\phi}_2 = \boldsymbol{\phi}(\mathbf{q}, l_2)$, and the Hermitian positive definite $n \times n$ matrices $\hat{w}_{1,2}(\mathbf{q})$ correspond to the first ($z = l_1$) and second ($z = l_2$) boundaries.

The specific form of matrices $\hat{a}$, $\hat{b}$, $\hat{c}$, $\hat{w}_1$, and $\hat{w}_2$ depends on the type of the liquid crystal and on the geometry under investigation. The expressions for these matrices will be given in Section 5.

## 3. GENERAL METHOD FOR CALCULATING THE CORRELATION FUNCTION FOR A FINITE CELL

We will be interested in the correlation function of fluctuations in the general case of the vectorial order parameter $\phi(\mathbf{q}, z)$:

$$g_{\alpha\beta}(\mathbf{q}; z, z_1) = \langle \phi_a(\mathbf{q}, z) \phi_\beta^*(\mathbf{q}, z_1) \rangle, \tag{3.1}$$

where the angle brackets ⟨…⟩ indicate statistical averaging and the superscript ⟨*⟩ indicates complex conjugation.

In order to calculate tensor $\hat{g}$, we must average the expression $\phi_\alpha(z)\phi^*_\beta(z_1)$ over all possible values of $\boldsymbol{\phi}(\mathbf{q}, z)$ with the weight function $\exp(-F_{\mathrm{tot}}(\boldsymbol{\phi})/k_BT)$.

Since the quantity $\phi(r)$ is real-valued, we have $\phi(-\mathbf{q}, z) = \phi^*(\mathbf{q}, z)$ and, hence, the quantities $\phi(\mathbf{q}, z)$ and $\phi(-\mathbf{q}, z)$ are not independent. It is convenient to write the expression of $F_{\mathrm{tot}}(\boldsymbol{\phi})$ in terms of independent variables. For this purpose, a transition is usually made from the integration over the entire spectrum of $\mathbf{q}$ to the integration over a half-space [25] which will be conditionally denoted as $\mathbf{q} \geq 0$:

$$F_{\mathrm{tot}}(\boldsymbol{\phi}) = 2\int\limits_{\mathbf{q}\geq 0}\frac{d\mathbf{q}}{(2\pi)^2}\Phi_{\mathrm{tot}}(\boldsymbol{\phi}), \tag{3.2}$$

where

$$\Phi_{\mathrm{tot}}(\boldsymbol{\phi}) = \Phi_{bk}(\boldsymbol{\phi}) + \Phi_{sf}(\boldsymbol{\phi}).$$

In the Gaussian approximation, the fluctuational modes $\boldsymbol{\phi}(\mathbf{q}, z)$ for various $\mathbf{q} \geq 0$ are independent. Consequently, we can confine our analysis to a certain fixed $\mathbf{q}$. The probability density for fluctuations $\boldsymbol{\phi}(z) = \boldsymbol{\phi}(\mathbf{q}, z)$ is given by

$$\rho(\boldsymbol{\phi}) = \frac{1}{Z}\exp\left[-2\frac{\Phi_{\mathrm{tot}}(\boldsymbol{\phi})}{k_BT}\right], \tag{3.3}$$

where the partition function has the form

$$Z = \int\exp\left[-2\frac{\Phi_{\mathrm{tot}}(\boldsymbol{\phi})}{k_BT}\right]\mathcal{D}\boldsymbol{\phi}. \tag{3.4}$$

Here, the continual integration symbol $\int\ldots\mathcal{D}\boldsymbol{\phi}$ corresponds enumeration of all possible pairs of functions $\mathrm{Re}(\boldsymbol{\phi}(\mathbf{q}, z))$ and $\mathrm{Im}(\boldsymbol{\phi}(\mathbf{q}, z))$ for $l_1 \leq z \leq l_2$. Pay attention to coefficient 2 in Eqs. (3.2)–(3.4). It appears in connection with the integration over the half-space $\mathbf{q} \geq 0$ in (3.2).

The pair correlation function in this case can be expressed through the continual integral:

$$g_{\alpha\beta}(z, z_1) = \int\phi_\alpha\phi^*_\beta(z_1)\rho(\boldsymbol{\phi})\mathcal{D}\boldsymbol{\phi}. \tag{3.5}$$

Along with (3.5), a representation of $g_{\alpha\beta}$ in the form of a functional derivative is also used. For this purpose, the generating function

$$Z(\boldsymbol{\epsilon}) = \int\exp\left[-\frac{2\Phi_{\mathrm{tot}}(\boldsymbol{\phi}) + \mathcal{S}(\boldsymbol{\phi}; \boldsymbol{\epsilon})}{k_BT}\right]\mathcal{D}\boldsymbol{\phi}, \tag{3.6}$$

is introduced, where the source is given by

$$\mathcal{S}(\boldsymbol{\phi}; \boldsymbol{\epsilon}) = \int\limits_{l_1}^{l_2}[\boldsymbol{\phi}(z)\cdot\boldsymbol{\epsilon}^*(z) + \boldsymbol{\phi}^*(z)\cdot\boldsymbol{\epsilon}(z)]dz, \tag{3.7}$$

and the function $\boldsymbol{\epsilon}(\mathbf{q}, z) = \boldsymbol{\epsilon}^*(-\mathbf{q}, z)$ denotes a fictitious external field which is assumed to be equal to zero in the final results. In this case, the correlation function $\hat{g}$ can be calculated through the average value $\bar{\boldsymbol{\phi}}$ in the presence of field $\boldsymbol{\epsilon}$,

$$\begin{aligned}\bar{\boldsymbol{\phi}}(z, \boldsymbol{\epsilon}) &= \frac{1}{Z(\boldsymbol{\epsilon})}\\ &\times\int\boldsymbol{\phi}(z)\exp\left[-\frac{2\Phi_{\mathrm{tot}}(\boldsymbol{\phi}) + \mathcal{S}(\boldsymbol{\phi}; \boldsymbol{\epsilon})}{k_BT}\right]\mathcal{D}\boldsymbol{\phi}\end{aligned} \tag{3.8}$$

as follows:

$$g_{\alpha\beta}(z, z_1) = -k_BT\frac{\delta\bar{\phi}_\alpha(z, \boldsymbol{\epsilon})}{\delta\epsilon_\beta(z_1)}\bigg|_{\boldsymbol{\epsilon}=0}. \tag{3.9}$$

Here, we assume that the symbol $\delta/\delta\boldsymbol{\epsilon}$ of functional differentiation for complex-valued $\boldsymbol{\epsilon}$ in (3.9) has the following meaning:

$$\frac{\delta}{\delta\boldsymbol{\epsilon}} = \frac{1}{2}\left(\frac{\delta}{\delta\mathrm{Re}(\boldsymbol{\epsilon})} - i\frac{\delta}{\delta\mathrm{Im}(\boldsymbol{\epsilon})}\right). \tag{3.10}$$

The peculiarity of our problem is that $\Phi_{\mathrm{tot}}(\boldsymbol{\phi})$ contains the surface term $\Phi_{sf}$ which depends on a finite number ($2n$) of degrees of freedom $\boldsymbol{\phi}_1 = \boldsymbol{\phi}(l_1)$ and $\boldsymbol{\phi}_2 = \boldsymbol{\phi}(l_2)$. Consequently, it is natural to divide the degree of freedom of $\boldsymbol{\phi}$ into two parts and present integral (3.5) as a continual integral with fixed values of $\boldsymbol{\phi}(l_j) = \boldsymbol{\phi}_j$ ($j = 1, 2$) at the ends of the interval $[l_1, l_2]$, followed by finite-dimensional integration with respect to $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$:

$$\begin{aligned}g_{\alpha\beta}(\mathbf{q}; z, z_1) &= \frac{1}{Z}\int\exp\left[-\frac{2\Phi_{sf}(\boldsymbol{\phi}_{1,2})}{k_BT}\right]d\boldsymbol{\phi}_1 d\boldsymbol{\phi}_2\\ &\times\underbrace{\int\ldots\int}_{\phi(l_j)=\phi_j}\exp\left(-\frac{2\Phi_{bk} + \mathcal{S}}{k_BT}\right)\phi_\alpha(z)\phi^*_\beta(z_1)\mathcal{D}\boldsymbol{\phi}.\end{aligned} \tag{3.11}$$

The continual integral in this expression corresponds to severe boundary conditions. Following Feynman [26], we carry out in (3.11) a shift of the integration variable, presenting function $\boldsymbol{\phi}(z)$ in the form of the sum

$$\boldsymbol{\phi}(z) = \boldsymbol{\phi}_0(z) + \boldsymbol{\eta}(z), \tag{3.12}$$

where $\boldsymbol{\phi}_0$ is a fixed function satisfying inhomogeneous boundary condition of the first kind

$$\boldsymbol{\phi}_0(l_1) = \boldsymbol{\phi}_1, \quad \boldsymbol{\phi}_0(l_2) = \boldsymbol{\phi}_2 \tag{3.13}$$

and the Euler equation

$$\hat{\mathcal{L}}_E\boldsymbol{\phi}_0(z) = \boldsymbol{\epsilon}(z), \tag{3.14}$$

corresponding to the condition for the minimum of the exponent $2\Phi_{bk}(\boldsymbol{\phi}) = \mathcal{S}(\boldsymbol{\phi}; \boldsymbol{\epsilon})$ for $l_1 < z < l_2$. In this case, the boundary conditions for the new integration variable $\boldsymbol{\eta}(z)$ in the continual integral are homogeneous:

$$\boldsymbol{\eta}(l_1) = \boldsymbol{\eta}(l_2) = 0. \tag{3.15}$$

For energy (2.9), the differential operator $\hat{\mathcal{L}}_E$ is given by

$$\hat{\mathcal{L}}_E = \hat{a}\frac{\partial^2}{\partial z^2} + (\hat{b} + \hat{b}^+)\frac{\partial}{\partial z} - \hat{c}. \tag{3.16}$$

Considering that the function $\Phi_{bk}(\boldsymbol{\phi})$ is quadratic, while $\mathcal{S}(\boldsymbol{\phi}; \boldsymbol{\epsilon})$is linear in the variable $\boldsymbol{\phi}$, and taking into account the equilibrium condition $\delta(2\Phi_{bk} + \mathcal{S}) = 0$ and the boundary conditions (3.15), we can write

$$\begin{gathered} 2\Phi_{bk}(\boldsymbol{\phi}) + \mathcal{S}(\boldsymbol{\phi}; \boldsymbol{\epsilon}) \\ \equiv 2\Phi_{bk}(\boldsymbol{\phi}_0) + \mathcal{S}(\boldsymbol{\phi}; \boldsymbol{\epsilon}) + 2\Phi_{bk}(\boldsymbol{\eta}). \end{gathered}$$

Here, only the quantity $\boldsymbol{\phi}_0(z) = \boldsymbol{\phi}_0(\boldsymbol{\phi}_{1,2})$ on the right-hand side is a function of parameters $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$. Considering that the surface energy $\Phi_{sf}(\boldsymbol{\phi}_{1,2})$ is also only a function of these parameters, the integrals with respect to $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, and $\boldsymbol{\eta}$ in Eq. (3.11) can be factorized. As a result, the degrees of freedom $\boldsymbol{\eta}(z)$ and $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$ are found to be independent, and we have

$$\begin{gathered} \langle \phi_\alpha(z)\phi^*_\beta(z_1)\rangle \\ = \langle \phi_{0\alpha}(z)\phi^*_{0\beta}(z_1)\rangle_{\boldsymbol{\phi}_{1,2}} + \langle \eta_\alpha(z)\eta^*_\beta(z_1)\rangle_{\boldsymbol{\eta}}, \end{gathered} \tag{3.17}$$

where

$$\begin{gathered} \langle \phi_{0\alpha}(z)\phi^*_{0\beta}(z_1)\rangle_{\boldsymbol{\phi}_{1,2}} = \int \phi_{0\alpha}(z, \boldsymbol{\phi}_{1,2}) \\ \times \phi^*_{0\beta}(z_1, \boldsymbol{\phi}_{1,2})\rho(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2)d\boldsymbol{\phi}_1 d\boldsymbol{\phi}_2, \end{gathered} \tag{3.18}$$

$$\begin{gathered} \langle \eta_\alpha(z)\eta^*_\beta(z_1)\rangle_{\boldsymbol{\eta}} \\ = \underbrace{\int\ldots\int}_{\boldsymbol{\eta}(l_j) = 0} \eta_\alpha(z)\eta^*_\beta(z_1)\rho(\boldsymbol{\eta})\mathcal{D}\boldsymbol{\eta}, \end{gathered} \tag{3.19}$$

and the distribution functions are given by

$$\rho(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2) \propto \exp\left[-2\frac{\Phi_{sf}(\boldsymbol{\phi}_{1,2}) + \Phi_{bk}(\boldsymbol{\phi}_{1,2})}{k_B T}\right], \tag{3.20}$$

$$\rho(\boldsymbol{\eta}) \propto \exp\left[-2\frac{\Phi_{bk}(\boldsymbol{\eta})}{k_B T}\right]. \tag{3.21}$$

Here, the following notation has been introduced:

$$\Phi_{bk}(\boldsymbol{\phi}_{1,2}) = \Phi_{bk}(\boldsymbol{\phi}_0(\boldsymbol{\phi}_{1,2})).$$

In relations (3.18) and (3.19), we put $\boldsymbol{\epsilon} = 0$. The integral in (3.18) is a finite-dimensional integral of the Gaussian type and can be evaluated easily. The integral (3.19) is a Gaussian continual integral and in principle, can also be calculated explicitly. However, it is more convenient to calculate the correlation function (3.19) by using the relation

$$\langle \eta_\alpha(z)\eta^*_\beta(z_1)\rangle_{\boldsymbol{\eta}} = -k_B T\frac{\delta\phi_{0\alpha}(z)}{\delta\epsilon_\beta(z_1)}, \tag{3.22}$$

which is a version of relation (3.9) for zero boundary conditions $\boldsymbol{\eta}(l_1) = \boldsymbol{\eta}(l_2) = 0$. Pay attention to the fact that the derivative $\delta\phi_{0\alpha}/\delta\epsilon_\beta$ is independent of the quantities $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, and $\boldsymbol{\epsilon}$ since the dependence of the solution $\phi_0$ of the Euler equation (3.13), (3.14) on the parameters $\boldsymbol{\phi}_1$, $\boldsymbol{\phi}_2$, and $\boldsymbol{\epsilon}$ is linear.

It should be noted that the quantity $\Phi_{bk}(\boldsymbol{\phi}_{1,2})$ appearing in formula (3.20) can be expressed in terms of the values of $\boldsymbol{\phi}_0(z)$ and $\boldsymbol{\phi}'_0(z)$ at the boundaries $z = l_1$ and $z = l_2$. Indeed, if we use the Euler equation (3.14) in formula (2.9), the integration by parts for $\boldsymbol{\epsilon} = 0$ gives

$$\Phi_{bk}(\boldsymbol{\phi}_0) = \frac{1}{2}(\overset{+}{\boldsymbol{\phi}_0}\hat{a}\boldsymbol{\phi}'_0 + \overset{+}{\boldsymbol{\phi}_0}\hat{b}\boldsymbol{\phi}_0)\Big|_{l_1}^{l_2}. \tag{3.23}$$

Thus, the calculation of the correlation function $\hat{g}_{\alpha\beta}(\mathbf{q}; z, z_1)$ boils down to the solution of the Euler equation (3.14) with the boundary conditions of the first kind (3.13), followed by the evaluation of the finite-dimensional integral (3.18) and the functional derivative (3.22).

An alternative method for calculating the correlation matrix $\hat{g}(\mathbf{q}; z, z_1)$ is based on the direct evaluation of the functional derivative (3.9). This is possible since the average value $\bar{\boldsymbol{\phi}}$ in relation (3.8) for a random Gaussian quantity can be obtained by solving the Euler equation corresponding to the condition for the minimum of the total action

$$2\Phi_{\text{tot}}(\boldsymbol{\phi}) + \mathcal{S} = 2\Phi_{bk} + 2\Phi_{sf} + \mathcal{S}$$

in the field $\boldsymbol{\epsilon}$ of the source. In the equilibrium equation

$$2\Phi_{bk} + 2\Phi_{sf} + \delta\mathcal{S}/2 = 0$$

we integrate by parts the first, second, and third terms in expression (2.9) for $\Phi_{bk}$. On the interval $l_1 < z < l_2$, the equilibrium equation can be reduced to the form

$$\hat{\mathcal{L}}_E\bar{\boldsymbol{\phi}}(z) = \boldsymbol{\epsilon}(z), \tag{3.24}$$

which is identical to (3.14). In equilibrium, we must take into account at the boundaries $z = l_{1,2}$ the contribution of the surface energy $\Phi_{sf}$ as well as the nonintegral terms emerging as a result of integrating $\Phi_{bk}$ by parts. The requirement that the first variation of action must vanish for $z = l_2$ and $z = l_1$ leads to

$$\begin{cases} \hat{a}\bar{\boldsymbol{\phi}}'(l_2) + (\hat{b} + \hat{w}_2)\bar{\boldsymbol{\phi}}(l_2) = 0 \\ \hat{a}\bar{\boldsymbol{\phi}}'(l_1) + (\hat{b} + \hat{w}_1)\bar{\boldsymbol{\phi}}(l_1) = 0. \end{cases} \tag{3.25}$$

Thus, in order to find $\bar{\boldsymbol{\phi}} = \bar{\boldsymbol{\phi}}(\mathbf{q}, z, \boldsymbol{\phi})$, we must solve the Euler equation (3.24) with the boundary conditions (3.25) of the third kind.

The equation for the correlation function $g_{\alpha\beta}$ can also be obtained directly. To this end, it is sufficient to take the functional derivative of the Euler equation (3.24) and of the boundary conditions (3.25) with respect to $d\boldsymbol{\epsilon}(z_1)$. This gives

$$\begin{cases} \hat{\mathcal{L}}_E g(z, z_1) = -k_B T\delta(z - z_1)\hat{1} \\ \hat{a}\hat{g}'(l_2, z_1) + (\hat{b} + \hat{w}_2)\hat{g}(l_1, z_1) = 0 \\ \hat{a}\hat{g}'(l_1, z_1) + (\hat{b} - \hat{w}_1)\hat{g}(l_1, z_1) = 0 \end{cases} \tag{3.26}$$

for any fixed $z_1$: $l_1 \le z_1 \le l_2$. The derivatives in Eqs. (3.26) are taken with respect to the first argument of the function $\hat{g}(z, z_1)$. A system of equations of the type (3.26) was used in [20, 23].

## 4. EXPLICIT EXPRESSION FOR CORRELATION FUNCTION

The solution of the inhomogeneous Euler equation

$$\hat{\mathcal{L}}_E \boldsymbol{\phi}(z) = \boldsymbol{\epsilon}(z)$$

can be easily obtained by the method of variation of arbitrary constants. For this purpose, we introduce a set of independent solution

$$\boldsymbol{\phi}_i(z) = e^{\lambda_i z}\mathbf{e}^{(i)} \tag{4.1}$$

($i = 1, \ldots, 2n$) of the homogeneous equation

$$\hat{\mathcal{L}}_E \boldsymbol{\phi}(z) = 0.$$

Vectors $\mathbf{e}^{(i)}$ are the solution of the system of equations

$$[\hat{a}\lambda_i^2 + (\hat{b} - \hat{b}^+)\lambda_i - \hat{c}]\mathbf{e}^{(i)} = 0, \tag{4.2}$$

while numbers $\lambda_i$ are the roots of the algebraic equation

$$\det[\hat{a}\lambda^2 + (\hat{b} - \hat{b}^+)\lambda - \hat{c}] = 0. \tag{4.3}$$

The Hermitian nature of matrices $\hat{a}$ and $\hat{c}$ and the anti-Hermitian nature of matrix $\hat{b} - \hat{b}^+$ imply that if $\lambda$ is a root of Eq. (4.3), $-\lambda^*$ is also its root. Consequently, numbers $\lambda_i$ can be ordered as follows: $\lambda_{j+n} = -\lambda_j^*$, $j = 1, \ldots, n$. If we introduce the $n \times 2n$ matrix $\hat{\Phi}(z)$ of fundamental solutions with columns formed by solutions $\boldsymbol{\phi}_i(z)$, i.e.,

$$\hat{\Phi}(z) = (\boldsymbol{\phi}_1(z), \boldsymbol{\phi}_2(z), \ldots, \boldsymbol{\phi}_{2n}(z)), \tag{4.4}$$

the general solution of Eq. (3.14) for $l_1 \le z \le l_2$ may be written in the form

$$\boldsymbol{\phi}(z) = \hat{\Phi}(z)\mathbf{C}_0 + \int_{l_1}^{l_2} \hat{k}(z, z')\boldsymbol{\epsilon}(z')dz', \tag{4.5}$$

where the kernel $\hat{k}(z, z')$ is an $n \times n$ matrix:

$$\hat{k}(z, z') = \theta(z - z')\hat{\Phi}(z)\hat{\Psi}(z'), \tag{4.6}$$

while the $2n \times n$ matrix $\hat{\Psi}(z)$ is defined with the help of block matrices:

$$\hat{\Psi}(z) = \begin{pmatrix} \hat{a}\hat{\Phi}'(z) \\ \hat{\Phi}(z) \end{pmatrix}^{-1} \hat{R}_{10}. \tag{4.7}$$

Here and below, we use the following notation:

$$\hat{R}_{01} = \begin{pmatrix} \hat{0} \\ \hat{1} \end{pmatrix}, \quad \hat{R}_{10} = \begin{pmatrix} \hat{1} \\ \hat{0} \end{pmatrix}$$

are $2n \times n$ matrices, $\hat{1}$ is the $n$th order unit matrix, $\hat{0}$ is the $n$th order zero matrix, $\mathbf{C}_0$ is an arbitrary $2n \times 1$ constant column vector, and $\theta(z)$ is the Heaviside function. It should be noted that by virtue of the identity $\hat{\Phi}(z)\hat{\Psi}(z) = \hat{0}$, the function $\hat{k}(z, z_1)$ is continuous for $z = z_1$.

The equilibrium solution $\boldsymbol{\phi}_0(z)$satisfies the boundary conditions (3.13). Substituting Eq. (4.5) into (3.13), we determine the components of vector $\mathbf{C}_0$. This gives

$$\begin{aligned} \boldsymbol{\phi}_0(z) &= \hat{\Phi}(z)\hat{M}\begin{pmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \end{pmatrix} \\ &+ \int_{l_1}^{l_2} [\hat{k}(z, z') - \hat{\Phi}(z)\hat{M}\hat{R}_{01}\hat{k}(l_2, z')]\boldsymbol{\epsilon}(z')dz', \end{aligned} \tag{4.8}$$

where $\hat{M} = \hat{M}(l_1, l_2)$ is the $2n \times 2n$ matrix

$$\hat{M} = \begin{pmatrix} \hat{\Phi}(l_1) \\ \hat{\Phi}(l_2) \end{pmatrix}^{-1}. \tag{4.9}$$

Calculating the functional derivative in (3.22), we obtain

$$\begin{aligned} &\langle \eta_\alpha(z)\eta_\beta^*(z_1)\rangle_{\boldsymbol{\eta}} \\ &= k_B T[\hat{\Phi}(z)\hat{M}\hat{R}_{01}\hat{k}(l_2, z_1) - \hat{k}(z, z_1)]. \end{aligned} \tag{4.10}$$

It should be noted that the right-hand side of this equation vanishes for $z = l_{1,2}$ or $z_1 = l_{1,2}$. For $\boldsymbol{\epsilon} \longrightarrow 0$, we obtain from Eq. (4.8)

$$\begin{aligned} &\langle \phi_{0\alpha}(z)\phi_{0\beta}^*(z_1)\rangle_{\boldsymbol{\phi}_{1,2}} \\ &= \left[\hat{\Phi}(z)\hat{M}\left\langle \begin{pmatrix} \boldsymbol{\phi}_1 \\ \boldsymbol{\phi}_2 \end{pmatrix}(\boldsymbol{\phi}_1^*, \boldsymbol{\phi}_2^*)\right\rangle_{\boldsymbol{\phi}_{1,2}} \hat{M}^+\hat{\Phi}(z_1)\right]_{\alpha\beta}. \end{aligned} \tag{4.11}$$

The mean value in Eq. (4.11) is calculated by formula (3.18). Using Eq. (4.8) for $\boldsymbol{\epsilon} = 0$ in formulas (2.11) and (3.23), we can write the expression for $\Phi_{bk} + \Phi_{sf}$ as a quadratic form in the variables $\boldsymbol{\phi}_1$ and $\boldsymbol{\phi}_2$:

$$\Phi_{bk} + \Phi_{sf} = \frac{1}{2}(\boldsymbol{\phi}_1^*, \boldsymbol{\phi}_2^*)\hat{F}_1\begin{pmatrix}\boldsymbol{\phi}_1\\ \boldsymbol{\phi}_2\end{pmatrix}, \tag{4.12}$$

where

$$\begin{aligned}\hat{F}_1 &= [\hat{R}_{01}\hat{a}\hat{\Phi}'(l_2) + \hat{R}_{10}\hat{a}\hat{\Phi}'(l_1)]\hat{M}\\ &+ \begin{pmatrix} -\hat{b} + \hat{w}_1 & \hat{0}\\ \hat{0} & \hat{b} + \hat{w}_2\end{pmatrix}.\end{aligned} \tag{4.13}$$

is a $2n \times 2n$ matrix. Here, we have used the relations

$$\hat{M}^+\hat{\Phi}^+(l_2) = \hat{R}_{01}, \quad \hat{M}^+\hat{\Phi}^+(l_1) = \hat{R}_{10}, \tag{4.14}$$

which follow from definition (4.9) of matrix $\hat{M}$. Using the standard formula for Gaussian averaging (see, for example, [25]), we obtain the following result in terms of the elements of the reciprocal matrix:

$$\begin{aligned}&\langle\phi_{0\alpha}(z)\phi^*_{0\beta}(z_1)\rangle_{\phi_{1,2}}\\ &= k_BT[\hat{\Phi}(z)\hat{M}\hat{F}_1^{-1}\hat{M}^+\hat{\Phi}^+(z_1)]_{\alpha\beta}.\end{aligned} \tag{4.15}$$

Thus, the complete correlation function calculated by formula (3.17) has the form

$$\begin{aligned}\hat{g}(z, z_1) &= k_BT[\hat{\Phi}(z)\hat{M}\hat{F}_1^{-1}\hat{M}^+\hat{\Phi}^+(z_1)\\ &- \hat{k}(z, z_1) + \hat{\Phi}(z)\hat{M}\hat{R}_{01}\hat{k}(l_1, z_1)].\end{aligned} \tag{4.16}$$

In order to calculate $\hat{g}(z, z_1)$, we can also use an alternative approach based on the solution of the Euler equation (3.24) (identical to (3.13)) with the boundary conditions (3.25) of the third kind.

Substituting Eq. (4.5) into (3.25), we obtain a system of linear equations defining the vector $\mathbf{C}_0$ for these boundary conditions, which leads to

$$\mathbf{C}_0 = -\hat{F}_2^{-1}\hat{R}_{01}\hat{X}_2\int_{l_1}^{l_2}\hat{\Psi}(z')\boldsymbol{\epsilon}(z')dz', \tag{4.17}$$

where the $2n \times 2n$ matrix $\hat{F}_2$ is of the block type:

$$\hat{F}_2 = \begin{pmatrix}\hat{X}_1\\ \hat{X}_2\end{pmatrix}, \tag{4.18}$$

and

$$\hat{X}_j = \hat{a}\hat{\Phi}'(l_j) + (\hat{b} + (-1)^j\hat{w}_j)\hat{\Phi}(l_j) \tag{4.19}$$

are $n \times 2n$ matrices.

Substituting Eq. (4.17) into Eq. (4.5), we find that the dependence of $\bar{\boldsymbol{\phi}}$ on $\boldsymbol{\epsilon}$ is linear. The derivative $\delta/\delta\boldsymbol{\epsilon}$ in Eq. (3.9) can be evaluated trivially. Taking into account Eq. (4.6), we obtain the correlation function in the form

$$\begin{aligned}\hat{g}(z, z_1) &= k_BT[\hat{\Phi}(z)\hat{F}_2^{-1}\hat{R}_{01}\hat{X}_2\hat{\Psi}(z_1) - \hat{k}(z, z_1)]\\ &= k_BT\hat{\Phi}(z)[\hat{F} - \theta(z - z_1)\hat{I}]\hat{\Psi}(z_1),\end{aligned} \tag{4.20}$$

where

$$\hat{F} = \hat{F}_2^{-1}\hat{R}_{01}\hat{X}_2,$$

and $\hat{I}$ is a unit matrix of the order $2n$.

Carrying out simple but cumbersome calculations based on the properties of an analog of the Wronskian for operator (3.16), it can be proved that expressions (4.16) and (4.20) are identical. For the sake of definiteness, we will use formula (4.20) in the subsequent analysis. Further simplification can be carried out by taking into account the explicit expression for functions $\hat{\Phi}(z)$ in formula (4.4).

Let us introduce an $n \times 2n$ matrix $\hat{U}$ with components $U_{\alpha i} = e_\alpha^{(i)}$ and a diagonal matrix $\hat{\Lambda}$ of the $2n$th order with numbers $\lambda_i$ on the diagonal, $\Lambda_{ij} = \lambda_i\delta_{ij}$. Here, the summation over index $i$ is absent. We can express all the matrices in expression (4.20) in terms of the matrices $\hat{U}$ and $\hat{\Lambda}$:

$$\begin{gathered}\hat{\Phi}(z) = \hat{U}e^{\hat{\Lambda}z}, \quad \hat{\Psi}(z) = e^{-\hat{\Lambda}z}\hat{V},\\ \hat{V} = \begin{pmatrix}\hat{a}\hat{U}\hat{\Lambda}\\ \hat{U}\end{pmatrix}^{-1}\hat{R}_{10},\\ \hat{X}_j = (\hat{a}\hat{U}\hat{\Lambda} + (\hat{b} + (-1)^j\hat{w}_j)\hat{U})e^{\hat{\Lambda}l_j}.\end{gathered} \tag{4.21}$$

Substituting these expressions into (4.20), we obtain

$$\hat{g}(z, z_1) = k_BT\hat{U}e^{\hat{\Lambda}z}[\hat{F} - \theta(z - z_1)\hat{I}]e^{-\hat{\Lambda}z_1}\hat{V}. \tag{4.22}$$

It was shown above that numbers $\lambda_i$ are connected through conditions $\lambda_{i+n} = -\lambda_i^*$. In the general case of a non-Hermitian matrix $\hat{b} - \hat{b}^+$, there is no such simple relation for vectors $\mathbf{e}^{(i)}$. However (see Section 5), for all main geometries of SLC-*A* and NLC, matrix $\hat{b} - \hat{b}^+$ is symmetric and, hence, imaginary ($\hat{b} = 0$ for smectics *A*). It follows from Eq. (4.2) that $\mathbf{e}^{(i+n)} = \mathbf{e}^{(i)*}$ in this case. Consequently, we can write matrices $\hat{\Lambda}$ and $\hat{U}$ in the block form:

$$\hat{\Lambda} = \begin{pmatrix}\hat{\lambda} & \hat{0}\\ \hat{0} & -\hat{\lambda}^*\end{pmatrix}, \quad \hat{U} = (\hat{u}, \hat{u}^*), \tag{4.23}$$

where $\hat{\lambda}$ is a diagonal $n$th order matrix with numbers $\lambda_1, \ldots, \lambda_n$ on the diagonal and $\hat{u}$ is an $n \times n$ matrix with elements $u_{\alpha i} = e_{\alpha}^{(i)}$. It is convenient to assume that $\mathrm{Re}\lambda_i > 0$, $i = 1, \ldots, n$.

In this case, it can easily be verified that matrix $\hat{V}$ in (4.21) can be written in the form

$$\hat{V} = \hat{D}\hat{\mathcal{T}}\hat{U}^{+}, \tag{4.24}$$

where

$$\hat{\mathcal{T}} = \begin{pmatrix} \hat{0} & \hat{1} \\ \hat{1} & \hat{0} \end{pmatrix}, \quad \hat{D} = \begin{pmatrix} \hat{d} & 0 \\ 0 & -\hat{d}^{*} \end{pmatrix},$$
$$\hat{d} = (\hat{u}^{T}\hat{s}\hat{u})^{-1}, \quad \hat{s} = 2\hat{a}\mathrm{Re}(\hat{u}\hat{\lambda}\hat{u}^{-1}). \tag{4.25}$$

Here, $\hat{\mathcal{T}}$ is the transposition matrix, $\hat{\mathcal{T}}^{2} = \hat{\mathcal{T}}$, $\hat{d}$ and $\hat{s}$ are $n \times n$ matrices, $\hat{d}$ being a diagonal matrix and $\hat{s}$ a real-valued matrix. The superscript $T$ indicates transposition.

Substituting Eqs. (4.23)–(4.25) into Eqs. (4.21) and (4.22), we can write the correlation matrix in the form

$$\hat{g}(z, z_1) = k_B T[\hat{g}^{(1)}(z, z_1) + \hat{g}^{(2)}(z, z_1)], \tag{4.26}$$

where

$$\begin{aligned} \hat{g}^{(1)}(z, z_1) &= \hat{u}e^{\lambda z}[\hat{f}^{(11)} - \hat{1}\theta(z - z_1)]\hat{d}e^{-\hat{\lambda}z_1}\hat{u}^{T} \\ &- \hat{u}^{*}e^{-\hat{\lambda}^{*}z} + [\hat{f}^{(22)} + \hat{1}\theta(z - z_1)]\hat{d}^{*}e^{\hat{\lambda}^{*}z_1}\hat{u}^{+}, \\ \hat{g}^{(2)}(z, z_1) &= \hat{u}^{*}e^{-\hat{\lambda}^{*}z}\hat{f}^{(21)}\hat{d}e^{-\hat{\lambda}z_1}\hat{u}^{T} \\ &- \hat{u}e^{\hat{\lambda}z}\hat{f}^{(12)}\hat{d}^{*}e^{\hat{\lambda}^{*}z_1}\hat{u}^{+}. \end{aligned} \tag{4.27}$$

Here, $\hat{f}^{(ij)}$ are $n \times n$ matrices defined by the block representation

$$\hat{F} = \begin{pmatrix} \hat{f}^{(11)} & \hat{f}^{(12)} \\ \hat{f}^{(21)} & \hat{f}^{(22)} \end{pmatrix}. \tag{4.28}$$

Pay attention to the fact that in view of identity

$$\hat{u}\hat{d}\hat{u}^{T} = \hat{u}^{*}\hat{d}^{*}\hat{u}^{+}$$

function $\hat{g}^{(1)}(z, z_1)$ is continuous for $z = z_1$ in spite of the presence of discontinuous $\theta$ functions in its definition.

In order to determine the blocks $\hat{f}^{(ij)}$, we must know the explicit expression for matrix $\hat{F}_2^{-1}$. Since $\hat{F}_2$ is a complete matrix $2n \times 2n$ in the general case, the corresponding reciprocal matrix for a nematic ($n = 2$) is rather cumbersome.

In order to overcome this difficulty, it is convenient to use perturbation theory in the small parameters $e^{-|\lambda_j|L}$ for samples that are not very thin.

We choose the origin of coordinates so that $l_1 = -L/2$ and $l_2 = L/2$. In this case, matrices $\hat{X}_{1,2}$ in (4.19) can be written in the block form:

$$\hat{X}_j = (\hat{X}_j^{(1)}e^{\hat{\lambda}l_j}, \hat{X}_j^{(2)}e^{-\hat{\lambda}^{*}l_j}), \tag{4.29}$$

$j = 1, 2$, where

$$\begin{aligned} \hat{X}_j^{(1)} &= \hat{a}\hat{u}\hat{\lambda} + (\hat{b} + (-1)^{j}\hat{w}_j)\hat{u}, \\ X_j^{(2)} &= -\hat{a}\hat{u}^{*}\hat{\lambda}^{*} + (\hat{b} + (-1)^{j}\hat{w}_j)\hat{u}^{*}. \end{aligned} \tag{4.30}$$

In the principal order in parameters $e^{-|\lambda_j|L}$, we obtain from (4.18), (4.28), and (4.29)

$$\begin{aligned} \hat{f}^{(11)} &= \hat{1}, \quad \hat{f}^{(22)} = \hat{0}, \\ \hat{f}^{(12)} &= e^{-\hat{\lambda}L/2}(\hat{X}_2^{(1)})^{-1}\hat{X}_2^{(2)}e^{-\hat{\lambda}^{*}L/2}, \\ \hat{f}^{(21)} &= -e^{-\hat{\lambda}^{*}L/2}(\hat{X}_1^{(2)})^{-1}\hat{X}_1^{(1)}e^{-\hat{\lambda}L/2}. \end{aligned} \tag{4.31}$$

The next corrections to matrices $\hat{f}^{(11)}$ and $\hat{f}^{(22)}$ are of the order of $e^{-2\lambda_j L}$ and to matrices $\hat{f}^{(12)}$ and $\hat{f}^{(21)}$ are of the order of $e^{-3\lambda_j L}$.

Substituting formulas (4.31) into (4.27), we obtain

$$\begin{aligned} \hat{g}^{(1)}(z, z_1) &= \hat{u}\hat{d}e^{\hat{\lambda}(z - z_1)}\hat{u}^{T}\theta(z_1 - z) \\ &+ \hat{u}^{*}\hat{d}^{*}e^{-\hat{\lambda}^{*}(z - z_1)}\hat{u}^{+}\theta(z - z_1), \\ \hat{g}^{(2)}(z, z_1) &= -\hat{u}^{*}e^{-\hat{\lambda}^{*}(L/2 + z)}\hat{Y}_1 e^{-\hat{\lambda}(L/2 + z_1)}\hat{u}^{T} \\ &- \hat{u}e^{\hat{\lambda}(z - L/2)}\hat{Y}_2 e^{\hat{\lambda}^{*}(z_1 - L/2)}\hat{u}^{+}, \end{aligned} \tag{4.32}$$

where

$$\begin{aligned} \hat{Y}_1 &= (\hat{X}_1^{(2)})^{-1}\hat{X}_1^{(1)}\hat{d}, \\ \hat{Y}_2 &= (\hat{X}_2^{(1)})^{-1}\hat{X}_2^{(2)}\hat{d}^{*} \end{aligned} \tag{4.33}$$

are Hermitial $n \times n$ matrices. It should be noted that as $L \longrightarrow \infty$, the term $\hat{g}^{(2)}$ in (4.32) tends to zero and, hence, $\hat{g} = \hat{g}^{(1)}$. Thus, the term $\hat{g}^{(1)}$ in Eq. (4.32) describes fluctuations in an unbounded medium in **q**, $z$-representation, while the term $\hat{g}^{(2)}$ describes the corrections associated with the boundedness of the sample and with anchoring to the surface.

## 5. CORRELATION FUNCTION OF FLUCTUATIONS IN SMECTICS *A* AND NEMATICS

Let us apply the general formulas derived above for determining correlation functions in specific cases.

### *5.1. Fluctuations in Smectics A*

For the elastic component of energy from formula (2.3), we have

$$\Phi_e = \frac{1}{2}\int dz\{B|\partial_z u_{\mathbf{q}}(z)|^2 + Kq^4|u_{\mathbf{q}}(z)|^2\}. \qquad (5.1)$$

In accordance with relation (2.4), the contribution of the external field in the **r**-representation has the form

$$F_f = \frac{1}{2}\chi_a H^2[1-(\boldsymbol{\nabla}_\perp u)^2],$$

for $\chi_a > 0$ and

$$F_f = -\frac{1}{2}\chi_a H^2\left(\frac{\partial u}{\partial x}\right)^2$$

for $\chi_a < 0$. In the latter case, we have directed the *x* axis along the external field $\mathbf{H} = (H, 0, 0)$.

The corresponding quantities $\Phi_f$ have the form

$$\Phi_f = \begin{cases} \frac{1}{2}\chi_a H^2 q^2\int dz|u_{\mathbf{q}}(z)|^2, & \chi_a > 0 \\ -\frac{1}{2}\chi_a H^2 q_1^2\int dz|u_{\mathbf{q}}(z)|^2, & \chi_a < 0, \end{cases} \qquad (5.2)$$

where $\mathbf{q} = (q_1, q_2, 0)$.

It follows from formula (2.7) that the surface energy density is given by

$$\Phi_{sf} = \frac{1}{2}\sum_{j=1,2}(\gamma_1^{(j)} + q^2\gamma_2^{(j)})|u_{\mathbf{q}}(l_j)|^2. \qquad (5.3)$$

In the given case, matrices $\hat{a}$, $\hat{b}$, $\hat{c}$, and $\hat{w}_j$ ($j = 1, 2$) in expression (2.9) are one-dimensional, i.e., are scalars:

$$a = B, \quad b = 0, \quad w_j = \gamma_1^{(j)} + \gamma_2^{(j)}q^2,$$
$$c = \begin{cases} Kq^4 + \chi_a H^2 q^2, & \chi_a > 0 \\ Kq^4 - \chi_a H^2 q_1^2, & \chi_a < 0. \end{cases} \qquad (5.4)$$

The characteristic equation is quadratic and has the roots

$$\lambda_1 = -\lambda_2 = \sqrt{\frac{c}{B}}. \qquad (5.5)$$

"Vectors" $\mathbf{e}^{(j)}$ are also scalars and can be put equal to unity: $e^{(1)} = 1$, $e^{(2)} = 1$. The "matrices" are $\hat{\lambda} = \lambda$, $\hat{u} = 1$, $d = (2B\lambda)^{-1}$. This gives

$$\hat{V} = \frac{1}{2B\lambda}\begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

$$\hat{X}_1 = \left((B\lambda - w_1)e^{-\lambda L/2}, -(B\lambda + w_1)e^{\lambda L/2}\right), \qquad (5.6)$$

$$\hat{X}_2 = \left((B\lambda + w_2)e^{\lambda L/2}, (-B\lambda + w_2)e^{-\lambda L/2}\right).$$

Substituting these relations into Eqs. (4.28)–(4.31), we obtain the correlation function (4.26), (4.27) in the form

$$\begin{aligned} g(\mathbf{q}; z, z_1) = \frac{k_B T}{2B\lambda\Delta}\{&(B^2\lambda^2 - w_1 w_2)\cosh[(z + z_1)\lambda] \\ &+ B\lambda(w_1 - w_2)\sinh[(z + z_1)\lambda] \\ &+ (B^2\lambda^2 + w_1 w_2)\cosh[(L - |z - z_1|)\lambda] \\ &+ B\lambda(w_1 + w_2)\sinh[(L - |z - z_1|)\lambda]\}, \end{aligned} \qquad (5.7)$$

where

$$\begin{aligned} \Delta = {} & (B^2\lambda^2 + w_1 w_2)\sinh(\lambda L) \\ & + B(w_1 + w_2)\lambda\cosh(\lambda L). \end{aligned}$$

Formula (5.7) describes both cases ($\chi_a > 0$ and $\chi_a < 0$).

Let us single out two modes of the behavior of a liquid crystal at the boundary that can be referred to as strong and weak anchoring. In the case of strong anchoring, fluctuations near the surface are suppressed. If, however, the energy of anchoring with the given surface is low, fluctuations in the vicinity of this boundary may be stronger than in the bulk. A qualitative reason behind this effect is that the retrieving force at the surface is determined by the surface as well as bulk contributions, the role of bulk forces at the boundary being smaller than in the bulk of the sample.

This is illustrated in Fig. 1. If we take the simplest case $\gamma_1^{(1,2)} = 0$, $\gamma_2^{(1)} = \gamma_2^{(2)} = \gamma$, a transition from strong to weak anchoring occurs for $\gamma \sim \sqrt{KB}$.

Let us analyze the effect of an external field on the correlation function for SLC-*A*.

The conditions $\lambda L \sim 1$ and $\chi_a H^2 \sim Kq^2$ lead to the characteristic value of the field in SLC-*A*:

$$H_S \sim (KB)^{1/4}(\chi_a L)^{-1/2}.$$

For the typical values $K \sim 10^{-6}$ dyne, $B \approx 3 \times 10^7$ dyn/cm², $\chi_a \sim 10^{-7}$ for $L \approx 3 \times 10^{-3}$ cm, this gives $H_S \sim 10^5$ G. It can be seen that the characteristic field for smectics is quite large.

Let us now consider nematics.

### 5.2. Homeotropic Orientation, $\chi_a > 0$

In this case, the equilibrium director vector $\mathbf{n}^0$ is parallel to the vector of the external field strength **H** and directed along the $z$ axis, $\mathbf{n}^0 = (0, 0, 1)$, $\mathbf{H} = (0, 0, H)$. The director fluctuation vector is given by

$$\delta\mathbf{n}(\mathbf{q}, z) = (\delta n_{1\mathbf{q}}(z), \delta n_{2\mathbf{q}}(z), 0).$$

Following [1], we choose the $x$ axis along the wave vector $\mathbf{q} = (q, 0, 0)$. Formula (2.2) for Frank's energy gives

$$\Phi_e = \frac{1}{2}\int dz[q^2(K_{11}|\delta n_{1\mathbf{q}}|^2 + K_{22}|\delta n_{2\mathbf{q}}|^2) + K_{33}(|\partial_z\delta n_{1\mathbf{q}}|^2 + |\partial_z\delta n_{2\mathbf{q}}|^2)]. \tag{5.8}$$

The contribution of the external field is

$$\Phi_f = \frac{1}{2}\chi_a H^2\int dz(|\delta n_{1\mathbf{q}}|^2 + |\delta n_{2\mathbf{q}}|^2). \tag{5.9}$$

This gives the following expressions for matrices $\hat{a}$, $\hat{b}$, and $\hat{c}$:

$$\hat{a} = K_{33}\hat{1}, \quad \hat{b} = \hat{0},$$

$$\hat{c} = \begin{pmatrix} K_{11}q^2 + \chi_a H^2 & 0 \\ 0 & K_{22}q^2 + \chi_a H^2 \end{pmatrix}. \tag{5.10}$$

For the surface energy in this geometry, model (2.5) is commonly used, in which matrices $\hat{w}_1$ and $\hat{w}_2$ in (2.11) are proportional to the unit matrix:

$$\hat{w}_1 = W_1\hat{1}, \quad \hat{w}_2 = W_2\hat{1}. \tag{5.11}$$

Thus, we can write

$$\Phi_{sf} = \frac{1}{2}\sum_{i,\,j\,=\,1,\,2}|\delta n_{i\mathbf{q}}(l_j)|^2 W_j. \tag{5.12}$$

It should be noted that if matrices $\hat{w}_1$ and $\hat{w}_2$ are not proportional to the unit matrix, they are nondiagonal and depend on the direction of vector **q** in the general case.

The roots of the corresponding characteristic equation are given by

$$\lambda_i = -\lambda_{i+2} = \sqrt{\frac{K_{ii}q^2 + \chi_a H^2}{K_{33}}}. \tag{5.13}$$

The eigenvectors have the form

$$\mathbf{e}^{(1)} = \mathbf{e}^{(3)} = \begin{pmatrix}1\\0\end{pmatrix}, \quad \mathbf{e}^{(2)} = \mathbf{e}^{(4)} = \begin{pmatrix}0\\1\end{pmatrix}. \tag{5.14}$$

In this case, matrix $\hat{u}$ from Eq.(4.23) is a unit matrix ($\hat{u} = \hat{1}$), while matrix $\hat{d}$ in Eq. (4.25) is given by

$$\hat{d} = \frac{1}{2K_{33}}\begin{pmatrix}\lambda_1^{-1} & 0 \\ 0 & \lambda_2^{-1}\end{pmatrix}. \tag{5.15}$$

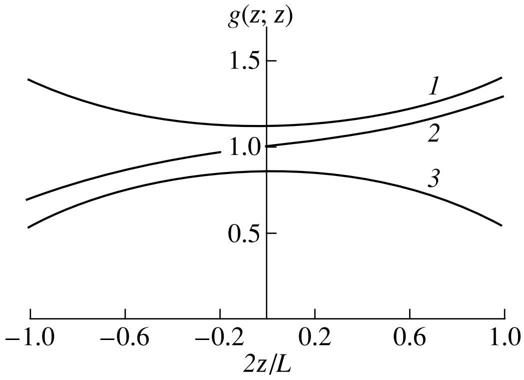

**Fig. 1.** Behavior of the correlation function $g(z, z)$ in SLC-$A$ for various values of surface tension. The following parameters are used: $\gamma_1^{(1,2)} = 0$, $B = 3 \times 10^7$ dyn/cm$^2$, $K = 10^{-6}$ dyn, $q = 10^5$ cm$^{-1}$, $L = 10^{-3}$ cm, $T = 300$ K, and $H = 0$. Curve *1* corresponds to weak anchoring at both surfaces, $\gamma_2^{(1)} = \gamma_2^{(2)} = 2.5$ erg/cm$^2$; curve *2* corresponds to strong anchoring at the surface $z = -L/2$, $\gamma_2^{(1)} = 10$ erg/cm$^2$ and weak anchoring at the surface $z = L/2$, $\gamma_2^{(2)} = 3$ erg/cm$^2$; and curve *3* corresponds to strong anchoring at both surfaces, $\gamma_2^{(1)} = \gamma_2^{(2)} = 15$ erg/cm$^2$. All the curves are normalized to $g(0, 0)$ for $\gamma_2^{(1)} = \gamma_2^{(2)} = 15$ erg/cm$^2$.

The block matrices $\hat{f}^{(ij)}$ can easily be evaluated and are diagonal:

$$\begin{aligned}\hat{f}^{(11)} &= \widehat{\mathrm{diag}}(\varphi_j^+/\Delta_j),\\ \hat{f}^{(22)} &= \widehat{\mathrm{diag}}(\varphi_j^-/\Delta_j),\\ \hat{f}^{(12)} &= \widehat{\mathrm{diag}}(\psi_j^-/\Delta_j),\\ \hat{f}^{(22)} &= \widehat{\mathrm{diag}}(\psi_j^+/\Delta_j),\end{aligned} \tag{5.16}$$

where the symbol $\widehat{\mathrm{diag}}(\mu_j)$ denotes a diagonal matrix with the elements $\mu_1, \mu_2, \ldots$ on the diagonal,

$$\begin{aligned}\varphi_j^{\pm} &= \frac{1}{2}[\pm(K_{33}^2\lambda_j^2 + W_1W_2) + K_{33}\lambda_j(W_1 + W_2)]e^{\pm\lambda_j L},\\ \psi_j^{\pm} &= \frac{1}{2}[\pm(K_{33}^2\lambda_j^2 - W_1W_2) + K_{33}\lambda_j(W_2 - W_1)],\\ \Delta_j &= [(K_{33}^2\lambda_j^2 + W_1W_2)\sinh(\lambda_j L) + K_{33}\lambda_j(W_1 + W_2)\cosh(\lambda_j L)].\end{aligned} \tag{5.17}$$

Substituting relations (5.14)–(5.17) into relations (4.26) and (4.27), we obtain the correlation matrix for director fluctuations, which is a diagonal matrix in the present case ($g_{12} = g_{21} = 0$):

$$g_{jj}(\mathbf{q}; z, z_1) = \frac{k_B T}{2K_{33}}\mathcal{E}_j(\mathbf{q}, z, z_1), \qquad (5.18)$$
$$j = 1, 2,$$

where

$$\mathcal{E}_j(\mathbf{q}, z, z_1) = \frac{1}{\lambda_j \Delta_j}$$
$$\times\{(K_{33}^2\lambda_j^2 + W_1W_2)\cosh[\lambda_j(L - |z - z_1|)]$$
$$+ K_{33}\lambda_j(W_1 + W_2)\sinh[\lambda_j(L - |z - z_1|)] \qquad (5.19)$$
$$+ (K_{33}^2\lambda_j^2 - W_1W_2)\cosh[\lambda_j(z + z_1)]$$
$$+ K_{33}\lambda_j(W_1 - W_2)\sinh[\lambda_j(z + z_1)]\}.$$

The characteristic value of the external magnetic field in NLC is $H_N \approx L^{-1}(K_{ii}/\chi_a)^{1/2}$ [1]. For the typical values $K \sim K_{ii} \sim 10^{-6}$ dyn and $\chi_a \sim 10^{-7}$, this gives $H_N \sim 10^3$ G for $L \approx 3 \times 10^{-3}$ cm. Consequently, in contrast to SLC-*A*, the effect of the field may be significant for the field strengths $H_N \sim 10^3$–$10^4$ G conventionally used in experiments.

The realistic values of the anchoring energy lie in the interval $W \sim 10^{-2}$–$10^{-4}$ erg/cm$^2$ [27, 28]. For large values of $q \sim 10^4$–$10^5$ cm$^{-1}$, weak anchoring will take place for both $g_{jj}$ modes. If, however, $q \sim 10^3$ cm$^{-1}$, strong anchoring may take place, and field effects also become significant.

### *5.3. Homeotropic Orientation, $\chi_a < 0$*

In this case, the external field vector **H** is perpendicular to the director $\mathbf{n}^0 = (0, 0, 1)$. We choose the *x* axis along $\mathbf{H} = (H, 0, 0)$. With such a choice of the axes, the director fluctuations and the wave vector have the form $\delta\mathbf{n} = (\delta n_1, \delta n_2, 0)$ and $\mathbf{q} = (q_1, q_2, 0)$.

The Frank energy is given by

$$\Phi_e = \frac{1}{2}\int dz[(K_{11}q_1^2 + K_{22}q_2^2)|\delta n_{1\mathbf{q}}|^2$$
$$+ (K_{11}q_2^2 + K_{22}q_1^2)|\delta n_{2\mathbf{q}}|^2 \qquad (5.20)$$
$$+ q_1q_2(K_{11} - K_{22})(\delta n_{1\mathbf{q}}\delta n_{2\mathbf{q}}^* + \delta n_{1\mathbf{q}}^*\delta n_{2\mathbf{q}})$$
$$+ K_{33}(|\partial_z\delta n_{1\mathbf{q}}|^2 + |\partial_z\delta n_{2\mathbf{q}}|^2)],$$

and the contribution of the external field has the form

$$\Phi_f = -\frac{1}{2}\chi_a H^2\int dz|\delta n_{1\mathbf{q}}|^2. \qquad (5.21)$$

The surface energy has the same form as (5.12) for $\chi_a > 0$.

Thus, in this geometry, we obtain the same matrices $\hat{a}$, $\hat{b}$, and $\hat{w}_j$ as in (5.10) and (5.11), while matrix $\hat{c}$ has the form

$$\hat{c} = \begin{pmatrix} A_2 & B \\ B & A_1 \end{pmatrix}, \qquad (5.22)$$

where

$$A_1 = K_{11}q_2^2 + K_{22}q_1^2, \quad B = q_1q_2(K_{11} - K_{22}),$$
$$A_2 = K_{11}q_1^2 + K_{22}q_2^2 - \chi_a H^2. \qquad (5.23)$$

The characteristic equation has the real roots

$$\lambda_i = -\lambda_{i+2}$$
$$= \left[\frac{1}{2K_{33}}((K_{11} + K_{22})q^2 - \chi_a H^2 + (-1)^{i+1}Q)\right]^{1/2}, \qquad (5.24)$$

where

$$Q = [q^4(K_{11} - K_{22})^2$$
$$- 2\chi_a H^2(K_{11} - K_{22})(q_1^2 - q_2^2) + \chi_a^2 H^4]^{1/2},$$

$i = 1, 2$. Except for the normalization, vectors $\mathbf{e}^{(i)}$ are defined as

$$\mathbf{e}^{(i)} = \mathbf{e}^{(i+2)} = \begin{pmatrix} B \\ K_{33}\lambda_i^2 - A_2 \end{pmatrix}. \qquad (5.25)$$

We could carry out calculations on the basis of formulas (4.26) and (4.27). However, the final result can be obtained much more simply. It should be noted that vectors $\mathbf{e}^{(1)}$ and $\mathbf{e}^{(2)}$ in (5.25) are orthogonal and real-valued. Consequently, matrix $\hat{u}$ in (4.23)) is orthogonal: $\hat{u}^{-1} = \hat{u}^T$. After the transition to the coordinate system with axes $\mathbf{e}^{(1)}$ and $\mathbf{e}^{(2)}$ in the $(x, y)$ plane, matrix $\hat{c}$ is diagonalized. Since matrices $\hat{w}_1$ and $\hat{w}_2$ are proportional to unit matrices in this geometry, they do not change upon a transition to the new system of coordinates. In this case, problem (3.26) for $\chi_a < 0$ (as well as the differential equation and the boundary conditions) formally coincides with the corresponding problem for $\chi_a > 0$. The only difference is that $\lambda_i$ in this case is defined by formulas (5.24) instead of formulas (5.13). Returning to the initial system of coordinates, we can write the correlation function for $\chi_a < 0$ in the form

$$\hat{g}_{\chi_a < 0} = \hat{u}\hat{g}_{\chi_a > 0}\hat{u}^{-1}, \qquad (5.26)$$

where $\hat{g}_{\chi_a > 0}$ is a diagonal matrix with the components $g_{jj}$ defined in (5.18).

Substituting relations (5.23) and (5.25) into (5.26), we obtain

$$g_{ll} = \frac{k_B T}{2QK_{33}} \sum_{j=1,2} (-1)^j (A_l - \lambda_j^2 K_{33}) \mathscr{E}_j(\mathbf{q}; z, z_1),$$
$$l = 1, 2, \tag{5.27}$$
$$g_{12} = g_{21} = \frac{k_B T B}{2QK_{33}} \sum_{j=1,2} (-1)^{j+1} \mathscr{E}_j(\mathbf{q}; z, z_1).$$

Systems with $\chi_a < 0$ are studied less thoroughly than those with $\chi_a > 0$. Such systems are interesting for our analysis in that the external field **H** suppresses only one of the fluctuating modes in them, while the other mode remains finite in the limit $H \longrightarrow \infty$ [24]. Having chosen a strong field for such a system with a small $q$, we may obtain weak anchoring at the boundary for one mode and strong anchoring for the other mode. These effects are illustrated in Figs. 2a and 2b. It should also be noted that the correlation function contains a crossed purely imaginary component $g_{12}(\mathbf{q}; z, z_1)$ presented in Fig. 2c.

### 5.4. Planar Orientation with $\chi_a > 0$

Let us direct the $y$ axis along the director ($\mathbf{n}^0 \parallel \mathbf{H}$). The director fluctuations and the wave vector in this system of coordinates have the form $\delta\mathbf{n} = (\delta n_1, 0, \delta n_2)$ and $\mathbf{q} = (q_1, q_2, 0)$.

The Frank energy is given by

$$\begin{aligned}\Phi_e = \frac{1}{2}\int dz\{&(K_{11}q_1^2 + K_{33}q_2^2)|\delta n_{1\mathbf{q}}|^2 \\ &+ K_{22}|\partial_z \delta n_{1\mathbf{q}}|^2 + (K_{22}q_1^2 + K_{33}q_2^2)|\delta n_{2\mathbf{q}}|^2 \\ &+ K_{11}|\partial_z \delta n_{2\mathbf{q}}|^2 + iq_1[K_{11}(\delta n_{1\mathbf{q}}\partial_z \delta n_{2\mathbf{q}}^* - \delta n_{1\mathbf{q}}^* \partial_z \delta n_{2\mathbf{q}}) \\ &- K_{22}(\delta n_{2\mathbf{q}} \partial_z \delta n_{1\mathbf{q}}^* - \delta n_{2\mathbf{q}}^* \partial_z \delta n_{1\mathbf{q}})]\}.\end{aligned} \tag{5.28}$$

The contribution of the external field is

$$\Phi_f = \frac{1}{2}\chi_a H^2 \int dz(|\delta n_{1\mathbf{q}}|^2 + |\delta n_{2\mathbf{q}}|^2). \tag{5.29}$$

The surface energy (2.6) in the given geometry can be written in the form

$$\Phi_{sf} = \frac{1}{2}\sum_{i,j=1,2} W_j^{(i)} |\delta n_{i\mathbf{q}}(l_j)|^2. \tag{5.30}$$

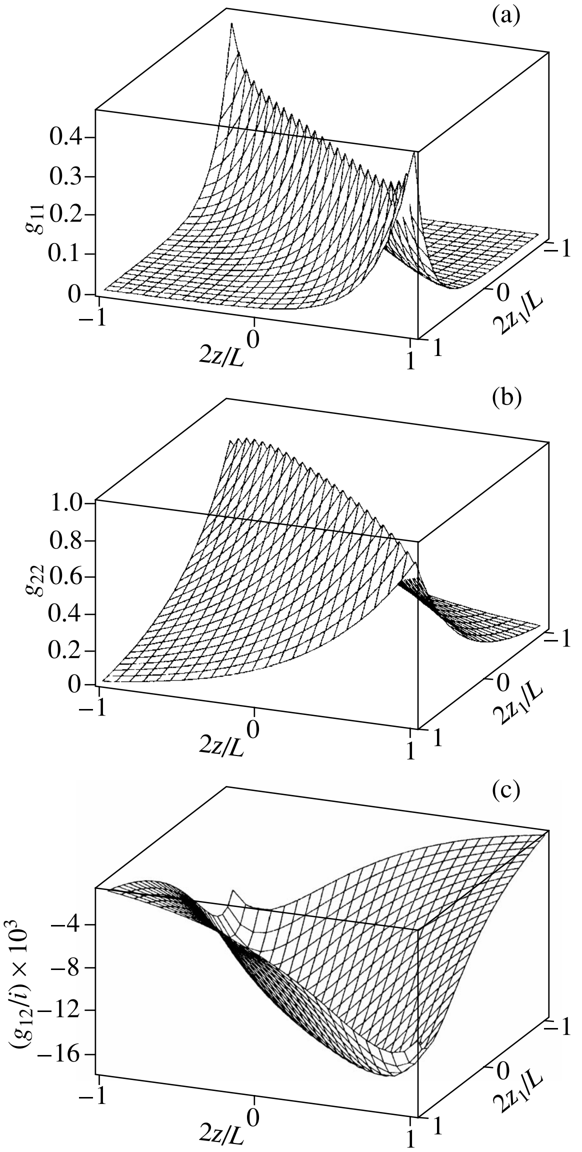

**Fig. 2.** Elements of the correlation matrix for NLC in the case of the homeotropic orientation with $\chi_a < 0$. The following parameters are used: $\chi_a = -10^{-7}$, $K_{11} = 10^{-6}$ dyn, $K_{22} = 0.5 \times 10^{-6}$ dyn, $K_{33} = 2 \times 10^{-6}$ dyn, $T = 300$ K, $H = 10^4$ G, $q_1 = 10^3$ cm$^{-1}$, $q_2 = 2 \times 10^3$ cm$^{-1}$, $W_1 = 5 \times 10^{-3}$ erg/cm$^2$, $W_2 = 4 \times 10^{-3}$ erg/cm$^2$, $L = 10^{-2}$ cm. (a) $g_{11}(z, z_1)$; (b) $g_{22}(z, z_1)$, and (c) $g_{12}(z, z_1)/i$. The curves are normalized to $g_{22}(0, 0)$.

Thus, matrices $\hat{a}$, $\hat{b}$, $\hat{c}$, and $\hat{w}_j$ ($j$ = 1, 2) in this geometry are given by

$$\hat{a} = \begin{pmatrix} K_{22} & 0 \\ 0 & K_{11} \end{pmatrix}, \quad \hat{b} = iq_1\begin{pmatrix} 0 & -K_{22} \\ K_{11} & 0 \end{pmatrix},$$
$$\hat{w}_j = \begin{pmatrix} W_j^{(1)} & 0 \\ 0 & W_j^{(2)} \end{pmatrix}, \quad \hat{c} = \begin{pmatrix} A_3 & 0 \\ 0 & A_4 \end{pmatrix}, \tag{5.31}$$

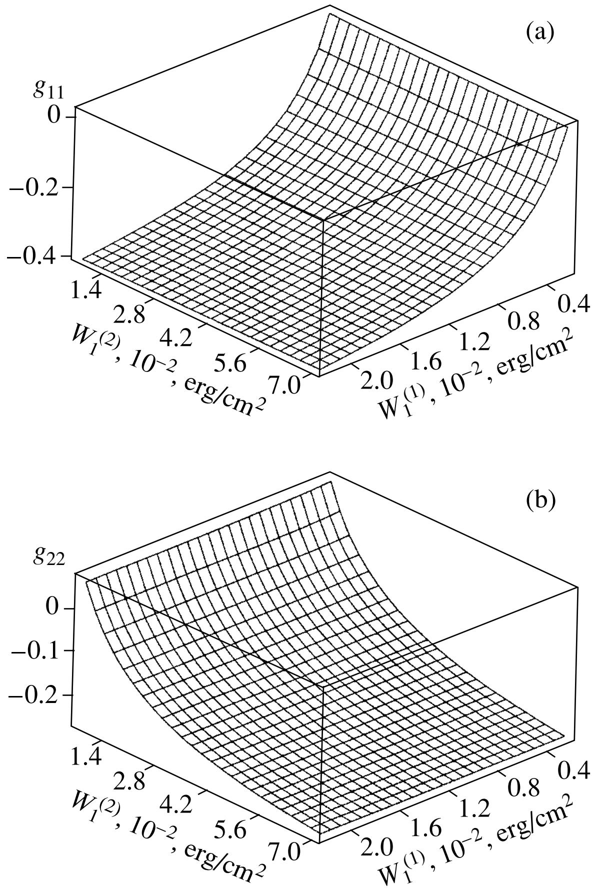

**Fig. 3.** Elements $g_{11}$ and $g_{22}$ of the correlation matrix (in arbitrary units) for NLC in the case of the planar orientation with $\chi_a > 0$ for $z = z_1 = L/2$ as functions of $W_1^{(1)} = W_2^{(1)}$ and $W_1^{(2)} = W_2^{(2)}$ for $H = 0$, $L = 10^{-2}$ cm. The remaining parameters are the same as in Fig. 2. (a) $g_{11}(L/2, L/2)$, (b) $g_{22}(L/2, L/2)$.

where

$$\begin{aligned} A_3 &= K_{11}q_1^2 + K_{33}q_2^2 + \chi_a H^2, \\ A_4 &= K_{22}q_1^2 + K_{33}q_2^2 + \chi_a H^2. \end{aligned} \tag{5.32}$$

The characteristic equation (4.3) has the following roots:

$$\lambda_1 = -\lambda_3 = \sqrt{\frac{A_3}{K_{11}}}, \quad \lambda_2 = -\lambda_4 = \sqrt{\frac{A_4}{K_{22}}}. \tag{5.33}$$

In this case, vectors $\mathbf{e}^{(i)}$ are identical (to within normalization):

$$\begin{aligned} \mathbf{e}^{(1)} &= \mathbf{e}^{(3)*} = \begin{pmatrix} q_1 \\ -i\lambda_1 \end{pmatrix}, \\ \mathbf{e}^{(2)} &= \mathbf{e}^{(4)*} = \begin{pmatrix} -i\lambda_2 \\ q_1 \end{pmatrix}. \end{aligned} \tag{5.34}$$

In this case, we have

$$\hat{u} = \begin{pmatrix} q_1 & i\lambda_2 \\ -i\lambda_1 & q_1 \end{pmatrix}. \tag{5.35}$$

Then matrix $\hat{d}$ in (4.25) is given by

$$\hat{d} = -\frac{1}{2\Omega}\begin{pmatrix} \lambda_1^{-1} & 0 \\ 0 & \lambda_2^{-1} \end{pmatrix}, \tag{5.36}$$

where $\Omega = K_{33}q_2^2 + \chi_a H^2$.

In contrast to the homeotropic geometry, matrix $\hat{F}_2$ in (4.18) has no special structure, and the components of the reciprocal matrix $\hat{F}_2^{-1}$ in relation (4.20) (and, hence, of matrices $\hat{F}^{(ij)}$ in relation (4.28)) are rather cumbersome. For this reason, it is convenient to use for numerical calculations the general expressions (4.26) and (4.27) for the correlation matrix.

It should be noted that in the planar orientation, coefficients $W_{1,2}^{(1)}$ and $W_{1,2}^{(2)}$ determine the contribution of modes $\delta n_1$ and $\delta n_2$, respectively, to the surface energy. We can expect that component $g_{11}$ is mainly determined by coefficient $W_j^{(1)}$ and is almost independent of $W_{1,2}^{(2)}$, while component $g_{22}$ is determined by coefficient $W_j^{(2)}$ and is almost independent of $W_{1,2}^{(1)}$. This effect, however, is masked by the interaction of modes $\delta n_1$ and $\delta n_2$ in the bulk energy. Nevertheless, for realistic values of parameters, the dependence of $g_{ll}$ on $W_j^{(l)}$ is observed to a high degree of accuracy. It can be seen from Fig. 3 that the value of $g_{ll}$ is indeed determined only by $W_j^{(l)}$. This dependence for $z = z_1 = \pm L/2$ is approximately defined by the formula

$$\begin{aligned} g_{ll}^{(2)}(\pm L/2, \pm L/2) &= -k_B T\frac{\lambda_1\lambda_2 - q_1^2}{2\lambda_l\Omega} \\ &\times \frac{W_j^{(l)}(\lambda_1\lambda_2 - q_1^2) - \lambda_l\Omega}{W_j^{(l)}(\lambda_1\lambda_2 - q_1^2) + \lambda_l\Omega}, \end{aligned} \tag{5.37}$$

where $l = 1, 2$, the upper sign corresponds to $j = 1$, and the lower sign, to $j = 2$. If $z$ and $z_1$ do not lie simultaneously on a face of the sample, component $g_{ll}$ becomes a function of both coefficients $W_{1,2}^{(l)}$.

The conditions of strong and weak anchoring at the boundaries $z = \pm L/2$ are determined by the sign of the derivatives

$$\left.\frac{\partial}{\partial z}g_{ll}(z, z)\right|_{z = \pm L/2}.$$

A transition from one case to another occurs for

$$W_j^{(l)} \approx \Omega\lambda_l/(\lambda_1\lambda_2 - q_1^2).$$

It should be noted that this condition also corresponds to the sign reversal of $g_{ll}^{(2)}$ in formula (5.37).

### *5.5. Planar Orientation with $\chi_a < 0$*

In this geometry, the external field vector is perpendicular to the director and may form an arbitrary angle α with the *x, y* plane. As in the case of planar geometry with $\chi_a > 0$, we direct the *y* axis along $\mathbf{n}^0$. Then $\mathbf{H} = H(\cos\alpha, 0, \sin\alpha)$, $\delta\mathbf{n} = (\delta n_1, 0, \delta n_2)$, and $\mathbf{q} = (q_1, q_2, 0)$.

The Frank energy and the surface energy in this case have the same form (5.28), (5.30) as in the previous case.

The contribution of the external field is given by

$$\Phi_f = -\frac{1}{2}\chi_a H^2 \int dz |\delta n_{1\mathbf{q}}\cos\alpha + \delta n_{2\mathbf{q}}\sin\alpha|^2. \quad (5.38)$$

Thus, matrices $\hat{a}$, $\hat{b}$, and $\hat{w}_j$ in this geometry are the same as in (5.31), while matrix $\hat{c}$ is given by

$$\hat{c} = \begin{pmatrix} A_5 & A_7 \\ A_7 & A_6 \end{pmatrix}, \quad (5.39)$$

where

$$\begin{aligned} A_5 &= K_{11}q_1^2 + K_{33}q_2^2 - \chi_a H^2\cos^2\alpha, \\ A_6 &= K_{22}q_1^2 + K_{33}q_2^2 - \chi_a H^2\sin^2\alpha, \\ A_7 &= -\chi_a H^2 \sin\alpha\cos\alpha. \end{aligned} \quad (5.40)$$

The characteristic equation in this case contains a term linear in λ:

$$\begin{aligned} &K_{11}K_{22}\lambda^4 - [A_5K_{11} + A_6K_{22} - q_1^2(K_{11} - K_2)^2]\lambda^2 \\ &\quad + 2iq_1A_7(K_{11} - K_{22})\lambda + A_5A_6 - A_7^2 = 0. \end{aligned} \quad (5.41)$$

The solutions of this equation $\lambda_1 = -\lambda_3^*$ and $\lambda_2 = -\lambda_4^*$ for $0 < \alpha < \pi/2$ and $K_{11} \neq K_{22}$ always contain an imaginary component. Nonstandardized vectors $\mathbf{e}^{(j)}$ are given by

$$\mathbf{e}^{(j)} = \begin{pmatrix} A_7 - iq_1(K_{11} - K_{22})\lambda_j \\ K_{22}\lambda_j^2 - A_5 \end{pmatrix}. \quad (5.42)$$

In the simplest cases when $\alpha = 0$ or $\alpha = \pi/2$. coefficient $A_7 = 0$ and Eq. (5.41) becomes biquadratic. Its solutions have the form

$$\begin{aligned} \lambda_j = -\lambda_{j+2} = \{&\frac{1}{2K_{11}K_{22}} \\ &\times [2K_{11}K_{22}q_1^2 + K_{33}(K_{11} + K_{22})q^2 \\ &- K_{ii}\chi_a H^2 + (-1)^{j+1}\sqrt{D}]\}^{1/2}, \end{aligned} \quad (5.43)$$

where

$$\begin{aligned} D = {} & K_{33}^2(K_{11} - K_{22})^2q_2^4 + K_{ii}^2\chi_a^2H^4 \\ & + 2(-1)^i(K_{11} - K_{22})\chi_a H^2(2K_{11}K_{22}q_1^2 + K_{ii}K_{33}q_2^2). \end{aligned}$$

Here, $i = 1$ corresponds to $\alpha = \pi/2$ and $i = 2$ to $\alpha = 0$. It should be noted that solutions (5.43) may be complex-valued.

Using formulas (5.43) and (5.42), we can derive analytic expressions for matrices $\hat{u}$ and $\hat{U}$ in (4.23). We shall not present the expression for the correlation matrix in this case since it is extremely cumbersome even in perturbation theory. However, numerical calculations of the elements of the correlation matrix on the basis of formulas (4.26), (4.27), or (4.32) are not difficult in this case also.

The peculiarity of the case under investigation is the complex valuedness of the eigenvalues $\lambda_j$ for $0 < \alpha < \pi/2$. However, the estimates obtained for realistic values of NLC parameters, the external field amplitude *H*, and the wave numbers $q_1$ and $q_2$ show that the inequality $\mathrm{Im}\lambda_j \ll \mathrm{Re}\lambda_j$ holds in all cases. For this reason, the components of matrices $\hat{g}(z, z_1)$ do not oscillate. Another feature of this case is the complex valuedness of all the component of the correlation matrix.

Concerning the effect of an external field, the situation in this case is similar to the homeotropic orientation with $\chi_a < 0$. Namely, the component of the director fluctuations which is perpendicular to **H** is not suppressed by the external field. In the given geometry, this component has the form of the linear combination

$$\delta\mathbf{n} = -\delta\mathbf{n}_1\sin\alpha + \delta\mathbf{n}_2\cos\alpha.$$

In particular, the field does not suppress the correlation matrix component $g_{22}$ for $\alpha = 0$ and $g_{11}$ for $\alpha = \pi/2$.

A problem emerging when the exact formulas (4.22) and (4.26) are used for numerical calculations is worth noting. The factors $\exp(\hat{\Lambda}l_j)$ appearing in formulas (4.21) contain both very large and very small components (in absolute value) which must be taken into account simultaneously without disregarding the latter components. For actual parameters of NLC, e.g., in solving an optical problem for a sample of thickness $L \sim 100$ μm, this necessitates that quantities of the order of $10^{\pm 2000}$ be takes into account in intermediate computations. This renders formula (4.22) inapplicable for numerical calculations. Such a problem does not emerge when thin samples with small values of $|\lambda_j|L$ are considered. In the case of large values of $\lambda L$, formulas (4.32) from perturbation theory can be used for numerical calculations. The calculations based on exact formulas (4.22) and formulas (4.32) in perturbation theory lead to virtually identical results up to values of $\lambda L \leq$ 2–3. The most significant discrepancy in the results of calculations made using these formulas are observed when $z$ and $z_1$ lie at the opposite ends of the sample ($L/2$ and $-L/2$). These discrepancies are eliminated if we take into account the second order in perturbation theory. However, the value of $g_{\alpha\beta}(L/2, -L/2)$ is negligibly small as compared to $g_{\alpha\beta}(z, z_1)$ for $z = z_1$ and hence these corrections are of no interest to us.

## 6. LIGHT SCATTERING

The measurement of the angular and polarization dependences of the intensity of scattered light is one of the methods of analysis of spatial correlation functions. From the point of view of optics, SLC-*A* and NLC are uniaxial media with the permittivity tensor

$$\varepsilon_{\alpha\beta}(\mathbf{r}) = \varepsilon_{\perp}\delta_{\alpha\beta} + \varepsilon_a n_{\alpha}(\mathbf{r}) n_{\beta}(\mathbf{r}), \qquad (6.1)$$

where $\varepsilon_a = \varepsilon_{\|} - \varepsilon_{\perp}$, and $\varepsilon_{\|}$ and $\varepsilon_{\perp}$ are the dielectric constants along and across the optical axis. The permittivity fluctuations $\delta\varepsilon_{\alpha\beta}$, which have the form

$$\delta\varepsilon_{\alpha\beta} = \varepsilon_a(n_{\alpha}^0 \delta n_{\beta} + n_{\beta}^0 \delta n_{\alpha}), \qquad (6.2)$$

in the approximation linear in $\delta\mathbf{n}$, lead to light scattering.

The intensity of scattering in an anisotropic planar layer can be written in the form (cf. [29])

$$I(\mathbf{e}^{(i)}, \mathbf{e}^{(j)}) = \frac{V I_0^{(i)} \omega^4}{16\pi^2 R^2 c^4} A^{(is)}(\mathbf{k}^{(s)}, \mathbf{k}^{(i)}) \times e_{\nu}^{(s)} e_{\mu}^{*(s)} D_{\nu\rho\mu\eta}(\mathbf{q}_{sc}; L) e_{\rho}^{(i)} e_{\eta}^{*(i)}, \qquad (6.3)$$

where $V$ *is* the scattering volume; $\omega$ is the cyclic frequency; $c$ is the velocity of light in vacuum;

$$D_{\nu\rho\mu\eta}(\mathbf{q}_{sc}; L) = \frac{1}{L} \times \int_{-L/2}^{L/2}\int_{-L/2}^{L/2} D_{\nu\rho\mu\eta}(\mathbf{q}_{sc\perp}; z, z_1) e^{-iq_{sc\|}(z - z_1)} dz dz_1, \qquad (6.4)$$

and $D_{\nu\rho\mu\eta}(\mathbf{q}_{sc\perp}; z, z_1)$ is the correlation function of permittivity tensor fluctuations, which has the following form in the coordinate representation:

$$D_{\nu\rho\mu\eta}(\mathbf{r}_1, \mathbf{r}_2) = \langle \delta\varepsilon_{\nu\rho}(\mathbf{r}_1) \delta\varepsilon_{\mu\eta}(\mathbf{r}_2) \rangle, \qquad (6.5)$$

$\mathbf{q}_{sc} = \mathbf{k}^{(i)} - \mathbf{k}^{(s)}$ being the scattering vector. Here, $i$ and $s$ denote the incident and scattered waves, respectively, $\mathbf{e}^{(i)}$ and $\mathbf{e}^{(s)}$ are the polarization vectors of these waves, and $\mathbf{k}^{(i)}$ and $\mathbf{k}^{(s)}$ are the wave vectors. In a uniaxial medium, $\mathbf{e}^{(i)}$, $\mathbf{k}^{(i)}$ and $\mathbf{e}^{(s)}$, $\mathbf{k}^{(s)}$ may assume two values each, corresponding to the ordinary and extraordinary waves. The quantity $A^{(is)}(\mathbf{k}^{(s)}, \mathbf{k}^{(i)})$, $\mathbf{k}^{(i)}$ is the angular factor, $R$ is the distance from the scattering volume to the point of observation, and $I_0^{(i)}$ is the intensity of incident light. The refraction at the sample boundary is not taken into account in formula (6.3). This problem is considered in [30].

In accordance with formula (6.2), the correlation function $\hat{D}(\mathbf{q}_{\perp}; z, z_1)$ in formula (6.3) is connected wit the correlation function of director fluctuations $\hat{G}$ through the following relation:

$$D_{\nu\rho\mu\eta} = \varepsilon_a^2(n_{\nu}^0 n_{\mu}^0 G_{\rho\eta} + n_{\nu}^0 n_{\eta}^0 G_{\rho\mu} + n_{\rho}^0 n_{\mu}^0 G_{\nu\eta} + n_{\rho}^0 n_{\eta}^0 G_{\nu\mu}), \qquad (6.6)$$

where $\hat{G} = \hat{g}$ for a nematic and

$$G_{\alpha\beta}(\mathbf{q}_{\perp}; (z, z_1)) = q_{\alpha} q_{\beta} g(\mathbf{q}_{\perp}; z, z_1)$$

for a smectic.

The integral with respect to $z$ and $z_1$ in formula (6.4) can be evaluated in the general form using formula (4.22). This gives

$$\int_{-L/2}^{L/2}\int_{-L/2}^{L/2} \hat{G}(\mathbf{q}_{\perp}; z, z_1) e^{-iq_{\|}(z - z_1)} dz dz_1 = k_B T \hat{U} \hat{J}^{-1}\left(4\sin\frac{\hat{J}L}{2}\hat{F}\sinh\frac{\hat{J}L}{2} + L\hat{J} + \hat{I} - e^{\hat{J}L}\right)\hat{J}^{-1}\hat{V}, \qquad (6.7)$$

where $\hat{J} = \hat{\Lambda} - iq_{\|}\hat{I}$ is a diagonal matrix.

For the extraordinary ray, the length of the wave vector $\mathbf{k}$ depends on direction, and the polarization vector $\mathbf{e}$ lying in the $(\mathbf{k}, \mathbf{n}^0)$ plane is not orthogonal to $\mathbf{k}$. We are interested above all in the effect of the anisotropy of the medium on the length of the scattering vector $\mathbf{q}_{sc} = \mathbf{k}^{(i)} - \mathbf{k}^{(s)}$ since the scattering intensity in an unbounded medium in zero field ($H = 0$) is proportional to $\sim 1/q_{sc}^0$ [1]. We will disregard anisotropy in the remaining quantities, which allows us to assume that $\mathbf{e} \perp \mathbf{k}$ and $A^{(is)}(\mathbf{k}^{(s)}, \mathbf{k}^{(i)}) = 1$.

Let us analyze the possibility of determining the anchoring energy from the data on light scattering in homeotropic and planar geometries of NLC with $\chi_a > 0$.

### *6.1. Homeotropic Orientation*

Let us consider the case when the incident ray is normal to the surface of the cell. The geometry of scattering is shown in Fig. 4. Here, the incident ray is ordinary and the scattered ray is extraordinary.

The angular dependence of the intensity of scattered light in this case has the form

$$I(\theta, \varphi) = C_0 \sin^2\theta \times \left[G_{(2)} + (G_{(1)} - G_{(2)})\cos^2\varphi \cos^2\frac{\theta}{2}\right], \tag{6.8}$$

where

$$C_0 = I_0^{(i)} V \left(\frac{\omega^2 \varepsilon_a}{4\pi R c^2}\right)^2$$

is a constant, and $G_{(j)}$ has the form

$$\begin{aligned} G_{(j)} &= \frac{k_B T}{K_{33}(\lambda_j^2 + q_\parallel^2)L} \\ &\times \left\{ L + \frac{q_\parallel^2 - \lambda_j^2}{\lambda_j(\lambda_j^2 + q_\parallel^2)} + \frac{K_{33}^2\lambda_j^2 - W_1 W_2}{\lambda_j \mathcal{K}_j^+} + 4e^{-\lambda_j L}\lambda_j \right. \\ &\times \frac{(W_1 W_2 - K_{33}^2 q_\parallel^2)\cos(Lq_\parallel) - K_{33} q (W_1 + W_2)\sin(Lq_\parallel)}{(\lambda_j^2 + q_\parallel^2)\mathcal{K}_j^+} \\ &\left. + \frac{e^{-2\lambda_j L}}{\lambda_j \mathcal{K}_j^+}\left[K_{33}^2\lambda_j^2 - W_1 W_2 - \left(\lambda_j L + \frac{\lambda_j^2 - q_\parallel^2}{\lambda_j^2 + q_\parallel^2}\right)\mathcal{K}_j^-\right]\right\} \\ &\times \left[1 - e^{-2\lambda_j L}\frac{\mathcal{K}_j^-}{\mathcal{K}_j^+}\right]^{-1}, \end{aligned} \tag{6.9}$$

where

$$\mathcal{K}_j^\pm = (K_{33}\lambda_j \pm W_1)(K_{33}\lambda_j \pm W_2),$$

$$q_\parallel = 2\sqrt{\varepsilon_\perp}\frac{\omega}{c}\sin^2(\theta/2),$$

and $\lambda_j$ are calculated using formula (5.13) with

$$q = q_\perp = \sqrt{\varepsilon_\perp}\frac{\omega}{c}\sin\theta.$$

For large values of $\lambda_j L$, the exponential terms in formula (6.9) can be omitted, and the quantity $G_{(j)}$ is determined by the first three terms. The first of these terms corresponds to the limit of an unbounded medium, $L \longrightarrow \infty$,

$$G_{(j)} = \frac{k_B T}{K_{33}}\frac{1}{\lambda_j^2 + q_\parallel^2} = \frac{k_B T}{K_{33}q_\parallel^2 + K_{jj}q_\perp^2 + \chi_a H^2},$$

which coincides with the well-known result [1].

The second term is associated with the finiteness of the volume in the exponential approximation, and we can put $\cosh x \approx \sinh x \approx \exp x$ in formula (5.19) while integrating of the spatially inhomogeneous correlation function of an unbounded medium, for which

$$g(\mathbf{q}, z, z_1) \approx g(\mathbf{q}, |z - z_1|).$$

For this reason, this term is independent of the anchoring energies $W_1$ and $W_2$.

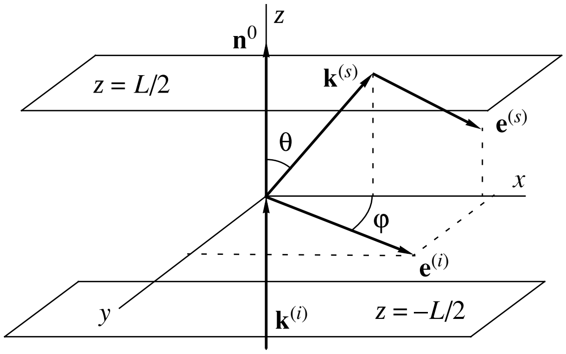

**Fig. 4.** Geometry of light scattering for the homeotropic orientation of NLC.

The third term in formula (6.9), which is a function of $W_1$ and $W_2$, appears due to integration of the third term in formula (5.19) between finite limits in the exponential approximation taking into account the spatial inhomogeneity of the correlation function. It should be noted that the spatially inhomogeneous term associated with the fourth term in formula (5.19) does not contribute to scattering in view of symmetry.

It should be noted that the third term in formula (6.9) is commensurate with the first two terms for samples that are not very thick.

It can be seen from formula (6.8) that angle φ determines only the relative contributions $\delta n_1$ and $\delta n_2$ to scattering. Consequently, the main information on parameters $W_1$ and $W_2$ is contained in the dependence of $I(\theta, \varphi)$ on angle θ, which is shown in Fig. 5. The figure presents the dependence of the scattering intensity calculated using formulas (6.8) and (6.9) for two sets of anchoring energy in the angular interval from 0.003 to 0.03 rad. The inset to Fig. 5a shows this dependence in an extended angular interval $0 \le \theta \le \pi/3$. Figures 5a and 5b depict three curves each, corresponding to different values of the external field: $H = 0$, 700, and 2000 G. It should be noted that when the first three terms in formula (6.9) are used, the deviation from exact results appears starting from angles $\theta \le 10^{-2}$ rad.

The difference in the behavior of the curves depicted in Figs. 5a and 5b shows that the measurement of the angular dependence of scattering intensity for different values of the applied field may serve as an effective method for determining anchoring energies.

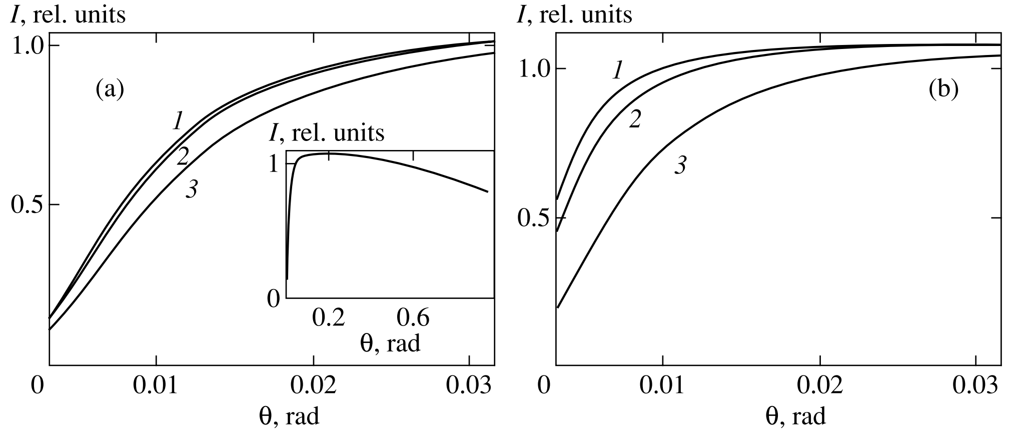

**Fig. 5.** Angular dependence of the intensity of scattered light for the homeotropic orientation of NLC in an external field for $L = 3 \times 10^{-3}$ cm, $\chi_a = 10^{-7}$, $k_0 = \sqrt{\varepsilon_\perp}(\omega/c) = 10^5$ cm$^{-1}$. The values of $K_{jj}$, $j$ = 1, 2, 3, are the same as in Fig. 2. (a) $W_1 = 10^{-3}$ erg/cm$^2$, $W_2 = 2 \times 10^{-3}$ erg/cm$^2$; (b) $W_1 = 10^{-4}$ erg/cm$^2$, $W_2 = 2 \times 10^{-4}$ erg/cm$^2$, $H$ = 0 (*1*), 700 G (*2*) and 2000 G (*3*).

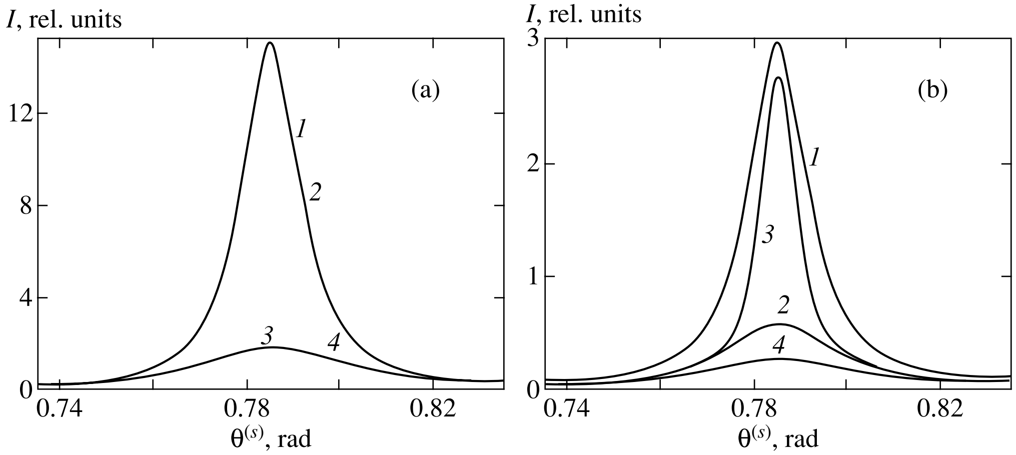

**Fig. 6.** Angular dependence of the scattering intensity of the (*ee*)-type for $\theta^{(i)} = \pi/4$ for the planar orientation of NLC. (a) $\varphi = \pi/2$, (b) $\varphi = \pi/20$. In our calculations, we assumed that $\varepsilon_\| = 3.5$, $\varepsilon_\perp = 2.5$, $L = 10^{-3}$ cm, and $H = 0$. The values of $K_{jj}$ ($j$ = 1–3) are the same as in Fig. 2. (*1*) $W_1^{(1)} = 2$, $W_1^{(2)} = 7$, $W_2^{(1)} = 1.9$, $W_2^{(2)} = 8.6$; (*2*) $W_1^{(1)} = 200$, $W_1^{(2)} = 7$, $W_2^{(1)} = 190$, $W_2^{(2)} = 8.6$; (*3*) $W_1^{(1)} = 2$, $W_1^{(2)} = 700$, $W_2^{(1)} = 1.9$, $W_2^{(2)} = 860$; (*4*) $W_1^{(1)} = 200$, $W_1^{(2)} = 700$, $W_2^{(1)} = 190$, $W_2^{(2)} = 860$. These values are in the units of $10^{-4}$ erg/cm$^2$. The intensities are normalized to the same value.

### *6.2. Planar Orientation*

In an analysis of light scattering in NLC, the most interesting case is when $q_{sc} \longrightarrow 0$ since the scattering intensity has the maximum value [1]. In view of the anisotropy of the medium, such a situation may be realized when the incident and scattered rays are of the same type; i.e., both rays are ordinary (*oo*) or both rays are extraordinary (*ee*). Since the scattering of the (*oo*) type is absent for geometrical reasons [1], we will consider the (*ee*)-type scattering. In contrast to the homeotropic situation, we have to consider here the case of oblique incidence since the intensity of light scattering through zero angle for normal or grazing incidence vanishes [29]. For the sake of simplicity, we choose the angle of incidence $\theta^{(i)} = \pi/4$, since in this case the extinction coefficient is close to its maximum [29], and confine the analysis to the case when the incident and scattered rays lie in the same plane with the normal to the surface. The azimuthal angle φ between the projections of vectors $\mathbf{k}^{(i)}$ and $\mathbf{k}^{(s)}$ on the *xy* plane and the *x* axis will be assumed to be arbitrary.

The relative contribution of the modes $\delta n_{1,2}$ of director fluctuations $\delta\mathbf{n} = (\delta n_1, 0, \delta n_2)$ to scattering is determined by the factor

$$e_\alpha^{(s)} e_\gamma^{(s)} D_{\alpha\beta\gamma\rho} e_\beta^{(i)} e_\rho^{(i)} = (\mathbf{e}^{(s)} \cdot \mathbf{n}^{(0)})^2 (\mathbf{e}^{(i)} \hat{g} \mathbf{e}^{(i)})$$
$$+ (\mathbf{e}^{(s)} \cdot \mathbf{n}^{(0)})(\mathbf{e}^{(i)} \cdot \mathbf{n}^{(0)})[(\mathbf{e}^{(i)} \cdot \hat{g} \mathbf{e}^{(s)}) + (\mathbf{e}^{(s)} \cdot \hat{g} \mathbf{e}^{(i)})]$$
$$+ (\mathbf{e}^{(i)} \cdot \mathbf{n}^{(0)})^2 (\mathbf{e}^{(s)} \cdot \hat{g} \mathbf{e}^{(s)}). \tag{6.10}$$

For $\varphi = \pi/2$, vectors $\mathbf{k}^{(i)} \approx \mathbf{k}^{(s)} \perp \mathbf{e}^{(i)} \approx \mathbf{e}^{(s)}$ for small-angle scattering lie in the $yz$ plane. Consequently, the convolutions of the type $(\mathbf{e}^{(i,s)} \cdot \hat{g}\mathbf{e}^{(i,s)})$ in formula (6.10) do not contain a contribution from mode $\delta\mathbf{n}_1$ (the director fluctuation component along the $x$ axis). Thus, the contribution of the $\delta\mathbf{n}_2$ mode and the surface energies $W_1^{(2)}$ and $W_2^{(2)}$ associated with it are measured directly in this geometry. This is illustrated in Fig. 6a depicting the dependence of scattering intensity on the angle $\theta^{(s)}$ between $\mathbf{k}^{(s)}$ and the $z$ axis. Formally, the figure shows the intensity curves corresponding to four sets of the values of surface energy. Since the $g_{22}$ component is virtually independent of $W_{1,2}^{(1)}$ (see Fig. 3 in Subsection 5.4), the pairs of curves in Fig. 6a corresponding to the same set $W_{1,2}^{(2)}$ but to different values of $W_{1,2}^{(1)}$ are indistinguishable. Thus, this geometry is most convenient for determining a pair of energy values of $W_{1,2}^{(2)}$.

In order to determine the other two values of anchoring energy $W_{1,2}^{(1)}$, we must measure the contribution of the $\delta n_1$ mode to scattering, i.e., make measurements for $\varphi \neq \pi/2$. It should be kept in mind in this case that the vectors $\mathbf{e}^{(s,i)}$ coincide in direction with the director vector $\mathbf{n}^0$ for $\varphi = 0$, and hence all convolutions of the type $(\mathbf{e}^{(i,s)} \cdot \hat{g}\mathbf{e}^{(i,s)})$ in formula (6.10) are equal to zero by virtue of the condition $\delta\mathbf{n} \perp \mathbf{n}^0$. For this reason, such measurements should be made for $0 < \varphi < \pi/2$. Figure 6b shows the angular dependence of scattering intensity for the same four sets of anchoring energies $W_{1,2}^{(1,2)}$ as in Fig. 6a. It can be seen that the scattering intensity starts depending on the anchoring energies $W_{1,2}^{(1)}$ also due to the contribution of the $\delta\mathbf{n}_1$ mode.

The effect of the external field in the planar geometry is illustrated in Fig. 7. It can be seen that the characteristic fields in which the small-angle scattering intensity decreases significantly are of the order of $H \sim 10^4$ G.

## 7. CONCLUSIONS

A general algorithm is proposed for calculating the correlation functions of orientation fluctuations in bounded samples of liquid crystals taking into account the effect of orienting surfaces and external fields. Spatial correlation functions are determined for plane-parallel cells in nematic and smectic liquid crystals without using simplifying assumptions of the one-constant approximation type. The cases of positive and negative anisotropy of permittivity or magnetic susceptibility are considered.

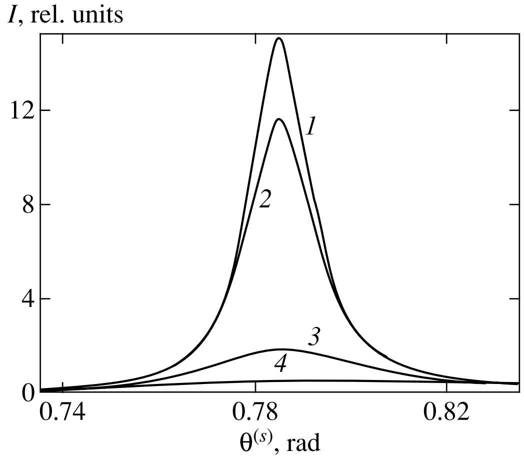

**Fig. 7.** Effect of external field on the angular dependence of the intensity of scattered light in the same geometry as in Fig. 6a. The parameters used in calculations are the same as in Fig. 6. The anchoring energies are assumed to be $W_1^{(1)} = 2$, $W_1^{(2)} = 7$, $W_2^{(1)} = 1.9$, $W_2^{(2)} = 8.6$ (in units of $10^{-4}$ erg/cm$^2$), $\chi_a = 10^{-7}$, $H = 0$ (*1*), $2 \times 10^3$ G (*2*), $10^4$ G (*3*), and $2 \times 10^4$ G (*4*).

The obtained results were used for calculating the angular dependence of the intensity of scattered light in various geometries. It turns out that the scattering intensity in bounded cells is very sensitive to the values of anchoring energy as well as to the magnitude of the applied field. It is shown, in particular, that in the case of the planar orientation of NLC, such experiments can be used for measuring separately the surface energies associated with the rotation of the director in the orienting plane and deviations from this plane.

The study of the intensity of light scattering in nematic liquid crystals with negative anisotropy may be interesting in connection with additional possibilities of studying the materials in which fluctuational modes vary in different ways in an external field.

The approach developed by us here can also be generalized to other geometries such as spherical or cylindrical, which are important for describing orientation fluctuations in liquid crystal drops encapsulated in a polymer matrix or porous media, as well as to more realistic surface potentials [27] differing from the Rapini potential (2.11).

## ACKNOWLEDGMENTS

The authors are grateful to V. A. Belyakov for fruitful discussions and helpful advice and to A. N. Shalaginov for a valuable exchange of views.

This work was supported by the ISSEP educational program.

## REFERENCES

1. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Clarendon Press, Oxford, 1993).
2. T. Bellini, N. A. Clark, C. D. Muzny, *et al.*, Phys. Rev. Lett. **69**, 788 (1992); N. A. Clark, T. Bellini, R. M. Malzbender, *et al.*, Phys. Rev. Lett. **71**, 3505 (1993); F. M. Aliev, Mol. Cryst. Liq. Cryst. **243**, 91 (1994).
3. A. Poniewierski and A. Samborski, Phys. Rev. E **53**, 2436 (1996).
4. P. Ziherl, A. Sarlah, and S. Zumer, Phys. Rev. E **59**, 602 (1998).
5. B. Alkhairalla, H. Allinson, N. Boden, *et al.*, Phys. Rev. E **59**, 3033 (1999).
6. S. Sprunt, M. S. Spector, and J. D. Litster, Phys. Rev. A **45**, 7355 (1992); R. Hall, K. Miyachi, D. Newton, *et al.*, Jpn. J. Appl. Phys. **31**, 329 (1992); T. Bellini, N. A. Clark, and D. W. Schaefer, Phys. Rev. Lett. **74**, 2740 (1995); S. Stallinga, M. M. Wittebrood, D. H. Luijendijk, and Th. Rasing, Phys. Rev. E **53**, 6085 (1996); A. N. Shalaginov, Phys. Rev. E **53**, 3623 (1996).
7. H. Li and M. Kardar, Phys. Rev. Lett. **67**, 3275 (1991).
8. H. Li, M. Paczuski, K. Kardar, and K. Huang, Phys. Rev. B **44**, 8274 (1991).
9. A. Adjari, L. Peliti, and J. Prost, Phys. Rev. Lett. **66**, 1481 (1991).
10. A. Ajdari, B. Duplantier, D. Hone, *et al.*, J. Phys. II **2**, 487 (1992).
11. D. Kang and C. Rosenblatt, Phys. Rev. E **53**, 2976 (1996).
12. S. Matsumoto, M. Houlbert, T. Hayashi, and K. Kubodera, Appl. Phys. Lett. **69**, 1044 (1996); L. Vicari, Phys. Rev. E **58**, 3280 (1998).
13. O. D. Lavrentivich, Phys. Rev. A **46**, R722 (1992); Yu. Panarin, S. T. Mac Lughadha, and J. K. Vij, Phys. Rev. E **52**, R17 (1995); B. Cull, Y. Chi, and S. Kumar, Phys. Rev. E **51**, 526 (1995); R. E. Kraig, P. L. Taylor, R. Ma, and D. Yang, Phys. Rev. E **58**, 4594 (1998); U. Kuhnau, A. G. Petrov, G. Klose, and H. Schmiedel, Phys. Rev. E **59**, 578 (1999).
14. R. J. Ondris-Crawford, G. P. Crawford, S. Zumer, and J. W. Doane, Phys. Rev. Lett. **70**, 194 (1993); L. Limat and J. Prost, Liq. Cryst. **13**, 101 (1993); M. Ambrozic and S. Zumer, Phys. Rev. E **59**, 4159 (1999).
15. P. G. de Gennes, C. R. Seances Acad. Sci. **266**, 15 (1968); J. Phys. (Paris), Colloq. **301**, C4–65 (1969).
16. V. L. Pokrovskiĭ and E. I. Kats, Zh. Éksp. Teor. Fiz. **73**, 774 (1977) [Sov. Phys. JETP **46**, 405 (1977)]; A. Yu. Val’kov and V. P. Romanov, Zh. Éksp. Teor. Fiz. **83**, 1777 (1982) [Sov. Phys. JETP **56**, 1028 (1982)].
17. B. Ya. Zel’dovich and N. V. Tabiryan, Zh. Éksp. Teor. Fiz. **81**, 1788 (1981) [Sov. Phys. JETP **54**, 922 (1981)].
18. T. Ya. Marusiĭ, Yu. A. Reznikov, V. Yu. Reshetnyak, *et al.*, Zh. Éksp. Teor. Fiz. **91**, 851 (1986) [Sov. Phys. JETP **64**, 502 (1986)].
19. K. Eidner, M. Lewis, H. K. M. Vithana, and D. L. Johnson, Phys. Rev. A **40**, 6388 (1989).
20. V. P. Romanov and A. N. Shalaginov, Zh. Éksp. Teor. Fiz. **102**, 884 (1992) [Sov. Phys. JETP **75**, 483 (1992)].
21. A. Poniewierski and A. Holyst, Phys. Rev. B **47**, 9840 (1993).
22. A. Rapini and M. Popoular, J. Phys. (Paris), Colloq. **30**, C4–54 (1969).
23. A. N. Shalaginov and V. P. Romanov, Phys. Rev. E **48**, 1073 (1993).
24. A. Yu. Val’kov and M. V. Romanov, Colloids Surf., A **148**, 179 (1999).
25. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1976; Pergamon, Oxford, 1980), Part 1.
26. R. P. Feynman, *Statistical Mechanics: A Set of Lectures* (Benjamin, Reading, 1972; Mir, Moscow, 1978, 2nd ed.).
27. L. M. Blinov, E. I. Kats, and A. A. Sonin, Usp. Fiz. Nauk **152**, 449 (1987) [Sov. Phys. Usp. **30**, 604 (1987)].
28. V. N. Matveenko and E. A. Kirsanov, *Surface Phenomena in Liquid Crystals* (Mosk. Gos. Univ., Moscow, 1991).
29. A. Yu. Val’kov, V. P. Romanov, and A. N. Shalaginov, Usp. Fiz. Nauk **164**, 149 (1994) [Phys. Usp. **37**, 139 (1994)].
30. M. Lax and D. F. Nelson, in *Proceedings of the 3rd Rochester Conference on Coherent and Quantum Optics*, Ed. by L. Mandel and E. Wolf (Plenum, New York, 1973), p. 415.

*Translated by N. Wadhwa*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 363–371.*
*Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 409–419.*
*Original Russian Text Copyright © 2001 by Strashnikova, Voznyĭ, Reznichenko, Gaĭvoronskiĭ.*

**SOLIDS**
**Structure**

# Optical Properties of Porous Silicon

**M. I. Strashnikova, V. L. Voznyĭ, V. Ya. Reznichenko, and V. Ya. Gaĭvoronskiĭ***
*Institute of Physics, National Academy of Sciences of Ukraine, pr. Nauki 46, Kiev, 03028 Ukraine*
**e-mail: vlad@iop.kiev.ua.*

Received November 14, 2000

**Abstract**—We have measured the absorption spectra and the dispersion of refractive index for porous silicon samples with different porosities in the energy range 1.5–3.5 eV at room temperatures. The experimental data are compared with the dependences calculated by using Bruggeman’s theory for the dielectric constant of a multicomponent system composed of crystal silicon, $SiO_2$, amorphous silicon, and voids (pores). The best agreement between the experimental and theoretical dependences is achieved for a significant percentage of $SiO_2$ in the porous silicon samples. © *2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

The flow of papers on porous silicon (por-Si) published in the past decade is growing like an avalanche. Attention has been focused on elucidating the nature of its visible photoluminescence at room temperature, which could be of great importance for applications (see, e.g., the review article [1]). The first and still one of the main versions of the nature of photoluminescence is the quantum-dimensional theory. Nanometer-sized wires and dots, in which quantum energy levels emerge above the band gap due to the quantum confinement of carriers, are believed to be formed in porous structures through electrochemical etching. This results in a blueshift of the absorption edge and, consequently, in the accompanying frequency shift of the luminescence that emerges during band-to-band carrier recombination. There are also important arguments for alternative points of view on the nature of photoluminescence, with the molecular theory being most popular among them.

As regards the study of fundamental optical characteristics for porous silicon (refractive indices $n$, absorption coefficients $\alpha$, and, consequently, the properties of its dielectric function $\varepsilon$), such studies are relatively few [2–10]. In most of the above studies, either the absorption spectra or the refractive indices were measured, and only in [3, 4] were both characteristics investigated simultaneously. We emphasize that a comprehensive study is important for elucidating the effects of the changed structure of the material when por-Si is produced on its optical characteristics. After all, the dielectric function is directly related to the band structure and selection rules for band-to-band transitions.

A comparison of the results obtained by different authors clearly reveals that the absolute values of refractive index vary over a wide range: from 3.99 [10] to 1.28 [9]. In addition, the dispersion curves $n(\lambda)$ in a wide spectral range differ in slope and structure in the ultraviolet near strong transitions in crystal silicon (3.4–4.2 eV) [3, 4]. These differences may be associated with different porosities of the samples analyzed.

Therefore, it seems of current interest to comprehensively study the dielectric function, to compare it with available theoretical models, and to correlate it with the porosity of por-Si samples.

## 2. EXPERIMENT

**Samples.** We analyzed two por-Si samples produced by a standard technique and separated from a silicon single-crystal substrate. Sample no. 1 was taken from a series of crystals whose luminescence was investigated in [11, 12]. It was in optical contact with a glass substrate and appeared a transparent bluish film. The crystal surface viewed through a microscope was smooth, shiny, and structureless. The sample thickness, as measured with a Linnik microinterferometer (MII-4) by the shift of the zero interference fringe during the reflection from the crystal and substrate surfaces, was 0.6 μm. As was shown in [12] using electron-microscope measurements, the samples of this series consisted of crystals 5–20 nm in size with a perfect single-crystal structure. They had an amorphous-phase impurity that gave clear halos in the electron diffraction pattern.

Sample no. 2 belonged to the series of crystals analyzed in [7]. Its thickness measured with a microscope “edge on” was 30 μm. It was in a free stand, with one of its corners being fixed to the substrate. The sample appeared as a yellowish transparent plate whose surface had a grainy structure through a microscope with a 1000-fold magnification. The views of these grains when focusing on the upper and lower crystal surfaces are shown in Fig. 1. The grain sizes (of the order of several tenths of a micron) are much larger than those for sample no. 1, as evidenced by the rugged edges of the interference fringes obtained with the MII-4 during the reflection from the sample surface.

1063-7761/01/9302-0363$21.00 © 2001 MAIK “Nauka/Interperiodica”

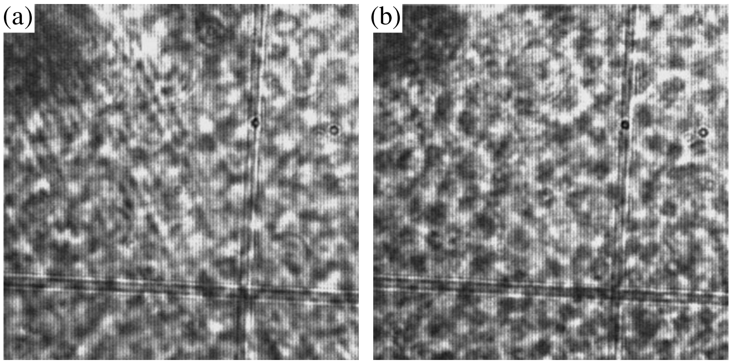

**Fig. 1.** Microphotographs of the (a) upper and (b) lower surfaces of sample no. 2.

Apart from analyzing the por-Si samples, we measured the optical constants of the silicon deposited on a glass substrate. The film was produced by a technique similar to that described in [13] through the deposition of amorphous germanium. The deposition conditions ensured the formation of an amorphous structure. The film surface viewed through a microscope was smooth and mirror-reflecting, while the interference fringes on the MII-4 were clear-cut. The film thickness was 0.06 μm.

The dispersion of refractive index was measured by using a Jamin two-beam interferometer combined with a spectrograph (DFS-13), whose reciprocal linear dispersion was 4 Å $\mathrm{mm}^{-1}$. The sample to be analyzed was placed in one of the two interferometer arms. In this case, a path difference $[n(\lambda) - 1]d$ arises between the interfering beams, where $n(\lambda)$ is the refractive index of the sample and $d$ is its thickness. This causes the entire fringe pattern and, in particular, the zero fringe to be shifted by $\Delta y$. The relation between the pattern displacement and the path difference is given by

$$[n(\lambda)-1]d = \left(\frac{\Delta y}{H}\right)\lambda, \qquad (1)$$

where $H$ is the fringe width. By sequentially photographing the patterns with and without the crystal and then measuring the zero-fringe displacement, we determined the refractive index.[1] Thus, the technique for measuring $n(\lambda)$ is based on a direct measurement of the change in phase of the light beam passed through the crystal plate, while the authors of [3, 4] used ellipsometry and calculations from the Kramers–Kronig relations based on measured reflectance spectra.

The absorption spectra were measured with an automated spectral setup that included a monochromator (MDR-6) with replaceable gratings. The reciprocal linear dispersion was 6.5 and 13 Å $\mathrm{mm}^{-1}$ in the wavelength ranges 200–500 and 500–900 nm, respectively. The light source was a halogen lamp. The cross-sectional area of the light spot focused on the sample surface was 1 $\mathrm{mm}^2$. The computer control of the experiment allowed the light-flux instability to be “kept” at 3%.

The absorption coefficients α for the free stand samples were calculated by taking into account multibeam interference using the relation

$$T = n\exp(-\alpha d)\left|\frac{t_1^2(n+i\kappa)}{1+r_1^2\exp(-i\Phi-\alpha d)}\right|^2, \qquad (2)$$

where $T$ and $d$ are the transmittance in intensity and thickness of the sample, respectively;

$$r_1 = \frac{(n-1)+i\kappa}{(n+1)+i\kappa} \text{ and } t_1 = \frac{2}{(n+1)+i\kappa}$$

are the complex amplitude reflectivity and transmittance at the air-sample interface for normally incident light; $\Phi = 4\pi nd/\lambda$ is the phase shift of the beam when it traverses the sample twice; and the extinction coefficients κ (the imaginary parts of complex refractive index $\tilde{n} = n + i\kappa$) are related to α by $\alpha = 4\pi\kappa/\lambda$.

When the film to be analyzed was in optical contact with the substrate, we performed our calculation by the method from [14]. The passage of a light beam through the film and the reflection of light from it were calculated by taking into account multibeam interference, while the contributions from the multiple beam passages through the thick transparent substrate were added incoherently. The transmittance $T$ of the sample on the substrate relative to the measured substrate transmittance for normal incidence is given by

$$\begin{gathered}T = T_f/[(1-R_gR_f)(1-R_g)],\\ T_f = n_g\exp(-\alpha d)\left|\frac{t_1t_2}{1-r_1r_2\exp(-i\Phi-\alpha d)}\right|^2,\\ R_f = \left|\frac{t_1r_1t_2\exp(-i\Phi-\alpha d)}{1-r_1r_2\exp(-i\Phi-\alpha d)}-r_2\right|^2.\end{gathered} \qquad (3)$$

Here, $T_f$ and $R_f$ are, respectively, the transmittance and reflectivity at the substrate-film interface in intensity with allowance for multibeam interference; $n_g$ is the

---

[1] A special method was developed for measuring crystals of large thickness, where a considerable zero-fringe displacement moved the fringe out of the visibility range.

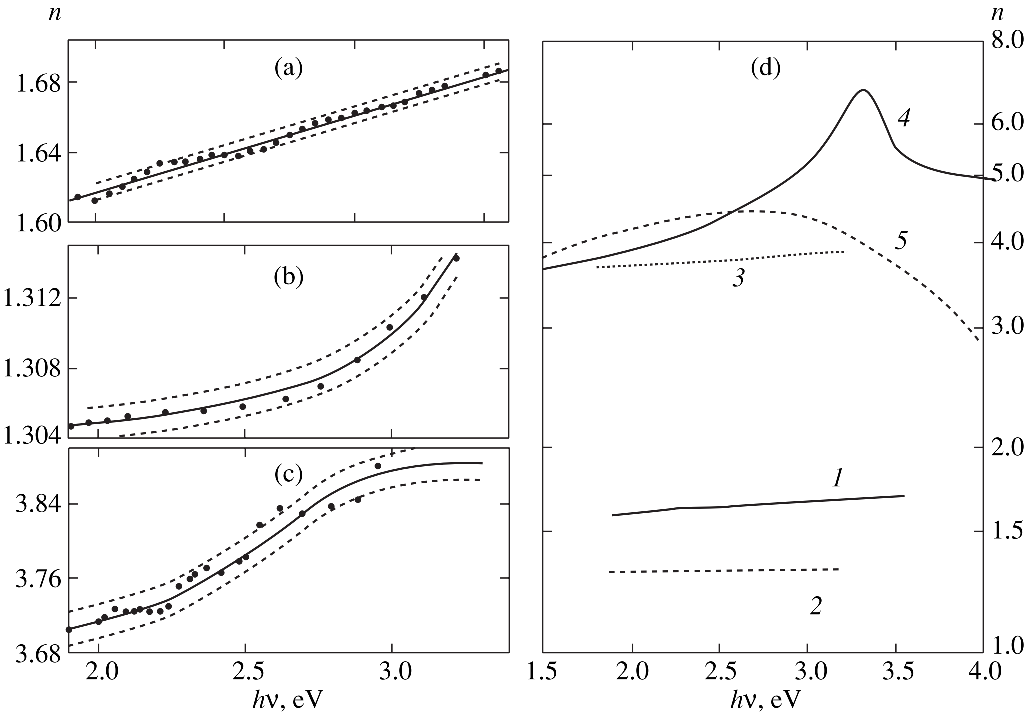

**Fig. 2.** The dispersion of refractive index for por-Si samples nos. 1 (a) and 2 (b) and for a-Si (c). (d) The combined data for por-Si no. 1 (curve *1*), por-Si no. 2 (curve *2*), a-Si (curve *3*), crystal Si from [15] (curve *4*), and amorphous Si from [16] (curve *5*).

refractive index of the substrate; $R_g = (n_g - 1)^2/(n_g + 1)^2$ is the reflectivity at the substrate-air interface; and

$$r_2 = \frac{(n - n_g) + i\kappa}{(n + n_g) + i\kappa} \quad \text{and} \quad t_2 = t_1(n + i\kappa)$$

are the complex reflectivity and transmittance at the film-substrate interface, respectively.

Eqs. (2) and (3) allow the absorption of light in the sample to be properly determined, because the contribution of reflections from the faces and multibeam interference can significantly distort the absorption coefficient estimated from simpler relations. In a porous medium, apart from the absorption and reflection of light from the faces, there is also light scattering. In our experiments, we did not measure the fraction of scattered light and, thereby, overestimated the absorption coefficient. Our samples were of a fairly high optical quality, and the contribution of scattered emission was not dominant.

Based on the experimentally derived spectral dependences of transmittances and refractive indices using the measured sample thickness, we calculated the absorption [$\alpha(h\nu)$] and extinction [$\kappa(h\nu)$] coefficients. Subsequently, we calculated the real, $\varepsilon'(h\nu) = n(h\nu)^2 - \kappa(h\nu)^2$, and imaginary, $\varepsilon''(h\nu) = 2n(h\nu)\kappa(h\nu)$, parts of dielectric constant $\varepsilon = \varepsilon' + i\varepsilon' = \tilde{n}^2$.

The photoluminescence spectrum was taken with the same setup. The third harmonic of a pulsed YAG : $Nd^{3+}$ (wavelength 353 nm) laser was used for the excitation. The excitation power was about 1 mW cm$^{-2}$, and the pulse duration was 10 ns.

## 3. RESULTS

The measured energy dependences of refractive index, $n(h\nu)$, for the two samples, the porous (por-Si) and amorphous (a-Si) silicon deposited on a glass substrate, are shown in Figs. 2a–2c. In each figure, the mean measurement error corresponds to the distance between the upper and lower dashed lines that bound the “predictable” domain. For sample no. 1, the experimentally measured dispersion of refractive index is fitted by

$$n = 1.491 + 0.055h\nu,$$

where $h\nu$ is in electronvolts. For sample no. 2, the fit is

$$n^2 = 1.301^2 + \frac{0.096}{3.612^2 - (h\nu)^2}.$$

Figure 2d shows all the measured experimental curves and the published data on the dispersion of refractive index for crystal silicon (c-Si) obtained by spectral ellipsometry [15] and for amorphous silicon (a-Si) calculated from Kramers–Kronig relations based on the measured reflectance spectrum $R(h\nu)$ [16]. For convenience, $n(h\nu)$ is presented on a logarithmic scale.

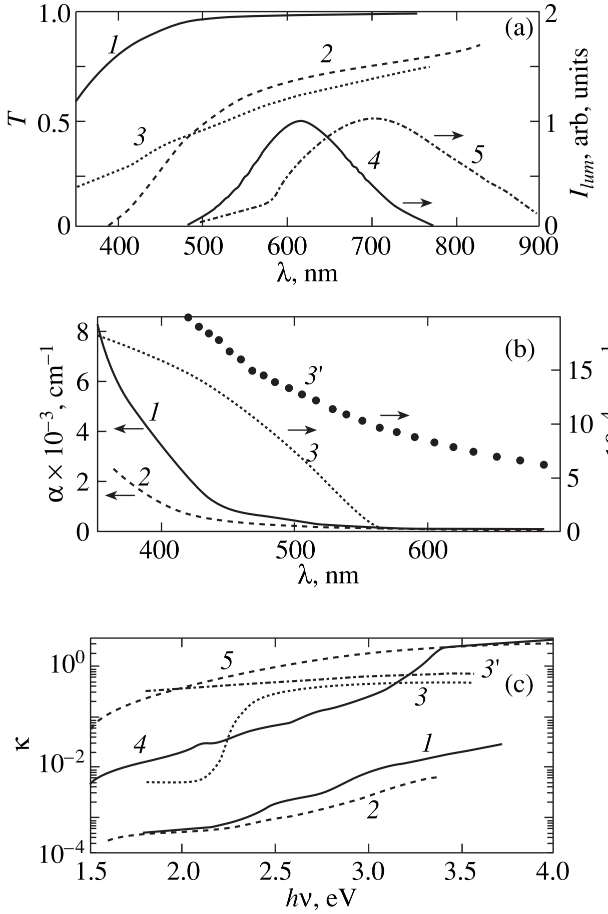

**Fig. 3.** (a) The transmittance curves for por-Si no. 1 (curve *1*), no. 2 (curve *2*), and a-Si (curve *3*) and the luminescence spectra of samples from the same series as por-Si no. 1 (curve *4*) and por-Si no. 2 (curve *5*). (b) The spectra of absorption coefficients, α(*h*ν). (c) The spectra of extinction coefficients, κ(*h*ν). The notation is the same as in Fig. 2d.

As can be seen from Fig. 2d, we obtained almost flat dispersion curves *n*(*h*ν) without a sharp rise at short wavelengths. They are very similar to the curves in [4] and [9] and differ sharply from those in [3], in which a significant rise followed by a peak on the *n*(*h*ν) curve near 3.5 eV was recorded. Much more of the crystalline phase may have remained in the sample analyzed in [3] than in our samples.

Figure 3a shows the transmittance curves *T*(λ) for the two por-Si samples and for the a-Si film deposited on a glass substrate. Also shown here for comparison are the luminescence spectra of samples belonging to the same series as por-Si no. 1 and por-Si no. 2 before their separation from the substrate. The measurements were made at room temperature. The results are seen to be consistent with the quantum-dimensional theory of photoluminescence; more specifically, the smaller the grain size in the porous structure, the larger the blueshift of the photoluminescence peak.

Figure 3b shows the absorption coefficients α(*h*ν) calculated from the transmittance curves, while Figure 3c (on a logarithmic scale for convenience) shows the corresponding dimensionless extinction coefficients κ(*h*ν) and published data for crystal and amorphous silicon. Since the correction for reflection for a-Si strongly affects α(*h*ν) and κ(*h*ν), the values without this correction being applied are indicated in Figs. 3b and 3c by curves *3*'. Finally, Fig. 4 shows ε'(*h*ν) and ε''(*h*ν) for our por-Si samples.

## 4. DISCUSSION

Our main objective was to establish a relationship between the optical properties of porous silicon and its composition, in particular, its porosity.

Some authors (see, e.g., [7]) estimated the porosity *p* of crystals, i.e., the fraction of voids in them, from a simple relation:

$$\alpha' = \alpha(1-p), \tag{4}$$

where α and α' are the absorption coefficients of the crystal and porous silicon, respectively. In our view, this relation is improper, because the energy absorbed in the material is proportional to the imaginary part of the dielectric function, ε'' = 2*n*κ, whereas κ character-

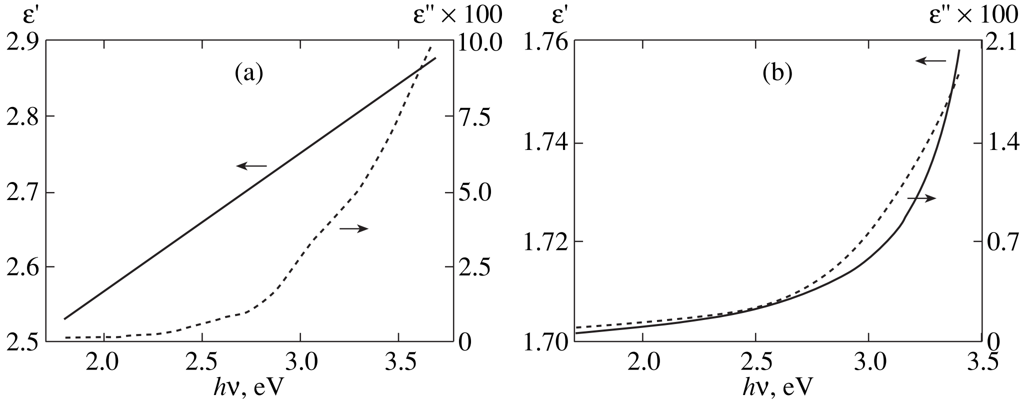

**Fig. 4.** Dielectric constants for por-Si samples nos. 1 (a) and 2 (b).

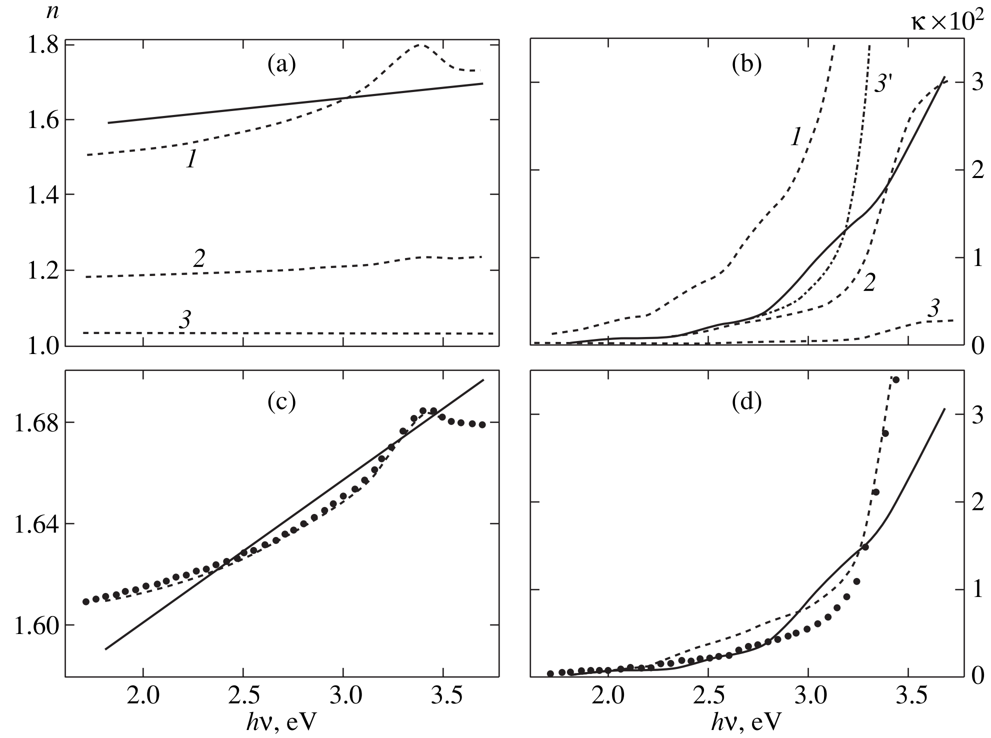

**Fig. 5.** Comparison of the experimentally measured $n(h\nu)$ and $\kappa(h\nu)$ curves for sample no. 1 (solid curves) with those calculated from Bruggeman's theory [17]: (a), (b) two-component system (curves *1–3*); (c), (d) three-component system (dotted curves) and four-component system (dashed curves). Dashed-dotted curve *3*' is the calculation using formula (4). Curves *1* correspond to the volume fraction $f_{c\text{-Si}}$ = 25.74%; curves *2*, to 12.12%; and curves *3* and *3*', to 2.15%.

izes only the rate of decrease in the amplitude of the light wave propagating through the medium.

When the scale length of the microstructure is smaller than the light wavelength (the long-wavelength limit), one may ignore the retardation effects and replace the multicomponent system by an effective medium with a macroscopic dielectric constant. By far the most accurate porosity estimate can be obtained in the approximation of Bruggeman's theory [17] for the effective medium, which is the most realistic for systems of irregularly shaped particles even for small porosities [18].

The effective dielectric constant $\varepsilon_{\text{eff}}$ of a multicomponent isotropic system is determined, according to [17], from the dielectric constants of its individual components, $\varepsilon_i$, whose volume fraction in the system is $f_i$:

$$\sum_i f_i \frac{\varepsilon_i - \varepsilon_{\text{eff}}}{\varepsilon_i + 2\varepsilon_{\text{eff}}} = 0, \quad \sum_i f_i = 1. \tag{5}$$

Based on this theory, we determined the porosity and composition of our porous silicon samples. We assumed that the crystals consisted mainly of crystal silicon, c-Si, but they could also contain an amorphous-phase (a-Si) impurity. We did not rule out the inclusion of quartz $SiO_2$ in the samples, which, according to [19], forms on the silicon pore surfaces. Thus, we considered a four-component system. The values of ε were taken from [15] for c-Si and from [16] for a-Si; we assumed that $\kappa = 0$ and $n = 1.5$ for $SiO_2$ (an approximately mean value between the refractive indices of α-$SiO_2$ and fused quartz) and $\varepsilon = 1$ for air.

Figures 5 and 6 show the calculated $n(h\nu)$ and $\kappa(h\nu)$ curves for the multicomponent system as a function of the volume fraction of its components, to reconcile them with the corresponding experimental curves for samples nos. 1 and 2.

At the first stage of our calculations, we considered a two-component system: c-Si + pores (Figs. 5a, 5b and 6a, 6b). It turned out that for the two crystals, no relation between c-Si and air allowed the dispersion curves of refractive indices and the absorption curves to be reconciled between themselves. The values of $n(h\nu)$ are always smaller than those measured experimentally if the volume fraction is such that the $\kappa(h\nu)$ curves are close. Conversely, if the $n(h\nu)$ curves can be reconciled, then the calculated $\kappa(h\nu)$ are much larger than those measured experimentally. Also shown in these figures for comparison are the absorption curves calculated using formula (4). They are lying considerably higher than those calculated from Bruggeman's theory.

Figures 5b and 6b clearly show how the absorption spectrum changes with $f_{c\text{-Si}}$: the band gap seemingly increases. Accordingly, recall that the popular opinion

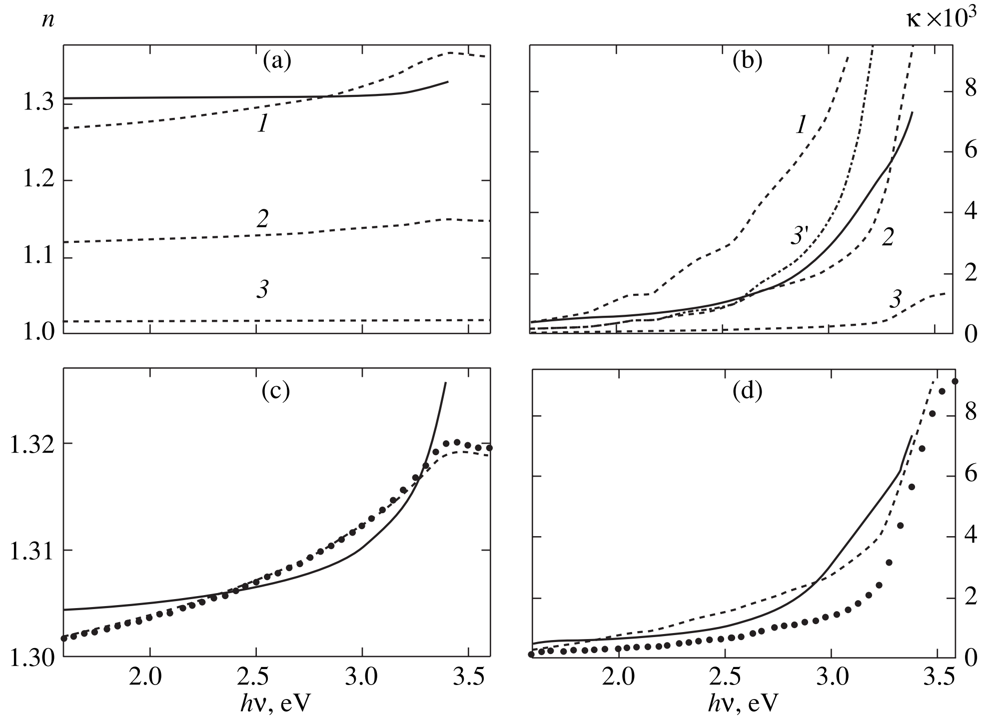

**Fig. 6.** Same as in Fig. 5 for por-Si sample no. 2. Curves *1* correspond to the volume fraction $f_{c\text{-Si}}$ = 16.7%; curves *2*, to 8.57%; and curves *3* and *3'*, to 1.287%.

about an increase in band gap with porosity and, consequently, with the appearance of a quantum level in filaments is based on the blueshift of the absorption edge. The edge shift is believed to result in crystal transparency in the visible spectral range. As we see from Fig. 3a, our samples also became transparent to visible light. However, if we estimate the absorption per unit remaining material, then the shift of the absorption edge is apparent, because invariable initial values of e for silicon with no quantum level, i.e., without an artificial increase in band gap, were laid in our calculations. Note in passing that neither the absorption spectra (Fig. 3) nor the dispersion curves of refractive index (Fig. 2) exhibit any features associated with the quantum levels responsible for the luminescence of these crystals (Fig. 3a). This may be determined by the spread in micrograin sizes, which gives a spread in absorption coefficients and strongly smoothens the resulting curve.

At the next stage of our calculations (three-component system), we introduced $SiO_2$ into the system to approach the experimental values of *n*(*h*ν). The results of this introduction are indicated in Figs. 5c, 5d and 6c, 6d by the dotted curves. As we see from the figures, the gap between the *n*(*h*ν) and κ(*h*ν) curves can be reduced by assuming that sample no. 1 contains up to 85% of $SiO_2$, while sample no. 2 contains about 49% of $SiO_2$ for c-Si volume fractions of 9 and 4%, respectively.

**Table 1**

| | *c*-Si | $SiO_2$ | *a*-Si | Porosity *p* |
|---|---|---|---|---|
| Sample 1 | 8.93% | 80.7% | 1.14% | 9.23% |
| Sample 2 | 3.2% | 49% | 0.53% | 47.23% |

At the final stage of our calculations (four-component system), we introduced an amorphous-phase (a-Si) impurity. As a result, the agreement between the experimental and theoretical curves improved, but we still failed to achieve close agreement. The best agreement between the measured and calculated *n*(*h*ν) and κ(*h*ν) curves (the dashed lines in Figs. 5c, 5d and 6c, 6d) was achieved for the fraction ratios presented in Table 1. Such volume fractions are obtained if our data and the published data [16] are used for the amorphous phase in samples nos. 1 and 2, respectively.

The possible large $SiO_2$ impurity in porous silicon was discussed in [19]. The authors pointed out that the formation of a 1.5-nm-thick oxide layer on the pore walls must result in a $SiO_2$ fraction in porous silicon exceeding 30%. In addition, an infrared absorption band was discovered at the $SiO_2$ characteristic oscillation frequency (approximately 1100 $cm^{-1}$). In the same frequency range (1060–1200 $cm^{-1}$), a strong absorption band of por-Si samples was also detected by Astrova *et al.* [6]. The authors emphasize that the smaller crystallites are surrounded by a much larger number of silicon bonds with oxygen and hydrogen. Finally, the spectra recently published in [20] exhibit a drastic increase in absorption at a frequency of 1100 $\text{cm}^{-1}$ during a long exposure in air (up to 450 days). Important information confirming the presence of a large fraction of $SiO_2$ impurity in porous silicon is also contained in [21]. An analysis of Auger spectra showed that an oxide phase is produced on the surface by anisotropic etching. The thickness of the recorded oxidized layer exceeded 100 nm. Based on two independent methods of investigation, ellipsometry and Rutherford backscattering, Belyakov *et al.* [9] concluded that their sample contained from 38.8 to 43.3% of $SiO_2$. As has already been noted, the refractive index and the behavior of $n(h\nu)$ determined in [9] are close to $n(h\nu)$ for our sample no. 2. In our case, the $SiO_2$ volume fraction also approaches its value in [9].

As for sample no. 1, the $SiO_2$ volume fraction that has to be substituted in Bruggeman's formulas to reconcile the $n(h\nu)$ and $\kappa(h\nu)$ curves proves to be very large (about 80%). This may seem surprising and may cast doubt on their validity. Therefore, we analyze the results by using data from [22]. The authors of [22] studied the change in the optical parameters of thin oxidized porous silicon films depending on how they were produced and thermally treated. The film composition was determined from the measured refractive index at $\lambda = 632.8$ nm by using the dependences calculated with Bruggeman's three-component model for an effective medium composed of layers with different initial porosities. Despite the fact that we did not subject our samples to special oxidation, it still was of interest to compare their parameters with those in [22], although the calculations in [22] were performed for a single wavelength, only for a three-component system, and disregarding the absorption.

An important assertion made in [22] is that the oxidized film is not an arbitrary mixture of three components and that the volume fraction of each of them changes during oxidation at the expense of another. The oxidation of each Si atom was shown to cause the volume of the solid phase associated with it to increase by a factor of 2.27. Therefore, when the initial two-component system Si + pores is oxidized, a new $SiO_2$ phase emerges and the volumes occupied by pores and silicon decrease. Since the possibility of oxidation depends on the volume of free space that $SiO_2$ can occupy, samples with different initial porosities $p_{\text{in}}$ behave differently. If $p_{\text{in}} > 56\%$, then after complete oxidation of the silicon skeleton, the sample remains porous and consists of $SiO_2$ and pores. If, alternatively, $p_{\text{in}} < 56\%$, then the pores completely disappear when the silicon skeleton has not yet been completely oxidized, resulting in a different two-component system: Si + $SiO_2$.

The possible compositions for the coexistence of three phases were calculated in [22] and shown in the plots there as the hatched part of the plane. A comparison of our data with those from [22] indicates that, first, both our samples fall within the region where three phases can coexist. Second, using relations from the above paper, we can determine the degree of oxidation and $p_{\text{in}}$ for our samples. The volume fraction of the silicon skeleton that oxidized is $x = f_{SiO_2}/2.27$. The total silicon volume fraction in the initial por-Si is the sum $f_{\text{Si}} = f_{c\text{-Si}} + f_{a\text{-Si}} + x$. The initial porosity is $p_{\text{in}} = 1 - f_{\text{Si}}$. The degree of oxidation is defined as $s = (x/f_{\text{Si}})100\%$. The results are given in Table 2. Thus, $p_{\text{in}} < 56\%$ in por-Si sample no. 1 and $p_{\text{in}} > 56\%$ in por-Si sample no. 2. This determined a very large difference in their compositions after a prolonged storage in air. However, in none of them did possible finite processes take place: there is no complete oxidation in por-Si no. 2 and no complete disappearance of pores in por-Si no. 1. Nevertheless, the degree of oxidation in both samples is very large.

**Table 2**

| | Volume fraction of skeleton (Si) oxidized, $x$ | Initial volume fraction of silicon, $f_{\text{Si}}$ | Initial porosity, $p_{in}$ | Degree of oxidation $s$ |
|---|---|---|---|---|
| Sample 1 | 35.5% | 46.2% | 53.8% | 77% |
| Sample 2 | 21.6% | 25.3% | 75.0% | 85% |

If two plots from [22] are used (s-n and p-n), then $p_{\text{in}} = 52.5$ and 53.5% for por-Si no. 1 and $p_{\text{in}} = 73$ and 74% for por-Si no. 2. These values are very close to those from Table 2. Therefore, the entire set of our data suggests that the processes in our crystals followed the scheme proposed in [22]. The recorded oxidation of the por-Si samples may be responsible for the attenuation of their photoluminescence with time.

**Comparison with theoretical calculations of the por-Si band structure.** While summarizing the results of fitting the experimental curves by theoretical dependencies, it should be emphasized that close agreement between them still cannot be achieved. We believe that this is because we used in our calculations the $\varepsilon$ values for pure silicon as the initial data for c-Si. This automatically assumes that the band structure and selection rules for band-to-band transitions and, consequently, the corresponding oscillator strengths in porous silicon remain the same as those in the initial single crystal. This, of course, is a rough approximation. When a porous structure is formed, the far order in the grating inevitably breaks down, which must weaken direct transitions at the center of Brillouin's zone $k = 0$ (3.4 eV, 4.2 eV). This precisely takes place during the formation of amorphous silicon, when, on the one hand, strong peaks of absorption and dispersion $n(h\nu)$ vanish in the ultraviolet, and, on the other hand, the absorption and $n$ increase in the red spectral range near the indirect absorption edge. This can be seen from Figs. 2d and 3c and suggests a reduction in the indirectness of amorphous silicon.

The electronic structure and optical properties of porous silicon were theoretically and consistently studied in detail in [23, 24]. In [24], the material was represented as a set of tightly bound quantum wires with lengths of the order of a micron and widths of several nanometers (the length-to-width ratio is of the order of 1000 : 1). The authors calculated the band gap and the position of the exciton level as a function of wire width, as well as the imaginary part of the dielectric function and the absorption spectra.

In [23], porous silicon was modeled by a supercell, in which "columns" of material were removed with a certain periodicity. The same porosity, i.e., the percentage of the removed material, was achieved by different methods: either narrow columns were removed frequently or wide columns were removed rarely. The authors analyzed the dependences of band structure and absorption spectrum (the imaginary part of $\varepsilon$) on the sizes of the removed column and on porosity. Their calculations showed that the position of the absorption edge is determined by two competing effects. First, transitions are assisted by the scattering of carriers on the lattice of pores, which effectively decrease the "indirectness" of por-Si and results in a redshift of the absorption edge. Second, quantum confinement of carriers increases the band gap. The latter effect significantly depends on the thickness of the silicon skeleton that remains between the pores.

It was of interest to compare our experimental data with theoretical calculations in order to find out whether the simplified theoretical models correspond to actual crystal structures.

The comparison indicates that the absolute values of $\varepsilon''$ and $\alpha$ (cm$^{-1}$) calculated in [24] and the shape of the $\alpha(h\nu)$ curves including exciton states are far from our experimental dependences. As for the polarization anisotropy in absorption relative to the direction of the quantum-wire axis predicted in [24], we could not detect it under our experimental conditions. Our structures were dots (grains) rather than wires.

The $\varepsilon(h\nu)$ curves calculated in [23] for samples with different porosities clearly show an increase in the red tail and a decrease in the short-wavelength peaks with increasing $p$, in qualitative agreement with our data. However, the minimum percentage of the remaining crystal material ($1-p$) for which the calculation was performed (about 23.4%) is considerably higher than that estimated in our samples, which complicates a comparison. The authors obtained a spectrum that was greatly blueshifted compared to c-Si. This spectrum consisted of several peaks that emerge during transitions between almost flat bands. Our measurements of the absorption spectrum do not cover this spectral range, which further complicates a comparison of experiment with theory. Unfortunately, neither $\varepsilon'(h\nu)$ nor $n(h\nu)$ curves are given in [23] and [24], with which a comparison can be made in the visible spectral range. Nevertheless, based on the measured $n(h\nu)$ curves, no ultraviolet shift of the absorption spectrum takes place in our samples, but only a general lowering of the $\kappa(h\nu)$ curve is observed. Thus, the resonant frequency $\omega_0 = 3.61$ eV estimated in sample no. 2 from the curvature of the $n(h\nu)$ curve (Fig. 2b), which characterizes the position of the maximum of an averaged absorption bands, falls between the first (3.4 eV) and second (4.2 eV) extrema of the $\kappa(h\nu)$ curve for crystal silicon [15].

## 5. CONCLUSION

We have comprehensively studied the properties of the dielectric function for two por-Si samples from different series by measuring the dispersion curves of refractive index $n(h\nu)$ and the absorption spectra $\alpha(h\nu)$ and $\kappa(h\nu)$. Based on the dependences derived from Bruggeman's theory [17], we determined the porosities of these samples and the possible percentages of silica ($SiO_2$) and the amorphous phase (a-Si) in them. The analysis was performed by assuming that the dielectric function $\varepsilon$ of crystal silicon c-Si was invariable during the formation of a porous structure. In this case, the $n(h\nu)$ and $\kappa(h\nu)$ curves can be roughly reconciled for each sample if they contain a large fraction of silica. The oxidation of the silicon skeleton to produce $SiO_2$ appears to be one of the main reasons for the "ageing" of porous silicon samples, i.e., the deterioration of their quality and the attenuation of photoluminescence.

At present, we know no theoretical calculations that could reconcile the absorption and dispersion of our porous silicon samples without including $SiO_2$ in their composition. The available theories that allow for quantum-dimensional effects in porous silicon (quantum wires, supercells) fail to describe our experimental data.

## ACKNOWLEDGMENTS

We wish to thank R.D. Fedorovich, D.B. Dan'ko, and V.Yu. Timoshenko for providing the samples analyzed. We are also grateful to O.V Sakhno and R.V. Hristosenko for help in carrying out the experiment.

## REFERENCES

1. S. V. Svechnikov, A. V. Sachenko, G. A. Sukach, *et al.*, Optoélektron. Poluprovodn. Tekh. **27**, 3 (1994).
2. Y. Kanemitsu, H. Uto, and Y. Masumoto, Phys. Rev. B **48**, 2827 (1993).
3. N. Koshida, H. Koyata, Y. Suda, *et al.*, Appl. Phys. Lett. **63**, 2774 (1993).
4. L.-Y. Chen, X.-Y. Hou, D.-M. Huang, *et al.*, Jpn. J. Appl. Phys., Part 1 **33**, 1937 (1994).
5. Y. H. Xie, M. S. Hybertsen, and W. L. Wilson, Phys. Rev. B **49**, 5386 (1994).
6. E. V. Astrova, A. A. Lebedev, A. D. Remenyuk, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 1649 (1995) [Semiconductors **29**, 858 (1995)].
7. D. G. Yarkin, E. A. Konstantinova, and V. Yu. Timoshenko, Fiz. Tekh. Poluprovodn. (St. Petersburg) **29**, 669 (1995) [Semiconductors **29**, 348 (1995)].

8. A. N. Obraztsov, V. A. Karavanskiĭ, H. Okushi, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 1001 (1998) [Semiconductors **32**, 896 (1998)].
9. L. V. Belyakov, T. L. Makarova, V. I. Sakharov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **32**, 1122 (1998) [Semiconductors **32**, 1003 (1998)].
10. O. V. Vakulenko, O. I. Datsenko, V. A. Makara, *et al.*, Ukr. Fiz. Zh. **43**, 348 (1998).
11. M. S. Brodin, V. M. Bykov, D. B. Dan'ko, *et al.*, Ukr. Fiz. Zh. **40**, 933 (1995).
12. M. S. Brodyn, V. N. Byckov, D. B. Dan'ko, *et al.*, Proc. SPIE **2780**, 324 (1995).
13. P. G. Borzyak and R. D. Fedorovich, Fiz. Tverd. Tela (Leningrad) **11**, 3020 (1960) [Sov. Phys. Solid State **2**, 2683 (1961)].
14. A. Penzkofer, E. Drotleft, and W. Holzer, Opt. Commun. **158**, 221 (1998).
15. D. E. Aspnes and A. A. Studna, Phys. Rev. B **27**, 985 (1983).
16. D. T. Pierce and W. E. Spicer, Phys. Rev. B **5**, 3017 (1972).
17. D. A. Bruggemann, Ann. Phys. (Leipzig) **24**, 636 (1935).
18. W. Theiss, Surf. Sci. Rep. **29**, 91 (1997).
19. H. Munder, C. Andrzejak, M. G. Berger, *et al.*, Appl. Surf. Sci. **56-58**, 6 (1992).
20. B. V. Kamenev, E. A. Konstantinova, P. K. Kashkarov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 753 (2000) [Semiconductors **34**, 728 (2000)].
21. E. F. Venger, T. Ya. Gorbach, L. A. Matveeva, *et al.*, Zh. Éksp. Teor. Fiz. **116**, 1750 (1999) [JETP **89**, 948 (1999)].
22. E. V. Astrova, V. B. Voronkov, F. V. Remenyuk, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **33**, 1264 (1999) [Semiconductors **33**, 1149 (1999)].
23. M. Cruz, M. R. Beltran, C. Wang, *et al.*, Phys. Rev. B **59**, 15381 (1999).
24. G. D. Sanders and Y. C. Chang, Phys. Rev. B **45**, 9202 (1992).

*Translated by V. Astakhov*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 372–379.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 420–429.*
*Original Russian Text Copyright © 2001 by Znamenskiĭ, Manykin, Orlov, Petrenko, Yukina, Malyukin, Borisov, Zhmurin, Lebedenko, Grinev.*

**SOLIDS**
**Structure**

# The Nature of Activation Centers in $Y_2SiO_5 : Pr^{3+}$, $Gd_2SiO_5 : Pr^{3+}$, and $Lu_2SiO_5 : Pr^{3+}$ Crystals

**N. V. Znamenskiĭ[1], *, É. A. Manykin[1], Yu. V. Orlov[1], E. A. Petrenko[1], T. G. Yukina[1], Yu. V. Malyukin[2], R. S. Borisov[2], P. N. Zhmurin[2], A. N. Lebedenko[2], and B. V. Grinev[2]**

[1]*Kurchatov Institute Russian Research Centre, pl. Akademika Kurchatova 1, Moscow, 123182 Russia*

[2]*Institute of Single Crystals, National Academy of Sciences of Ukraine, pr. Lenina 60, Kharkov, 61001 Ukraine*
** e-mail: zarja@isssph.kiae.ru*

Received February 22, 2001

**Abstract**—A complex study of the energy spectra and relaxation channels for the excitation energy of activation centers in $Y_2SiO_5 : Pr^{3+}$, $Lu_2SiO_5 : Pr^{3+}$, and $Gd_2SiO_5 : Pr^{3+}$ was performed. An analysis of the low-temperature optical spectra showed that the energy parameters and the character of field splitting of the $^1D_2$ and $^3H_4$ activator ion terms were substantially different in crystals of different crystallographic types. The pseudosymmetry effect was observed in splitting of the $^1D_2$ and $^3H_4$ terms of $Pr^{3+}$ ions situated in nonequivalent crystal lattice cation sites of $Y_2SiO_5$ and $Lu_2SiO_5$. Activator ions nonuniformly populated nonequivalent cation sites of the $Y_2SiO_5$ crystal lattice. At high activator ion concentrations (>1 at. %), luminescence decay in $Y_2SiO_5$ could not be described by a simple exponential time dependence. The complex luminescence decay law was caused by activator ion excitation energy migration and capture by acceptors. The role of energy acceptors was played by activator ion dimers. © *2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

Oxiorthosilicate crystals offer much promise for the development of new lasing and scintillating materials, in the first place, because of their large isomorphic capacity [1–5]. Effective scintillating systems have already been found among oxiorthosilicates and have been extensively studied; these are $Y_2SiO_5 : Ce^{3+}$, $Gd_2SiO_5 : Ce^{3+}$, and $Lu_2SiO_5 : Ce^{3+}$ [2, 6–8]. The mechanism of scintillating response in these system is, however, not well understood. For creating solid-state lasers with the mechanism of population inversion through frequency conversion toward higher frequencies [9], of special interest is activation of oxiorthosilicates by $Pr^{3+}$ ions. Diode IR pumping of $Pr^{3+}$ ions can be used for multifrequency laser generation in the blue-green spectral region [3, 10].

Rare-earth metal oxiorthosilicates form monoclinic crystals [1, 2, 4, 5, 11]. Oxiorthosilicate crystals, however, form two structural classes in cation sites depending of the radius of the rare-earth metal ion [4, 5, 11]. The first and second classes are formed by the La…Tb and Dy…Yb (including Y) ions, respectively. $Y_2SiO_5$ (YSO) and $Lu_2SiO_5$ (LSO) crystals have the same structural type [1, 4, 5, 11]. The YSO unit cell parameters are $a$ = 14.43 Å, $b$ = 10.41 Å, $c$ = 6.733 Å, $\beta$ = 122.13°, and $V$ = 856.1 Å$^3$; for the LSO lattice, $a$ = 14.33 Å, $b$ = 10.32 Å, $c$ = 6.671 Å, $\beta$ = 122.3°, and $V$ = 833.8 Å$^3$ [1, 5, 11]. $Gd_2SiO_5$ (GSO) crystals are of a different structural type. Their unit cell parameters are $a$ = 9.16 Å, $b$ = 7.09 Å, $c$ = 6.83 Å, $b$ = 107.58°, and $V$ = 422.9 Å$^3$ [5]. In the YSO lattice, one of the cation sites is a distorted octahedron with four bound and two free oxygen atoms; the Y–O distance varies in the range 2.21–2.33 Å [4, 5, 11]. The second cation site is a coordination polyhedron formed by five bound and two free oxygen atoms [4, 5, 11]. In GSO, the coordination polyhedron of one cation site contains four bound and three free oxygen atoms (the mean Gd–O distance is 2.39 Å), and the coordination polyhedron of the second cation site contains eight bound and one free oxygen atoms (the mean Gd–O distance is 2.29 Å) [4, 5]. Nonequivalence of cation sites in oxiorthosilicates is determined by the special features of the structure of coordination polyhedra and differences in the mean distances between the rare-earth metal ion and oxygen atoms in polyhedron vertices [1, 2, 4, 5, 11]. The introduction of rare-earth metal ions as activators into oxiorthosilicate crystals results in the formation of two types of optical centers [5, 12].

A complex structure of spectral lines was determined for both types of $Pr^{3+}$ optical centers in the YSO : $Pr^{3+}$ crystal by the method of spectral hole burning [13]. The spectral lines of $Pr^{3+}$ optical centers comprised several spectral contours, which coalesced into one inhomogeneously broadened spectral contour [13]. The complex structure of spectral lines is evidence that the $Pr^{3+}$ ion can have several different states (positions) in YSO cation sites of both types. Thermally stimulated $Pr^{3+}$ ion tran-

1063-7761/01/9302-0372$21.00 © 2001 MAIK “Nauka/Interperiodica”

sitions between these states create an additional channel for phase relaxation on resonance optical transitions [14].

The practical aspects of the use of activated oxiorthosilicates are related to several fundamental problems concerning the microstructure of activation centers and their energy spectrum, the dynamics of electronic transitions, and energy relaxation channels. The large isomorphic capacity of oxiorthosilicates with respect to activator ions (see above) has not been given a satisfactory consistent explanation as yet. It remains unclear how activator ions populate nonequivalent cation sites in oxiorthosilicate lattices. Another open question is whether or not electronic excitation energy exchange can occur between the centers that arise as a result of activator ion localization in nonequivalent cation sites of oxiorthosilicate crystal lattices.

## 2. EXPERIMENTAL TECHNIQUE

The optical spectra of the crystals were recorded on an automated spectrofluorimeter based on an NDR-23 monochromator. The luminescence spectra were excited by an organic dye frequency-tunable laser. Luminescence decay was recorded by time-correlated single photon counting [15].

The YSO : $Pr^{3+}$, LSO : $Pr^{3+}$, and GSO : $Pr^{3+}$ crystals were grown by the Chokhral'skii method. The concentration of activator ions was 0.3 at. % in LSO and GSO and 0.3, 0.6, and 1.8 at. % in YSO.

Low-temperature measurements were taken using a helium optical cryostat. The crystals were placed in helium vapors.

## 3. EXPERIMENTAL RESULTS

The crystal field characteristics of nonequivalent cation sites in oxiorthosilicates and the conditions of populating them by activator ions were studied spectroscopically. In particular, the low-temperature optical spectra of $Pr^{3+}$ activator ions were recorded. The $^1D_2$, $^3P_0$, $^3P_1$, $^1I_6$, and $^3P_2$ states of the $Pr^{3+}$ ion were observed in the visible optical spectrum region [16, 17]. The spectral lines corresponding to optical transitions between the fundamental $^3H_4$ term and the $^1I_6$, $^3P_1$, and $^3P_2$ terms were strongly mixed, which complicated their correct interpretation [16, 17]. At low temperatures, the spectral lines of optical transitions between the Stark sublevels of the fundamental $^3H_4$ term and the $^3P_0$ and $^1D_2$ terms could fairly easily be assigned [12, 16, 17]. We used the spectral region of optical transitions involving the $^3H_4$ and $^1D_2$ terms on the assumption that the special features of the ligand field microstructure should influence the splitting of the $^1D_2$ term.

Fragments of the absorption spectra of YSO, LSO, and GSO crystals activated by $Pr^{3+}$ are shown in Fig. 1. The spectral lines of each absorption spectrum could be divided into two groups according to their widths (Fig. 1).

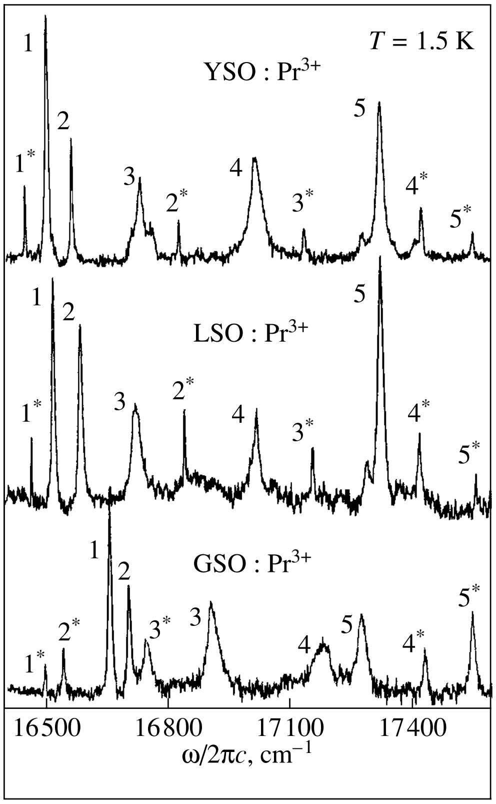

**Fig. 1.** Fragments of optical absorption spectra of YSO : $Pr^{3+}$, LSO : $Pr^{3+}$, and GSO : $Pr^{3+}$. Spectral lines in groups are marked by numbers and numbers with asterisks.

With GSO : $Pr^{3+}$, this property of the spectrum was less manifest (Fig. 1c). Each group comprised five spectral lines. The lines progressively broadened as their number increased (Fig. 1). This behavior is characteristic of spectral lines corresponding to one multiplet of rare-earth ions [16–18].

Selectively exciting luminescence into separate groups of spectral lines (Fig. 1) gave two different luminescence spectra for each crystal (Figs. 2, 3). The line at the highest frequency in the luminescence spectra of both types (Figs. 2, 3) was in resonance coincidence with the line observed at the lowest frequency in the absorption spectra, 1 or 1* (Fig. 1). When luminescence of one type was selectively excited, we did not observe luminescence of the other type.

Spectral lines whose intensity depended on temperature could be observed in the optical spectra of the crystals as temperature increased. For instance, the $\eta$, $\zeta$, $\xi$, and $\gamma$ lines (Fig. 3) were recorded at $T$ = 80 K. The $\eta$ and $\zeta$ lines were in resonance coincidence with spectral lines 2 and 3, and the $\xi$ spectral line coincided with spectral line 2* of the corresponding luminescence spectrum (Fig. 3). The $\eta$, $\zeta$, and $\xi$, spectral lines corresponded to absorption from thermally populated Stark components of the $^3H_4$ term. At $T$ = 80 K, two $^3H_4$ term

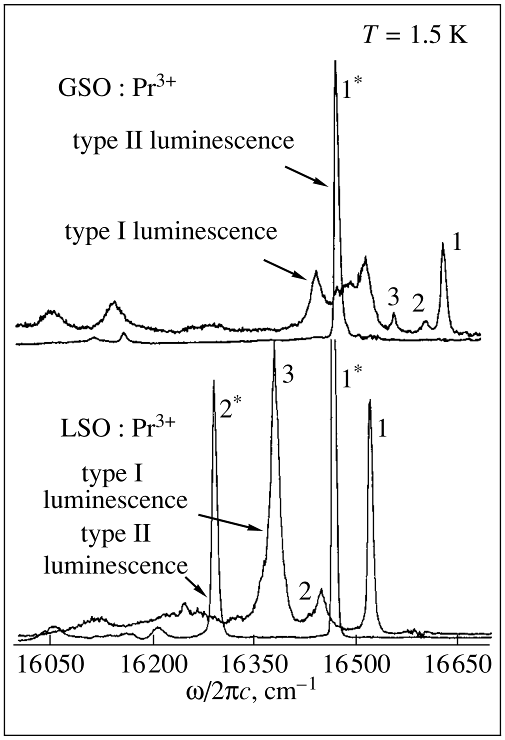

**Fig. 2.** Fragments of luminescence spectra of LSO : $Pr^{3+}$ and GSO : $Pr^{3+}$ crystals obtained by selectively exciting groups of spectral lines.

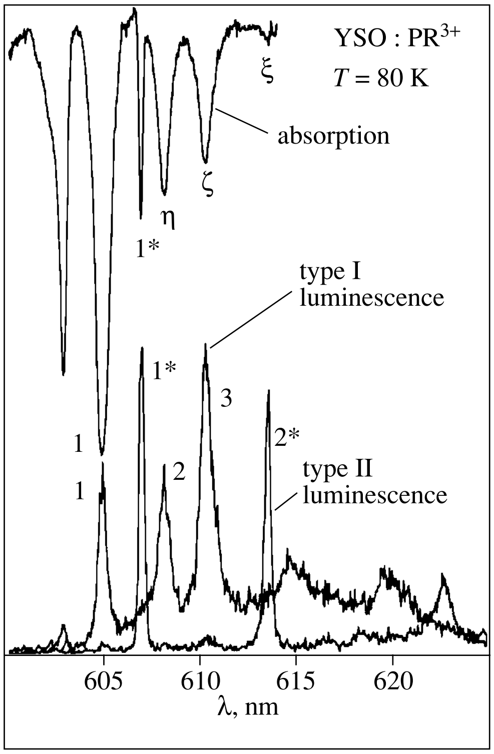

**Fig. 3.** Fragments of optical spectra of YSO : $Pr^{3+}$ crystals at 80 K.

Stark components were populated in optical centers of the first type, and only one Stark component was populated in centers of the second type. For a similar reason, the luminescence spectra of the first type contained the $\gamma$ spectral line. This line is caused by luminescence from the second thermally populated Stark component of the $^1D_2$ term. An analysis of the temperature-dependent spectral lines allows the spectral lines in the optical spectra of the crystals to be unambiguously interpreted and assigned.

To inquire into the mode of the distribution of activator ions over nonequivalent cation sites, we studied the concentration dependence of the optical spectra of YSO : $Pr^{3+}$ crystals. In the first place, an increase in the concentration of activator ions resulted in usual concentration broadening of spectral lines (Figs. 1, 4). The integrated intensity of absorption by YSO : $Pr^{3+}$ increased proportionally to the total concentration of activator ions. To reduce errors, we studied this dependence for two spectral lines, 1 and 1* (Figs. 1, 4). The total area under spectral lines 1 and 1* increased proportionally to the concentration of activator ions. The areas under lines 1 and 1*, however, changed differently. The area of spectral line 1 increased in the ratio 1 : 2.2 : 6.5, and that of spectral line 1*, in the ratio 1 : 1.6 : 3.2.

At a high concentration of activator ions, the spectra contained well-defined spectral lines $\delta_1$ and $\delta_2$ (Fig. 4). The frequency intervals between spectral lines 1* and $\delta_1$ and between 1 and $\delta_2$ were 8.9 and 21.4 $cm^{-1}$, respectively. The same lines, although virtually obscured by noise, were present in the absorption spectrum of the sample with an intermediate concentration of $Pr^{3+}$ ions. In the

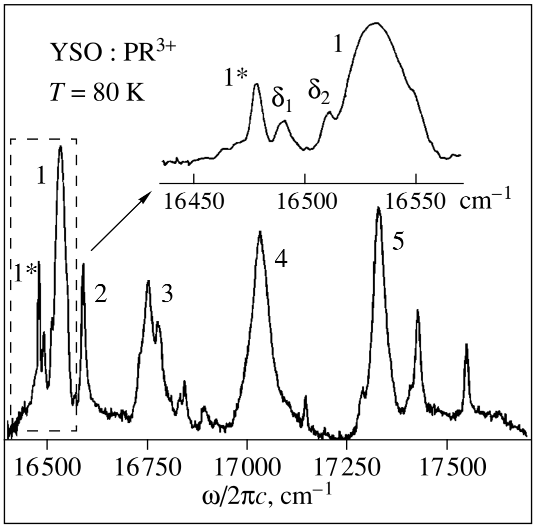

**Fig. 4.** Absorption spectrum of YSO : $Pr^{3+}$ crystals with a high (1.8 at. %) concentration of activator ions.

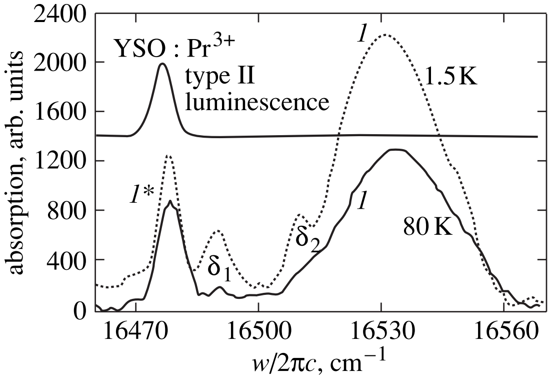

**Fig. 5.** Temperature dependences of the intensity of spectral lines of activator ion dimers in YSO : Pr$^{3+}$.

crystal with the lowest concentration of activator ions, the $\delta_1$ and $\delta_2$ spectral lines were absent. The intensity of the $\delta_1$ and $\delta_2$ spectral lines depended on temperature. At 80 K, these lines virtually disappeared (Fig. 5). We did not observe spectral lines with the same frequencies as $\delta_1$ and $\delta_2$ in the luminescence spectrum (Fig. 5). Note that spectral lines 5, 3, 4*, and 2* were also characterized by a certain structure (Fig. 1). This structure was, however, independent of the concentration of activator ions.

Luminescence decay after pulsed excitation in YSO : Pr$^{3+}$, LSO : Pr$^{3+}$, and GSO : Pr$^{3+}$ crystals with the lowest concentration of activator ions (0.3 at. %) was described by an exponential time dependence for both optical centers (Fig. 6). The decay constants did not depend on temperature in the temperature range 1.5–80 K. The decay constants for two optical centers were $\tau^{(1)} = 108$ and $\tau^{(2)} = 145$ μs in YSO, $\tau^{(1)} = 64$ and $\tau^{(2)} = 82$ μs in LSO, and $\tau^{(1)} = 39$ and $\tau^{(2)} = 50$ μs in GSO. In all crystals, the ratio between the luminescence decay constants for centers of the first and second type equaled 1.3.

The shape of the luminescence decay curve for optical centers of the first type changed appreciably in the YSO : Pr$^{3+}$ crystal with a 0.6 at. % concentration of activator ions. Precisely these ions make up the larger fraction of activator ions. At a 1.8 at. % concentration of activator ions, a strong change in the luminescence decay curve was observed for optical centers of both types (Fig. 7). The following tendency was observed for optical centers of the first type. The higher the concentration of activator ions was, the stronger luminescence decay deviated from the exponential law. The deviation was more manifest at the initial decay stage (Fig. 7).

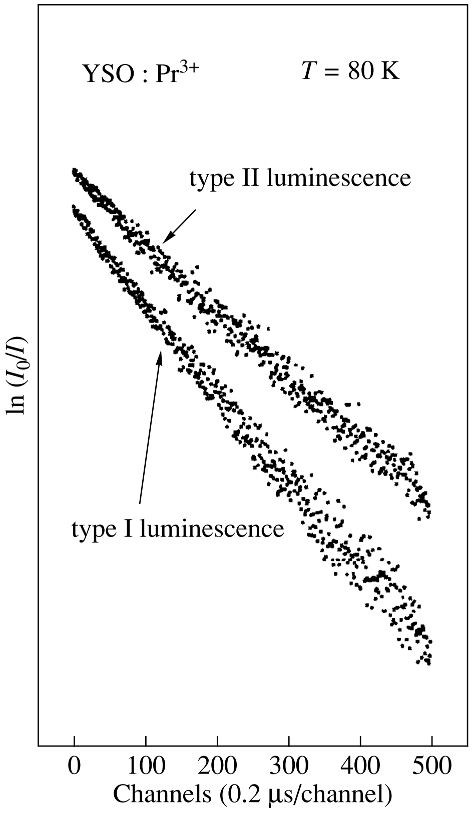

**Fig. 6.** Luminescence decay curves for optical centers in YSO : Pr$^{3+}$ with the lowest concentration of activator ions (0.3 at. %).

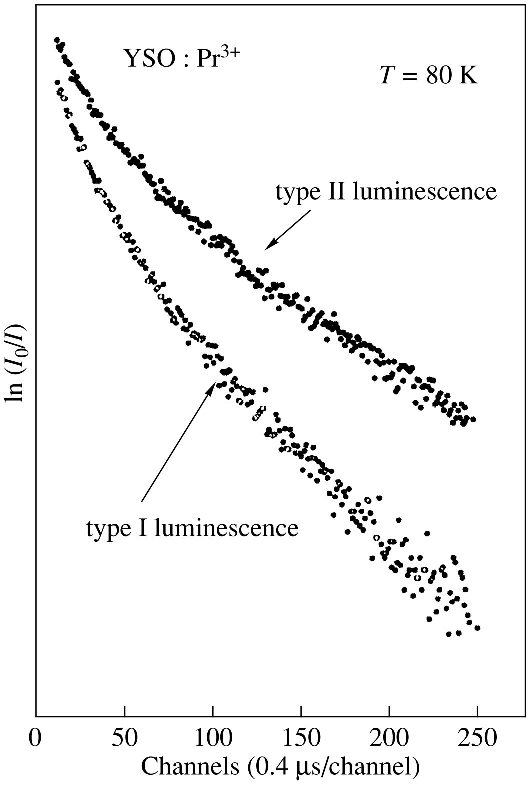

**Fig. 7.** Luminescence decay curves for optical centers in YSO : Pr$^{3+}$ with the highest concentration of activator ions (1.8 at. %).

**Table 1**

| | | YSO : $Pr^{3+}$ | | LSO : $Pr^{3+}$ | | GSO : $Pr^{3+}$ | |
|---|---|---|---|---|---|---|---|
| | | type I | type II | type I | type II | type I | type II |
| Term $^1D_2$ | 1 | 0 (16529.2) | 0 (16477.3) | 0 (16521.7) | 0 (16466.9) | 0 (16657.9) | 0 (16496.7) |
| | 2 | 59.6 | 364 | 66 | 373.7 | 42.1 | 43.8 |
| | 3 | 224.9 | 667.9 | 202.2 | 688.4 | 248 | 248.7 |
| | 4 | 501.2 | 948 | 498.2 | 951.7 | 525.9 | 940 |
| | 5 | 801.3 | 1071.5 | 801.8 | 1091.4 | 621.8 | 1056.7 |
| Term $^3H_4$ | 1 | 0 | 0 | 0 | 0 | – | – |
| | 2 | 88.4 | – | 73.4 | 176 | – | – |
| | 3 | 146.3 | – | 140.7 | 263 | – | – |
| | 4 | 207.9 | – | 198.2 | 308.7 | – | – |
| | 5 | 263.7 | – | 254 | 332.2 | – | – |
| | 6 | 286.2 | – | 275 | 351.7 | – | – |
| | 7 | 344.1 | – | 288 | 381.5 | – | – |
| | 8 | 366.3 | – | 304 | 430 | – | – |
| | 9 | 388.5 | 418 | 406 | 499.7 | – | – |

Note: Splitting parameters of the $^1D_2$ and $^3H_4$ terms are in $cm^{-1}$ with respect to the positions of spectral lines 1 and 1*.

For the crystal with the highest concentration of activator ions, the shape of the luminescence decay curve for optical centers of both types depended on temperature. In the crystal with an intermediate concentration of activator ions, a weak temperature dependence was only observed for optical centers of the first type.

## 4. ANALYSIS OF EXPERIMENTAL RESULTS

The presence of spectral lines with essentially different widths (Fig. 1) and selective excitation of luminescence of two types with different decay constants lead us to assert that the two groups of spectral lines in the absorption spectra of the crystals (Fig. 1) correspond to $Pr^{3+}$ optical centers of two types. These centers are formed as a result of $Pr^{3+}$ substitution in nonequivalent cation sites of YSO, LSO, and GSO [1, 2, 4, 5]. The presence of five spectral lines in each group (Fig. 1) is evidence of very low symmetry of cationic sites in YSO, LSO, and GSO. The degeneracy of the $^1D_2$ term of the $Pr^{3+}$ ion is completely removed under the crystal field action. At helium temperatures, only the lowest Stark component of the $^3H_4$ fundamental term is populated, and the five spectral lines of each group (Fig. 1) can therefore be unambiguously related to the five Stark components of the $^1D_2$ term. This allows us to determine the energy parameters of the splitting of the $^1D_2$ term of $Pr^{3+}$ ions under the action of YSO, LSO, and GSO crystal fields. The results are summarized in Table 1. The parameters and the character of the splitting of the $^1D_2$ term are different for different $Pr^{3+}$ centers. In YSO and LSO crystals of the same crystallographic type [1, 2, 4, 5], the energy parameters of the splitting of the $^1D_2$ term and the arrangement of the Stark components are similar (see Fig. 1 and Table 1). The unambiguous correspondence of the spectral lines in the absorption spectra (Fig. 1) to the Stark components of the $^1D_2$ term gives grounds for analyzing the special features of the arrangement of both. For optical centers of the first type in YSO and LSO crystals, the smallest energy interval of ~60 $cm^{-1}$ separates spectral lines 1 and 2. Line 3 is situated close to lines 1 and 2. The interval between the first three lines and lines 4 and 5 is substantially larger, about 290 $cm^{-1}$ (Fig. 1). For optical centers of the second type, the smallest energy interval separates lines 4* and 5*. Line 3* is situated somewhat lower in energy. Lines 1* and 2* are separated from line 3* by an interval of ~300 $cm^{-1}$. A qualitative conclusion can be drawn that the $^1D_2$ Stark components of optical centers of the second type are arranged in an order inverse to that characteristic of optical centers of the first type.

Certain splitting characteristics of the $^3H_4$ term of activator $Pr^{3+}$ ions in YSO and LSO crystals are similar to those observed for the $^1D_2$ term. The energy parameters of splitting of the $^3H_4$ term for two optical centers can be found from the arrangement of lines in the low-temperature luminescence spectra (Figs. 2, 3). Unfortunately, some spectral lines corresponding to optical transitions to the high-energy Stark components of the $^3H_4$ term could not be interpreted unambiguously (Figs. 2, 3). For this reason, Table 2 contains incomplete data on the splitting of the $^3H_4$ term. The $^3H_4$ term splitting parameters in YSO and LSO are very close to each other. The energy

**Table 2**

| $T$, K | | YSO 0.3 at. % $Pr^{3+}$ | | YSO 0.6 at. % $Pr^{3+}$ | | YSO 1.8 at. % $Pr^{3+}$ | |
|---|---|---|---|---|---|---|---|
| | | type I | type II | type I | type II | type I | type II |
| 77 | $\tau_0$, s | $108 \times 10^{-6}$ | $145 \times 10^{-6}$ | $108 \times 10^{-6}$ | $145 \times 10^{-6}$ | $108 \times 10^{-6}$ | $145 \times 10^{-6}$ |
| | $\alpha$ | 0 | 0 | 0.4 | 0.14 | 2.3 | 1.98 |
| | $\beta$ | 0 | 0 | 0.19 | 0 | 1.24 | 0.99 |
| | $c_a$, $cm^{-3}$ | – | – | $3.1 \times 10^{19}$ | $10^{19}$ | $1.8 \times 10^{20}$ | $1.5 \times 10^{20}$ |
| | $D$, $cm^3\ s^{-1}$ | – | – | $6.2 \times 10^{-12}$ | – | $6.9 \times 10^{-12}$ | $3.5 \times 10^{-12}$ |
| 1.5 | $\tau_0$, s | – | – | – | – | $108 \times 10^{-6}$ | $145 \times 10^{-6}$ |
| | $\alpha$ | – | – | – | – | 2.3 | 1.98 |
| | $\beta$ | – | – | – | – | 0.73 | 0.54 |
| | $c_a$, $cm^{-3}$ | – | – | – | – | $1.8 \times 10^{20}$ | $1.5 \times 10^{20}$ |
| | $D$, $cm^3\ s^{-1}$ | – | – | – | – | $5.7 \times 10^{-12}$ | $2.5 \times 10^{-12}$ |

intervals separating the two lowest $^3H_4$ Stark components are substantially different for optical centers of the first and second types, as with $^1D_2$ term splitting. Qualitatively and ignoring frequency intervals, the arrangement of the $^1D_2$ and $^3H_4$ term Stark components can be said to be quasi-inverse for two $Pr^{3+}$ optical centers in YSO and LSO crystals. This effect is similar to the splitting of the $^2D$ term of the $Ti^{3+}$ ion in ligand fields of tetrahedral and octahedral symmetry [19, 20]. The phenomena under consideration can be treated in quasi-symmetry terms, because nonequivalent cation sites have the lowest point symmetry group possible [1, 2, 4, 5].

Unlike YSO : $Pr^{3+}$ and LSO : $Pr^{3+}$ crystals, GSO : $Pr^{3+}$ gives absorption spectra in which spectral lines are arranged similarly for optical centers of both types: the first group is 1–2–3 (1*–2*–3*), and the second one is 4–5 (4*–5*) (Fig. 1).

It follows from a comparison of the splitting parameters of the $^1D_2$ and $^3H_4$ terms that the $Pr^{3+}$ optical centers of the second type (narrower spectral lines, Fig. 1) in YSO, LSO, and GSO crystals experience stronger crystal field perturbation. It might seem that the luminescence decay constant should be lower for centers of the second type. Indeed, parity selection rule restrictions on optical electric dipole transitions within the $f$ shell of rare-earth metal ions are removed by crystal field [17–20]. It is likely that, in the systems under consideration, an important role is played not only by the amplitude but also by crystal field quasi-symmetry in the region of admixture center localization. The contribution of odd harmonics to ligand field expansion is therefore smaller for optical centers of the second type than for first-type centers. Precisely odd crystal field harmonics remove restrictions on optical electric dipole transitions within the $f$ shell [17–20].

The concentration dependence of the $\delta_1$ and $\delta_2$ spectral lines allows them to be assigned to activator ion associates. At a comparatively low activator ion concentration, 1.8 at. %, these associates can be expected to be dimers. One or two dimer states can be observed in optical spectra depending on the mutual orientation of the dipole moments of ions in the dimer [21]. For $Pr^{3+}$ optical centers of the first and second types, the lower and higher dimer energy states, respectively, are observed. It follows that the dipole moments of interacting ion optical transitions in second-type centers have exactly opposite orientations [21]. Dimer energy states usually [21, 22] experience strong radiationless relaxation. For this reason, the luminescence spectra do not contain spectral lines whose frequencies coincide with those of $\delta_1$ and $\delta_2$ (Fig. 5).

The temperature dependence of the intensities of spectral lines $\delta_1$ and $\delta_2$ is evidence of a collective character of the excited state of dimers. Generally [21], the state of a dimer is described by a wave function of the form $\phi = a_1(t)\varphi_1 + a_2(t)\varphi_2$, where $\varphi_1$ and $\varphi_2$ are the wave functions describing the states of monomers, and $a_1(t)$ and $a_2(t)$ are complex functions of time. If the relative phase of $a_1(t)$ and $a_2(t)$ experiences stochastic disturbances at a frequency exceeding $V_{dd}/\hbar$ ($V_{dd}$ is the dipole-dipole interaction value in the dimer, and $\hbar$ is the Planck constant), the states of the dimer collapse. One of the reasons for $a_1(t)$ and $a_2(t)$ phase disturbances is scattering of phonons on admixture centers. Heating the crystal therefore decreases the intensity of the $\delta_1$ and $\delta_2$ spectral lines (Fig. 5). For a similar reason, dimer states are not formed if the initial states are subject to strong radiationless relaxation. This is characteristic of the $^1D_2$ term Stark components that lie higher in energy than the metastable state [18]. Dimer states are therefore not formed for the Stark components related to spectral lines 2 (2*), 3 (3*), 4 (4*), and 5 (5*).

The complex nonexponential luminescence decay law for YSO : $Pr^{3+}$ crystals with a high activator ion concentration (1.8 at. %) and the dependence of the shape of the luminescence decay curve on the concentration of activator ions and temperature (Fig. 6, 7) are consequences of simultaneous action of two relaxation mechanisms [23–26]. The first one operates by activator ion excitation energy transfer to acceptors [23, 24]. The second mechanism involves migration of electronic excitation energy [23, 24]. Activator ions ($Pr^{3+}$) donate electronic excitation energy to acceptor centers (the nature of these centers will be discussed below). Before the system luminesces a photon or there occurs energy transfer to an acceptor center, electronic excitation migrates over $Pr^{3+}$ ions of the same type. Migration and transfer of activator ion electronic excitation energy are caused by dipole-dipole interactions between the corresponding pairs [23–26]. If both mechanisms are operative, the luminescence decay curve for donors ($Pr^{3+}$ activator ions) is described by the dependence [24]

$$I_d(t) = I_0\exp\left(-\frac{t}{\tau_0} - \alpha\sqrt{\frac{t}{\tau_0}} - \beta\frac{t}{\tau_0}\right), \qquad (1)$$

$$\alpha = 7.4R_0^3c_a, \qquad (2)$$

$$\beta = 8.6R_0^{3/2}(D\tau_0)^{3/4}c_a, \qquad (3)$$

where $\tau_0$ is the luminescence decay constant for donors in the absence of acceptors, $R_0$ is the critical radius of electronic excitation energy transfer, $D$ is the diffusion coefficient of electronic excitation energy, and $c_a$ is the concentration of acceptors.

If $\alpha = 0$ and $\beta = 0$, (1) describes luminescence decay in samples with a low concentration of activator ions. The $\tau_0$ values determined for luminescence of the first and second type were given above. Correctly approximating the luminescence decay curves of the YSO : $Pr^{3+}$ crystals with medium and maximum activator ion concentrations by (1) requires the $\alpha$ and $\beta$ parameters to be assigned some physical meaning and their values to be specified.

Unfortunately, neither $c_a$ nor $R_0$ are known. A reasonable estimate of $R_0$ can, however, be obtained. Second-type luminescence decay in the YSO : $Pr^{3+}$ sample with an intermediate concentration of activator ions (0.6 at. %) is close to exponential and can be described by (1) with $\alpha = 0.14$ and $\beta = 0$. It follows that energy transfer to acceptors is in the nascent state, and migration is absent. First-type optical centers participate in electronic excitation energy migration. If some part of activator ions are acceptors (this will be shown below), the mean distance between second-type optical centers equals the critical radius of energy transfer. The mean distance between second-type optical centers can be estimated taking into account that the unit cell contains eight $Y_2SiO_5$ formula units and only 37% of the total concentration of activator ions are involved in the formation of second-type optical centers. It follows that the concentration of second-type optical centers (donors) equals $2.3 \times 10^{20}$ cm$^{-3}$, and the mean distance between them is 12 Å. Substituting $R_0 \sim 12$ Å and $\alpha = 0.14$ into (2) yields the concentration of acceptors, $c_a \sim 1.3 \times 10^{19}$ cm$^{-3}$. The concentration of acceptors is one order of magnitude lower than the concentration of donors, $c_d = 2.3 \times 10^{20}$ cm$^{-3}$. It follows that uncontrolled impurities cannot play the role of acceptors, because their concentration in the raw material used to activate YSO crystals is three orders of magnitude lower than the concentration of $Pr^{3+}$ ions. This raises the question of the nature of acceptors. The presence of the $\delta_1$ and $\delta_2$ lines in the absorption spectrum (Fig. 4) and their absence in the luminescence spectrum (Fig. 5) allows activator ion dimers to be treated as effective acceptors. Considering the topology of the YSO lattice and the concentration of $Pr^{3+}$ admixture ions equal to 0.6 at. %, we arrive at the conclusion that, among optical centers of the second type, there are about 1.9% ion pairs. Their concentration is $\sim 4.6 \times 10^{18}$ cm$^{-3}$, which is close to the calculated concentration of acceptors ($\sim 1.3 \times 10^{19}$ cm$^{-3}$).

The results obtained by approximating the kinetics of luminescence of the samples by (1) at various temperatures are summarized in Table 2. In (1), both $\alpha$ and $\beta$ parameters were varied. The $\alpha$ and $\beta$ parameter values and $R_0 = 12$ Å were used to calculate the concentration of acceptors and the diffusion coefficient of electronic excitation energy for each system and temperature. The $\alpha$ value and $R_0 = 12$ Å were used to determine $c_a$. The $D$ value was found from the concentration of acceptors and the $\beta$ parameter. For optical centers of the first type, the diffusion coefficients of electronic excitation energy were almost equal in samples with the highest and intermediate activator ion concentrations. The diffusion coefficients for first- and second-type optical centers were, however, different. As expected, the diffusion coefficient of electronic excitation energy decreased as temperature lowered. It follows that diffusion of electronic excitation energy was a thermally activated process [23, 24].

The areas under spectral lines 1 and 1* in the absorption spectra of YSO : $Pr^{3+}$ were different increasing functions of the concentration of activator ions (Fig. 4). This and the special features of luminescence decay for two types of optical centers led us to conclude that nonequivalent cation sites were nonuniformly populated by activator ions. Energy transfer between optical centers of two types in YSO : $Pr^{3+}$ did not occur in the temperature range 1.5–80 K. This finding requires special comments. Because the frequency gap between the metastable levels of two optical centers in YSO : $Pr^{3+}$ is about 66 cm$^{-1}$ (Fig. 1), energy transfer with the participation of phonons might well occur at 80 K [26]. Such energy transfer between rare-earth metal ions in solid-state matrices is observed under less favorable conditions [26].

The totality of the experimental data obtained in this work, their analysis, and spectral hole burning [13] and photon echo [14] experiments for YSO : $Pr^{3+}$ shed light on the microscopic nature of the large isomorphic capacity of oxiorthosilicates [1, 2, 4, 11]. Cation sites in oxiorthosilicate crystals are in a certain sense fairly "loose" because of high mobility of free oxygen atoms in coordination polyhedra. This makes it possible to optimize (from the point of view of free energy minimization) the geometry of activation complexes ("activator ion + oxygen polyhedron") in doping oxiorthosilicate crystals by various rare-earth metal ions.

## 5. CONCLUSION

An analysis of the low-temperature optical spectra of the YSO : $Pr^{3+}$, LSO : $Pr^{3+}$, and GSO : $Pr^{3+}$ crystals shows that the energy parameters and the character of field splitting of the $^1D_2$ and $^3H_4$ activator ion terms are substantially different for crystals of different crystallographic types. The pseudosymmetry effect is observed in YSO : $Pr^{3+}$ and LSO : $Pr^{3+}$ for the splitting of the terms of $Pr^{3+}$ ions situated in nonequivalent cation sites of the crystal lattices. Activator ions nonuniformly populate nonequivalent cation sites of the YSO crystal lattice. At high activator ion concentrations (above 1 at. %), activator ion luminescence decay in YSO : $Pr^{3+}$ is not described by a simple exponential time dependence. The complex decay law is caused by co-occurrence of two relaxation processes involving migration and capture by traps of activator ion excitation energy. Energy acceptors are activator ion dimers.

## REFERENCES

1. J. Felsche, Struct. Bonding (Berlin) **13**, 99 (1973).
2. C. L. Melcher, R. A. Manente, C. A. Peterson, and J. S. Schweitzer, J. Cryst. Growth **128**, 1001 (1993).
3. *Physics and Spectroscopy of Crystals*, Ed. by A. A. Kaminskiĭ (Nauka, Moscow, 1986).
4. B. A. Maksimov, Yu. A. Kharitonov, V. V. Ilyukhin, and N. V. Belov, Kristallografiya **15**, 926 (1970) [Sov. Phys. Crystallogr. **15**, 806 (1971)].
5. G. V. Anan'eva, A. M. Korovkin, T. I. Merkulyaeva, *et al.*, Izv. Akad. Nauk SSSR, Neorg. Mater. **17**, 1037 (1981).
6. C. L. Melcher and J. S. Schweitzer, IEEE Trans. Nucl. Sci. **39**, 502 (1992).
7. P. Dorenbos, C. W. E. van Eijk, A. J. J. Bos, and C. L. Melcher, J. Lumin. **60-61**, 979 (1994).
8. H. Suzuki, T. A. Tombrello, C. L. Melcher, and J. S. Schweitzer, Nucl. Instrum. Methods Phys. Res. A **320**, 263 (1992).
9. G. Huber, E. Heumann, T. Sandrock, and K. Peterman, J. Lumin. **72-74**, 1 (1997).
10. M. Malinowski, M. F. Joubert, and B. Jacquier, J. Lumin. **60-61**, 179 (1994).
11. B. A. Maksimov, Yu. A. Kharitonov, V. V. Ilyukhin, and N. V. Belov, Dokl. Akad. Nauk SSSR **183**, 1072 (1968) [Sov. Phys. Dokl. **13**, 982 (1968)].
12. Yu. V. Malyukin, N. V. Znamenskiĭ, É. A. Manykin, *et al.*, Fiz. Nizk. Temp. **24**, 571 (1998) [Low Temp. Phys. **24**, 432 (1998)].
13. K. Holliday, M. Croci, E. Vauthey, and U. P. Wild, Phys. Rev. B **47**, 14741 (1993).
14. Yu. V. Malyukin, E. A. Manykin, N. I. Znamensky, and E. A. Petrenko, Zh. Éksp. Teor. Fiz. **115**, 704 (1999) [JETP **88**, 385 (1999)].
15. D. V. O'Connor and D. Phillipe, *Time-correlated Single Photon Counting* (Academy, New York, 1984).
16. H. H. Caspers, H. E. Rast, and R. A. Buchanan, J. Chem. Phys. **43**, 2124 (1965).
17. N. A. Kulagin and D. T. Sviridov, *Computing Methods of Electronic Structures of Free and Impure Ions* (Nauka, Moscow, 1986).
18. *Spectroscopy of Solids Containing Rare Earth Ions*, Ed. by A. A. Kaplyanskii and R. M. Macfarlane (North-Holland, Amsterdam, 1987).
19. I. B. Bersuker, *Electron Structure and Properties of Coordination Compounds* (Khimiya, Leningrad, 1986).
20. A. M. Stouneham, *Theory of Defects in Solids: the Electronic Structure of Defects in Insulators and Semiconductors* (Clarendon Press, Oxford, 1975; Mir, Moscow, 1979).
21. R. Hochstrasser, *Molecular Aspects of Symmetry* (Benjamin, New York, 1966; Mir, Moscow, 1969).
22. *Spectroscopy and Excitation Dynamics in Condensed Molecular Systems*, Ed. by V. M. Agranovich and R. M. Hochstrasser (North-Holland, Amsterdam, 1983; Nauka, Moscow, 1987).
23. M. J. Weber, Phys. Rev. B **4**, 2932 (1971).
24. V. M. Agranovich and M. D. Galanin, *Electronic Excitation Energy Transfer in Condensed Matter* (Nauka, Moscow, 1978; North-Holland, Amsterdam, 1982).
25. M. Nikl, J. A. Mares, E. Minokova, and K. Blazek, J. Lumin. **60-61**, 971 (1994).
26. X. X. Zhang, P. Hong, M. Bass, *et al.*, J. Lumin. **60-61**, 878 (1994).

*Translated by V. Sipachev*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 380–392.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 430–444.*
*Original Russian Text Copyright © 2001 by Belyakov, Kats.*

**SOLIDS**
**Structure**

# Surface Anchoring and Pitch Variation in Thin Smectic *C** Layers in an Electric Field

**V. A. Belyakov[a, *] and E. I. Kats[a, b]**

[a]*Landau Institute of Theoretical Physics, Russian Academy of Sciences,*
*Chernogolovka, Moscow oblast, 142432 Russia*
[b]*Institute Laue–Langevin, F-38042, Grenoble, France*
**e-mail: bel@cpd.landau.ac.ru*

Received April 9, 2001

**Abstract**—The variations of the pitch of smectics *C** in thin planar layers in an external electric field and their dependence on the surface anchoring are investigated theoretically. The proposed mechanism of the change in the number of half-turns of the helical structure in a finite-thickness layer upon a change in the applied field is the slip of the director on the surface of the layer through the potential barrier of surface anchoring. The equations describing the pitch variation in an external field and, in particular, the hysteresis in the jumpwise variations of the pitch for opposite directions of field variation are given and analyzed for arbitrary values of the field. For weak fields, it is found that the pitch variation in the layer is of a universal nature and is determined by only one dimensionless parameter, $S_d = K_{22}/dW$, where $K_{22}$ is the Frank torsion modulus, $W$ is the surface anchoring potential, and $d$ is the layer thickness. The possibility of direct determination of the form of the anchoring potential from the results of corresponding measurements is considered. Numerical calculations for the deviation of the director from the direction of alignment on the layer surface and pitch variations, as well as the points of pitch jumps and hysteresis in the field, are made for the Rapini model anchoring potential for values of the parameters for which the pitch variation weakly depends on the direction of the field applied in the plane perpendicular to the spiral axis of smectics *C**. The changes in the pitch variation in stronger fields are discussed, and the optimal conditions for observing the discovered effects are formulated. © *2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

Interest in the detailed analysis of smectics *C** possessing unique properties, since they are not only characterized by local anisotropy of dielectric parameters but also exhibit spontaneous electric polarization, has increased considerably in recent years (see, for example, [1]). This interest is partially explained by general physical factors (these crystals are characterized by a large number of various nontrivial phases; see, for example, monograph [2]). Besides, numerous additional advantages of smectics *C** over traditional nematic crystals have not been realized as yet in applications. Since smectics *C** are widely used in developing information display systems, controllable optical transparencies, and for many other purposes, an analysis of their properties in confined geometries is extremely vital. The study of the optical parameters of thin layers of smectics *C**, including the influence of external agencies on them, provides information on the change in their structure in thin layers and on the dynamics of these changes, which is important for applications as well as for understanding the physics of liquid crystals.

It is well known that in the presence of surface anchoring forces in thin planar layers of smectics *C**, jumpwise pitch variations take place upon a continuous variation of an external action [1, 3]. Such jumpwise changes were studied most comprehensively for the pitch of cholesteric liquid crystals upon the variation of temperature [4, 5]. In particular, these changes are manifested in precise measurements in linear and nonlinear optics of liquid crystals [6, 7]. Moreover, jumpwise variation of the pitch of helical structures of smectics *C** display a temperature hysteresis [4, 5]. This leads to bistability of liquid crystals, which is important from the viewpoint of numerous applications of liquid crystals.

Phenomena similar to the above-mentioned temperature hysteresis must exist in an external field applied to thin layers of smectics *C** due to adhesive forces acting on their surfaces (see, for example, [8,9]). However, the corresponding changes in the parameters of these crystals (especially the hysteresis of these changes) in external electric or magnetic fields have been studied less comprehensively, although the jumps and hysteresis in the variation of parameters in an external field are important from the physical point of view.

The present work is devoted to an analysis of pitch variations in thin layers of smectics *C** in an external electric field. This problem is more complicated than the problem of temperature variations of the pitch in thin layers with surface anchoring since the field induces not only a change in the spiral pitch in these crystals, but also distortions of the spiral itself (it becomes unharmonic). In particular, the so-called frustrations take place in smectics *C** due to a competition between enantiomorphism facilitating a nonuniform helical ordering of molecules and the action of an external field striving to align the molecules along the field. The compromise between these opposite tendencies (and, hence, elimination of frustrations) may be achieved in the system due to the formation of domain walls or solitons separating the regions with a practically uniform orientation. For example, an ideal spiral of the director field is transformed into a lattice of $2\pi$ solitons in a magnetic field (or in an electric field in the case of cholesteric liquid crystals which display no spontaneous polarization). This means that the regions of uniform orientation of the director along the field become separated by domain walls in which the orientation changes by $2\pi$, while in smectics *C** in an electric field, a $\pi$-soliton lattice is formed.

In this work, we determine the conditions under which the simple continual theory of elasticity taking into account surface anchoring is applicable for describing the structural variations of smectics *C** in layers in an external field. For these conditions, a general analysis of pitch variation (including pitch jumps) in a field, as well as of the hysteresis of these variations, is carried out. It is shown that in weak fields, pitch variations exhibit a universal behavior in a layer upon the application of an electric field, which is similar to temperature variations of the pitch in smectics *C** [10]. Numerical calculations of pitch variation and of the hysteresis of pitch jumps in smectics *C** in an electric field are made for the case of strong surface anchoring and the effects accessible for experimental observation are outlined.

## 2. BASIC EQUATIONS

Let us consider the behavior of a spiral pitch in a thin planar layer of smectic *C** in an electric field perpendicular to the spiral axis. A similar problem for a bulk cholesteric was considered for the first time in [11, 12], for a bulk smectic *C** in [13], and for a cholesteric layer with the spiral axis perpendicular and parallel to its surface in [14] and [15], respectively. We will assume that the forces of surface anchoring acting on the two surfaces of the layer are identical and that the directions of the axes of the director alignment on both surfaces are preset (and generally different). The problem of the spiral unwinding in a layer of a smectic *C** in the presence of surface anchoring forces (in particular, the jumps of the spiral pitch) differs qualitatively from the corresponding problem for a bulk crystal. In the case of confined geometry with surface anchoring, simple equilibrium models of liquid crystals do not always provide a correct description of the behavior of a smectic *C** under the action of an external factor. In particular, the points at which the parameters of liquid crystals in a layer experience jumpwise variation determined using such models may not reflect the actual course of the process. A correct description of director field variations in such cases requires an analysis of the liquid crystal dynamics. A simple illustration of this fact is the behavior of a spiral in a layer of a smectic *C** with an infinitely strong surface anchoring in an external field. In the simple continual theory taking the surface anchoring into account, the variations of an external field cannot change the number of turns of a helix in the layer; they simply deform the helix. In actual practice, however, the number of helix turns changes under the action of an external factor (field, temperature, etc.) even in the case of a very strong anchoring and the system passes to an equilibrium state corresponding to the intensity of the external action. In the models taking into account the dynamics of liquid crystals, such transitions are possible due to fluctuations in the orientation of the director. The example with and infinitely strong surface anchoring describes an extreme situation in which the simple theory is completely inapplicable. In the general case of a finite force of surface anchoring, fluctuations displace (on the scale of the external effect) the points of jumps in the parameters of a liquid crystal layer, determined by using the continual approach, towards a decrease in the observed hysteresis in a change in the parameters of the layer. In some cases, it is fluctuations that determine the jump mechanism, while in other cases the fluctuations lead to the abovementioned displacement of the jump points determined using the continual approach. The specific mechanism of a jump in the given case depends on the parameters of the problem of liquid crystals in a layer. We will assume in the subsequent analysis that the jump mechanism is associated with overcoming the potential barrier of surface anchoring forces by the director on the surface; for this reason, the problem will be solved for the corresponding range of liquid crystal parameters in a layer (see [10]).

It should be noted that in contrast to [11, 12, 14], where spontaneous ferroelectric polarization in liquid crystals was absent, we do not assume that the spontaneous polarization in smectics *C** is equal to zero. We write the expression for the free energy in a layer placed in an electric field in the form

$$F(\mathbf{E}) = F_S + \int F_V(\mathbf{E})dV, \tag{1}$$

where $F_V(\mathbf{E})$ is the volume density of free energy of smectics *C** in an external electric field $E$, $F_S$ is the surface free energy, and integration is carried out over the layer volume.

The surface free energy can be expressed in terms of the surface anchoring potential and in the simplest case has the form

$$F_S = \sum W_S(\varphi_{iS}), \tag{2}$$

where summation is carried out over both surfaces of the layer, $W_S(\varphi)$ is the surface anchoring potential, and $\varphi_{iS}$ is the angle of deviation of the director on the *i*th surface of the layer from the alignment direction.

It was mentioned above that smectics $C^*$ are unique systems with spontaneous electric polarization; for this reason, it is necessary to take into account in the general case both the linear (in an electric field) contribution to the free energy, associated with nonzero spontaneous polarization, and the quadratic contribution associated with anisotropy of permittivity. Moreover, it is well known [3, 16, 17] that any deformation of the director field in liquid crystals may lead to the emergence of electric dipole polarization (the so-called flexoelectric effect). In the case of strong deformations of the director field, the corresponding contribution to the free energy of liquid crystals in an electric field may become comparable to the above-mentioned dielectric and ferroelectric contributions.

Surface anchoring, as well as an external field, leads to frustrations; however, a compromise between the opposite tendencies in this case may be reached more easily since enantiomorphism is significant in the entire volume of the liquid crystal, while surface anchoring acts in a narrow surface layer, $\xi \approx K/W$, where $K$ is the characteristic elastic modulus and $W$ is the characteristic anchoring potential.

In the general case, $W$ is a function of the polar and azimuthal angles formed by the director with the normal to the surface, $W(\theta, \varphi)$, and the problem of minimization of the total free energy becomes very complicated. Since we aim mainly at determining qualitative dependences, we assume that $W(\theta, \varphi)$ can be factorized ($W(\theta, \varphi) = W_1(\theta)W_2(\varphi)$) so that $W_1(\theta) = W_1(\theta - \theta_0)$, where $\theta_0$ is the angle of tilt of the director, and the potential $W_1(\theta)$ is such that a very strong anchoring in the polar angle $(\theta - \theta_0)$ takes place. As regards $W_2(\varphi)$, the anchoring energy for this function can be arbitrary.

We do not assume that such a simplified model of surface energy is necessarily applicable to all known smectics $C^*$. Our aim is much more modest: to consider the consequences of our simplified model and to compare them with experimental data. It should also be noted that in the case of cholesteric liquid crystals, our model anchoring potential (with $\theta_0 = 0$) is quite justified from the physical point of view.

Let us first consider cholesteric liquid crystals with zero spontaneous polarization. In this case, the volume component of the free energy density $F_V(\mathbf{E})$ has the familiar form

$$\begin{aligned} F_V(\mathbf{E}) = \frac{1}{2}[&K_{11}(\operatorname{div}\mathbf{n})^2 + K_{22}(\mathbf{n}\cdot\operatorname{curl}\mathbf{n} - q_0)^2 \\ &+ K_{33}[\mathbf{n}\times\operatorname{curl}\mathbf{n}]^2] + \frac{\varepsilon_a(\mathbf{n}\cdot\mathbf{E})^2}{8\pi} \\ &+ e_1\mathbf{E}\cdot[\mathbf{n}\times\operatorname{curl}\mathbf{n}] + e_3\mathbf{n}\operatorname{div}\mathbf{n}, \end{aligned} \quad (3)$$

where $K_{ii}$ are elastic moduli, $q_0 = 2\pi/p_0$, $p_0$ being the pitch of the cholesteric spiral, which has an equilibrium value for a bulk liquid crystal, and $e_1$ and $e_3$ are the flexoelectric coefficients.

Smectics $C^*$ have the spontaneous polarization $\mathbf{P}_s$, whose direction is unambiguously determined by symmetry considerations: $\mathbf{P}_s = P_s\mathbf{n}\times\mathbf{l}$, where $\mathbf{l}$ is a unit vector perpendicular to smectic layers.

Thus, expression (3) for the free energy of these crystals must be supplemented with the additional term $\mathbf{P}_s\cdot\mathbf{E}$. Taking into account the symmetry properties of smectic $C^*$, we can write its director in the form $\mathbf{n} = \mathbf{l}\cos\theta + \mathbf{c}\sin\theta$, where $\theta$ is the angle of tilt of molecules in smectic layers, which is determined by intermolecular forces, and $\mathbf{c}$ is the so-called $\mathbf{c}$ director which singles out a certain direction in the plane of smectic layers.

The elastic energy of a smectic $C^*$ also contains the contribution associated with the compression of smectic layers, which, however, is not affected by an external field and hence can be disregarded. The orientation energy component important for our analysis can be presented in the form of an expansion in the gradients of $\mathbf{l}$ and $\mathbf{c}$ (it is also convenient to introduce the third vector $\mathbf{p} = \mathbf{c}\times\mathbf{l}$ to have the unit vectors of the local system of coordinates).

The general expression for the free energy of a smectic $C^*$ is very cumbersome in view of its low symmetry (e.g., it contains 14 flexoelectric terms alone!). Considering that the values of phenomenological constants appearing in the relevant terms are not known exactly, we will consider the simplest situation, in which torsion is the only admissible deformation, which leads to the following expression for the volume density of free energy [13, 16] (in the case of cholesteric liquid crystals, we must put $\mathbf{P}_s = 0$):

$$\begin{aligned} F_V(\mathbf{E}) &= \frac{K_{22}d}{2}\left(\frac{d\varphi}{dz} - \frac{2\pi}{p_0}\right)^2 \\ &+ EP_s\cos\varphi + \frac{\varepsilon_a E^2}{16\pi}\cos 2\varphi, \end{aligned} \quad (4)$$

where $K_{22}$ is the elastic torsion modulus, $p_0$ is the equilibrium value of the spiral pitch in a bulk smectic $C^*$ in zero field, $\varphi$ is the azimuthal angle of molecular orientation measured from the direction of the applied field, $\varepsilon_a$ is the dielectric anisotropy of the liquid crystal, $\mathbf{P}_s$ is its spontaneous polarization, and $d$ is the layer thickness.

The same type of expression for free energy in the general case also corresponds to the boundary regions of the layer. In order to derive this expression, we must write all possible invariants constructed on the basis of the three vectors $\mathbf{p}$, $\mathbf{c}$, and $\mathbf{l}$ for both surfaces of the layer. Besides, in the general case, we must take into account the fact that the angle of tilt of molecules on the surface might differ from its value in the bulk of the layer, and the order parameter does not remain constant over the layer thickness [18]. Moreover, a domain structure may be formed on the surface of a layer of ferroelectric liquid crystals. The physical mechanism responsible for this domain structure is associated with the dipole interaction. Besides, owing to the dipole interaction, the polarization can modify the spatial distribution of the order parameter. This interaction must be added to the surface and volume energies by introducing the so-called depolarization field $\mathbf{E}_d$ [19] in the form of the term $-\mathbf{E}_d(\mathbf{P}_s + \mathbf{P}_f)/2$, where $\mathbf{P}_s$ and $\mathbf{P}_f$ are the spontaneous and flexoelectric polarizations, respectively.

For a thin layer having a thickness $d$ and being unbounded on the surface, this field differs from zero only if the direction of the total polarization $\mathbf{P} = \mathbf{P}_s + \mathbf{P}_f$ is not parallel to the surface of the layer. Introducing the angle $\chi$ between the normal to the surface and the polarization vector $\mathbf{P}$, we can write the expression for the depolarization field in the form [19]

$$\mathbf{E}_d(z) = -4\pi\varepsilon_0\cos^2\chi\left[\mathbf{P}(z) - \frac{l}{d}\int\mathbf{P}(z)dz\right], \tag{5}$$

where $\varepsilon_0$ is the isotropic permittivity component and the integration is carried out over the layer thickness.

We may consider in principle the behavior of a smectic $C^*$ in a field, taking into account all the above-mentioned contributions to the free energy of the layer. However, a comparison with the experimental data requires the knowledge of many phenomenological parameters, most of which are unfortunately unknown. For this reason, we will not minimize the free energy of the layer in its most general form in the presence of unknown parameters since the theory is divested of its predictability in this case. Instead, we will be using the "principle of minimum requirements"; i.e., we will carry out our analysis using the simplest (but not trivial) model, which ensures experimentally verifiable predictions under relatively weak limitations on generality. This means that we will use an expression for free energy in the form (4), assuming that the direction of the polarization vector $\mathbf{P}$ is rigidly connected with director $\mathbf{n}$ and disregarding the dipole interaction. In this case, the surface anchoring energy assumes the standard form (2) and is a function of the azimuthal angle $\varphi_S$ between the director on the surface and the alignment direction only. The equation describing the spatial variations of the director orientation in the field, i.e., $\varphi(z)$, where the $z$ axis determines the direction of the spiral axis and is perpendicular to the layer surface, can be obtained in the conventional way [11–13] from the condition for the minimum of the free energy (1).

The volume component of free energy given by the integral in relation (1), which will be denoted by $F_{\mathrm{IV}}(\mathbf{E}, N, \varphi_{1S}, \varphi_{2S})$, where $N$ is the number of half-turns of the director over a layer of a finite thickness and $\varphi_{1S}$ and $\varphi_{2S}$ are the angles formed by the director with the direction of alignment on the layer surfaces, turns out to be a function of not only magnitude of the applied field $\mathbf{E}$, but also the mutual orientation of the director on the layer surfaces and the direction of the applied field. The procedure of minimization of the free energy (1) in this case is reduced to determining the minimum value of the sum of the surface energy and the functional $F_{\mathrm{IV}}(\mathbf{E}, N, \varphi_{1S}, \varphi_{2S})$ as a function of $\varphi_{1S}$ and $\varphi_{2S}$. It should be noted that the minimum of functional $F_{\mathrm{IV}}(\mathbf{E}, N, \varphi_{1S}, \varphi_{2S})$ for fixed values of $\varphi_{1S}$ and $\varphi_{2S}$ is realized for a certain function $\varphi(z)$, describing the director distribution over the layer thickness and corresponding to the solution of the initial problem for an infinitely strong surface anchoring with preset direction of alignment ($\varphi_{1S}$ and $\varphi_{2S}$) on the surfaces. This means that the equilibrium configuration of the spiral in the field is defined by the equations

$$\begin{aligned}\frac{\partial}{\partial\varphi_{1S}}(F_S + F_{\mathrm{IV}}(\mathbf{E}, N, \varphi_{1S}, \varphi_{2S})) &= 0,\\ \frac{\partial}{\partial\varphi_{2S}}(F_S + F_{\mathrm{IV}}(\mathbf{E}, N, \varphi_{1S}, \varphi_{2S})) &= 0.\end{aligned} \tag{6}$$

The equilibrium value of free energy $F(\mathbf{E})$ for a fixed value of field $\mathbf{E}$ can be found by substituting the function $\varphi(z)$ as well as the quantities $\varphi_{1S}$ and $\varphi_{2S}$ determined from Eqs. (6) into expression (1). The system of equations defining the equilibrium values of $\varphi_{1S}$ and $\varphi_{2S}$ for a fixed value of field $\mathbf{E}$ is given by equations containing the surface anchoring potential $W_S(\varphi_i)$, the volume density of free energy $F_V(\mathbf{E})$, and the function $\varphi(z)$ defining the equilibrium distribution of the director in the layer for the field $\mathbf{E}$:

$$\begin{aligned}\left(\frac{dW_S(\varphi)}{d\varphi}\right)_{\varphi = \varphi_{1S}} + \left(F_V(\mathbf{E})\frac{dz}{d\varphi}\right)_{\varphi = \varphi_{1S}} &= 0,\\ \left(\frac{dW_S(\varphi)}{d\varphi}\right)_{\varphi = \varphi_{2S}} + \left(F_V(\mathbf{E})\frac{dz}{d\varphi}\right)_{\varphi = \varphi_{2S}} &= 0.\end{aligned} \tag{7}$$

Equations (6) and (7) solve, in the general form, the problem of determining the effect of the electric field on the director distribution in the layer of a smectic $C^*$ of finite thickness in the presence of surface anchoring. It can be seen from system (7) that the decisive factor in the solution of the above-formulated problem is the distributions of the director in the layer and of the angles of deviation of the director from the direction of alignment on the layer surfaces.

## 3. INFINITELY STRONG SURFACE ANCHORING

It follows from the previous section that the solution of the problem of field distortion of the director distribution in a layer of a smectic $C^*$ of finite thickness for infinitely strong anchoring at the surface is a stage in the solution of the corresponding problem in the case of finite surface anchoring forces. In this case, we must retain only the volume term in formula (1) for free energy, assuming that the orientations of the director on both surfaces of the layer are preset, i.e., assuming that angles $\varphi_{1S}$ and $\varphi_{2S}$ are known. Then the conditions for the minimum of free energy (1) are determined by the same Euler equations as for a bulk smectic $C^*$ [13]:

$$K_{22}\left(\frac{d\varphi}{dz}\right)^2 + EP\sin\varphi + \frac{\varepsilon_a E^2}{8\pi}\sin 2\varphi = 0. \qquad (8)$$

The first integral in this equation is the same as for a bulk smectic $C^*$:

$$\left(\frac{d\varphi}{dz}\right)^2 - \frac{2EP}{K_{22}}\cos\varphi - \frac{\varepsilon_a E^2}{4\pi K_{22}}\cos^2 2\varphi = C_1, \qquad (9)$$

where the constant $C_1$ has to be determined.

Thus, the equation for the function $\varphi(z)$ assumes the form

$$\frac{d\varphi}{dz} = \pm\left[C_1 + \frac{2EP}{K_{22}}\cos\varphi + \frac{\varepsilon_a E^2}{4\pi K_{22}}\cos^2 2\varphi\right]^{1/2}. \qquad (10)$$

Integration of this equation defines the function $\varphi(z)$ in an implicit form with the help of the following relation:

$$z = \pm\int\left[C_1 + \frac{2EP}{K_{22}}\cos\varphi + \frac{\varepsilon_a E^2}{4\pi K_{22}}\cos^2 2\varphi\right]^{-1/2} d\varphi + C_2. \qquad (11)$$

Constants $C_1$ and $C_2$ in this relation are determined by the equilibrium distribution of the director in a layer of the liquid crystal for zero value of field $\mathbf{E}$:

$$C_1 = \frac{(\varphi_{1S} - \varphi_{2S})^2}{d^2}, \quad C_2 = \frac{d}{2(\varphi_{1S} - \varphi_{2S})^2}, \qquad (12)$$

where $z = 0$ corresponds to the middle of the layer; for this reason, angles $\varphi_{1S}$ and $\varphi_{2S}$ are measured not from the direction of preferred orientation on the surface, but from point $z = 0$ and, hence, have opposite signs. They also contain an angle incursion associated with the rotation of the director in the bulk of the layer upon a change in the coordinate along the $z$ axis from its center to the surface. Consequently, if, for example, an integral number of half-pitches fit into the layer thickness in zero field, and the direction of alignment and surface anchoring are identical for both surfaces, we obtain $C_1 = (2\pi/p_0)^2$ and $C_2 = 0$ (the pitch in the layer in zero field in this case coincides with the pitch $p_0$ in a bulk liquid crystal in zero field).

The function $\varphi(z, E)$ defined unambiguously in this way describes the distortion of the structure of the director in the layer upon field variations in the case when the orientations of the director on the layer surfaces are independent of the field and also determines the changes in the spiral deformation and the volume component of free energy upon a change in the values of $\varphi_{1S}$ and $\varphi_{2S}$. It is appropriate to note here that for an infinitely strong surface anchoring, the number $N$ of half-turns of the spiral in the layer thickness does not depend on the field and coincides with the initial number of half-turns in zero field. The action of the field in this case is reduced only to the deformation of a helix in the layer.

In order to solve the problem of the action of the field on the director structure in a layer of a liquid crystal for finite surface anchoring forces, we must substitute the obtained functions of the field $\varphi(z, \mathbf{E}, \varphi_{1S}, \varphi_{2S})$, containing $\varphi_{1S}$ and $\varphi_{2S}$ as parameters, into system (7), which determines the angles $\varphi_{1S}$ and $\varphi_{2S}$ of deviation of the director on the surface, which have equilibrium values for each value of field $\mathbf{E}$, and the deformation of the helix in the bulk of the layer corresponding to these angles. For finite anchoring forces, the number of half-turns of a helix fitting into the thickness of a layer may change under the action of the field (the helix is not only deformed but also uncoiled). The above remark concerning the independence of the number of helical turns on the applied field for an infinitely strong surface anchoring refers to an idealized model of a smectic $C^*$, which rules out the formation of defects in the field of the director. In actual experiments, the number of helical pitches fitting into the layer thickness may also change with the applied field, but this occurs due to the formation of a defect layer in the field of the director in the bulk, followed by the relaxation of the director field to its equilibrium configuration.

## 4. PITCH JUMPS INDUCED BY FIELD VARIATIONS

A general analysis of Eqs. (1)–(12) shows that as in the case of temperature-induced variations of the pitch in a layer of a smectic $C^*$ with surface anchoring [10], the change in the helical pitch induced by field variations exhibits, along with smooth variations as a function of the field, jumps at certain values of the field (see also [14]). At the points corresponding to a certain value of the field, the number of turns in the helix of a smectic $C^*$ in a finite-thickness layer experiences a jump. It was found that the points corresponding to the jumps do not coincide for opposite directions of field variation. This means that we are dealing with a hysteresis similar to that observed long ago in the confocal texture of a smectic $C^*$ [8].

The above formulas describe the pitch jump in the mechanism of director slip over the surface through the surface anchoring potential barrier. However, it was observed in [4, 5] and confirmed theoretically in [10] that the jump mechanism, i.e., the helical configurations through which a transition between states differing in the number of turns over the layer thickness occurs in actual experiments in a field, depend on the parameter $S_d = K_{22}/dW$. For values of this parameter larger than the critical value $S_{dc} = 1/2\pi$, the mechanism of transition is associated with the slip of the director through the surface anchoring potential barrier in the layer, and the above formulas describe the real jump process. For values of $S_d$ smaller than $S_{dc}$, the form of the jump is studied insufficiently and may be associated with the dynamics of the liquid crystal and with the formation of defects in the planar texture of a smectic $C^*$ (these defects have already been mentioned in connection with infinitely strong anchoring, for which $S_d = 0$), followed by their relaxation. For this reason, there exists a limitation (in parameter $S_d$) on the applicability of the formulas derived above to real experiments. We will confine our subsequent analysis of the behavior of a layer of a smectic $C^*$ in an external field to the values of parameter $S_d$ exceeding $S_{dc}$, for which the mechanism of director slip through the barrier of surface anchoring forces “operates.”

Since the jump (and hysteresis) phenomena are consequences of metastable states existing in the system, it would be interesting to analyze the height of the barrier between these states and its dependence on the field strength. In analogy with the case of temperature hysteresis [10], the expression for the barrier height between helical configurations differing by a half-turn can be presented in the form

$$B(\mathbf{E}, S_d) = F_c(\mathbf{E}, \varphi_c, S_d) - F(\mathbf{E}, \varphi_S(\mathbf{E}), S_d), \qquad (13)$$

where $\varphi_c$ is a certain critical angle of deviation of the director on the surface, beyond which the helical configuration in the layer changes jumpwise. (Here and below, we consider for simplicity the symmetric problem under the assumption that the alignment directions on both surfaces coincide; in this problem, $\varphi_{1S} = -\varphi_{2S} = \varphi_S$ and the orientation of the director on the surfaces of the layer is defined by the same parameter $\varphi_S$.) The equilibrium value of the angle $\varphi_S(E)$ of deviation of the director on the surface is determined by the solution of Eqs. (7), while the free energy $F(\mathbf{E}, \varphi_S(\mathbf{E}), S_d)$, which has the equilibrium value for a given field $\mathbf{E}$, is defined by formula (1) after the substitution of the director distribution over the cell thickness determined from Eqs. (6) and (7). The free energy $F_c(\mathbf{E}, \varphi_c, S_d)$ is also determined by formula (1) upon the substitution of the director distribution over the cell thickness determined from Eqs. (6) and (7) for field $\mathbf{E}$ under the condition that the angle of deviation of the director on the surface coincides with the critical value $\varphi_c$. In other words, the free energy $F_c(\mathbf{E}, \varphi_c, S_d)$ contains the volume component $F_{\mathrm{IV}}(\mathbf{E}, N, \varphi_S)$ corresponding to an infinitely strong surface anchoring for the orientation of the director corresponding to the critical angle. As expected, as the angle $\varphi_S(\mathbf{E})$ attains the critical value, the barrier height vanishes and a jumpwise transition between configurations differing in the number of helical turns by a half-turn occurs in the layer. In the case of temperature-induced variations of the pitch, the critical angle is determined only by the shape of the surface anchoring potential [10]. It is clear from general considerations that in the presence of pitch jumps in the field, the value of the critical angle depends not only on the form of the surface potential, but also on the magnitude and direction of the field (if it is strong).

The value of the pitch $p_d(\mathbf{E}_j)$ in a layer immediately before its jump can also be expressed in terms of the critical angle $\varphi_c$ whose value is now determined by the form of the anchoring potential and by the applied field. Naturally, the value of the field $\mathbf{E}_j$ for which a jump takes place can be used for determining the angle $\varphi_0(\mathbf{E}_j)$ of free rotation of the director in the absence of surface anchoring as well as the value of the helix pitch $p(\mathbf{E}_j)$ for this field $\mathbf{E}_j$ in a bulk smectic $C^*$ [13].

It was mentioned above that the critical angle in the case of temperature-induced pitch variations is the same for all values of the parameters of the layer (probably except for the parameter specifying the form of the surface anchoring potential), while the calculation of the critical angle for a nonzero field is a separate problem. It can only be stated beforehand that the value of $\varphi_c$ for the symmetric situation under investigation does not exceed $\pi/2$. As in the case of temperature-induced variations of the pitch [10], the condition for the attainment of the critical value in the system is the vanishing of the height of the barrier between helical configurations in the layer differing by half-turn of the helix, which is equivalent to the absence of the system resistance at this point to the rotation of the entire configuration of the helix in a layer as a whole. This means that the derivative of the equilibrium free energy with respect to $\varphi_{1S}$ or $\varphi_{2S}$ vanishes for a fixed difference $\varphi_{1S} - \varphi_{2S}$ for a certain value of field $\mathbf{E}_j$. Consequently, the corresponding field $\mathbf{E}_j$ and the value of $\varphi_c$ can be determined from the condition

$$\delta(F_S + F_{1V}(\mathbf{E}, N, \varphi_{1S}, \varphi_{2S})) = 0 \qquad (14)$$

under the additional requirement that the variation is taken at the point $\varphi_{1S} = -\varphi_{2S}$ for $\varphi_{1S} - \varphi_{2S} = \text{const} = 2\varphi_c$ and that $\varphi_S$ is determined by function $\varphi(\mathbf{E})_S$.

## 5. STRONG SURFACE ANCHORING

Generally speaking, the value of free energy (1) depends not only on the magnitude of the applied field, but also on the mutual orientation of the alignment directions on the surfaces of the layer and the direction of the applied field. However, this dependence may be very weak and can be neglected for a strong surface anchoring and a large number of turns in the layer thickness $d$. In order to demonstrate this, we divide the integral in expression (1) into two regions: the internal region containing an integral number of turns in the helix and the surface region containing less than one helix turn. The free energy of the internal part of the layer attains its minimum value for the helical configuration realized in the corresponding problem for a bulk smectic $C^*$ [13] and does not depend on the direction of the field in the plane perpendicular to the helix axis. If the volume free energy of the second (surface) region is much smaller than the surface free energy $F_S$, it can be disregarded. The possibility of such disregard is determined by the condition

$$\frac{W}{p_0} > EP + \frac{\varepsilon_a E^2}{16\pi}. \tag{15}$$

Naturally, in the case of very weak fields, condition (15) holds in all cases. In order to estimate the upper field limit for the fulfillment of condition (15), we can put $E = \alpha E_c$, where $E_c$ is the critical field of unwinding of the helix, which is determined by the parameters of the smectic $C^*$ [3, 13, 16], and $\alpha$ is a certain numerical coefficient determined by condition (15). It should be emphasized that when condition (15) is satisfied, the deviation of the director from the preferred orientation direction on the surface becomes universal and independent of the details of the director field configuration in the bulk of the layer.

It should also be noted that we presume the homogeneous state of the liquid crystal in the layer. For this reason, domain walls in the layer of the liquid crystals are disregarded in our analysis. In the approximation formulated above, the angle of deviation of the director from the direction of the alignment as a function of the field (under the assumption that the orientation is the same on both surfaces of the layer) can be calculated on the basis of relations similar to the equations for the temperature-induced variations of the director orientation [10]:

$$\frac{\partial W_S(\varphi)}{\partial \varphi} + \frac{2K_{22}}{d}(\varphi - \varphi_0(p(E))) = 0, \tag{16}$$

where we have omitted the subscript $S$ on the angle of deviation of the director on the surface, $p(E)$ is the pitch in the field $E$ in the bulk smectic $C^*$ [13], and $\varphi_0(p(E))$ is the angle of deviation of the director on the surface from the alignment direction in the absence of anchoring for a harmonic helix with pitch $p(E)$.

The values of the pitch $p_d(E_j)$ in a layer just before its jump can also be expressed in terms of the critical angle $\varphi_c$, whose value under the assumptions made here are determined quite exactly by the form of the anchoring potential (naturally, the angle $\varphi_0(E_j)$ of free rotation of the director; i.e., the helix pitch $p(E_j)$ in a bulk smectic $C^*$ for the jump field $E_j$, can also be determined from the value of $E_j$):

$$p_d(E_j) = \frac{2d}{N + 2\varphi_c/\pi}, \quad p(E_j) = \frac{2d}{N + 2\varphi_0(E_j)/\pi}, \tag{17}$$

where $N$ is the number of half-turns of the helix in a layer of thickness $d$ in the initial equilibrium configuration of the helix in zero field, $E_j$ is the value of the field corresponding to the pitch jump, and the angle of free rotation of the helix for this value of the jump field is given by

$$\varphi_0(E_j) = \varphi_c + \left(\frac{\partial W_S(\varphi)/\partial\varphi}{2WS_d}\right)_{\varphi = \varphi_c}. \tag{18}$$

Thus, if we know the shape of the surface anchoring potential and the behavior of the helix pitch in a bulk smectic $C^*$ as a function of the applied field, the above relations determine the behavior of the pitch in a finite-thickness layer as a function of the applied field provided that the field is weak and satisfies the inequality (15). The same formulas can also be used for solving the inverse problem, namely, the reconstruction of the surface anchoring potential from the experimentally determined field dependence of the helix pitch in a finite-thickness layer.

## 6. FIELD-INDUCED HYSTERESIS OF PITCH VARIATIONS IN A LAYER

In order to describe field-induced variations of the pitch in a layer, we must specify the shape of the surface anchoring potential $W_S(\varphi)$ [20, 21]. However, even without specifying the shape of this potential, we can qualitatively determine the variations of the pitch in a layer of a smectic $C^*$ as a function of the applied field. In order to analyze the field dependence of the pitch in the layer and, in particular, to determine the values of the field at the points of pitch jump corresponding to the change in the number of half-turns of the helix of the smectic $C^*$ in the layer by unity, we present the results of the corresponding analysis for a strong surface anchoring for which we can assume that the critical angle $\varphi_c$ is determined by the shape of the surface anchoring potential. The orientation of director on the layer surface coincides with the alignment direction for the field value such that an integral number of half-pitches $p(E)$ of the helix fits into the layer thickness for a bulk smectic $C^*$ in an external field $E$; i.e., $\varphi(E) = 0$. When the field deviates from this value, the values of $\varphi_0(E)$ and $\varphi(E)$ differ from zero. For a certain value of the field, the free energies of two helical configurations with a number of half-turns in the layer differing by unity become equal, but the transition between these configurations takes place not at this point on the field scale, but upon a further change in the field due to the presence of a surface anchoring potential barrier between these configurations. Thus, a field hysteresis emerges during the transition; i.e., the points corresponding to the transition on the field scale are different for different directions of field variation. In actual practice, the pitch jump due to hysteresis takes place for the field value determined by the condition $\varphi(E) = \varphi_c$, where $\varphi_c$ is the critical angle at which the jumpwise change in the helix pitch takes place.

For example, in the case of strong anchoring, we can readily conclude that the solution of Eq. (16) for $\varphi$ is an odd function of $\varphi_0(E)$ and is bounded by the values $\varphi_c$ proceeding from the natural assumption that potential $W_S(\varphi)$ is an even function of $\pm\varphi_c$, viz., the angle of deviation of the director from the preferred orientation on the surface of the layer. (It should be recalled that in this case, $\varphi_0(E)$ and $\varphi(E)$ are measured from the direction of alignment and might have different values depending on the magnitude of field $E$.) For example, when angle $\varphi$ attains one of the limiting values of $\varphi_c$, the pitch of a cholesteric helix in the layer changes jumpwise and angle $\varphi$ assumes a new value also in a jump. If the point of the jump corresponded to the angle $\varphi_0(E_{j+})$ of free rotation of the director, the value of this angle determining the value of $\varphi$ after the jump is equal to $\varphi_0(E_{j+}) - \pi/2$, but in a new helical configuration with the number of half-turns in the layer thickness differing by unity from the previous value. If the field varies in the opposite direction, the pitch jump occurs when angle $\varphi$ attains the other limiting value $(-\varphi_c)$ and the corresponding value of free rotation angle is $\varphi_0(E_{j-}) = -\varphi_0(E_{j+})$ (where $E_{j-}$ is the field value corresponding to the jump in this direction of field variation). This relation leads to the expression connecting the values of pitches of the helix in the field for a bulk smectic $C^*$ at the points corresponding to the pitch jump in the field for opposite directions of its variation [5]:

$$\frac{1}{p(E_{j+})} + \frac{1}{p(E_{j-})} = \frac{N-1/2}{d},$$

$$d\left(\frac{1}{p(E_{j+})}\right) - \frac{1}{p(E_{j-})} \tag{19}$$

$$= \left(\frac{(\partial W_S(\varphi)/\partial\varphi)}{\pi W S_d}\right)_{\varphi = \varphi_c} + \frac{4\varphi_c/\pi - 1}{2},$$

where $N$ is the number of half-pitches in the layer for the initial configuration of the helix (before the jump for the field value $E_{j+}$ and after the jump for field values $E_{j-}$).

The fist relation from (19) is quite universal and does not depend in any way on the surface anchoring force at first glance. In fact, the values of the jump fields $E_{j+}$ and $E_{j-}$ depend on the surface anchoring, to be more precise, on the parameter $S_d$. This expression itself determines the relation between the values of the fields of the pitch jumps for opposite directions of field variation, i.e., the field hysteresis in pitch variations, which does not depend on the specific shape of the surface potential, and possesses predictability which can be verified experimentally. Namely, if the value of the pitch is measured experimentally at the jump point for a certain direction of field variation in the case of a strong surface anchoring, the value of the pitch at the jump point for the opposite direction of field variation can be calculated using formula (19) and subjected to experimental verification. In particular, it follows from formula (19) that in the limit of vanishingly weak surface anchoring (i.e., for $E_{j+} \longrightarrow E_{j-}$), the value of the pitch at the jump point is $d/(N - 1/2)$ as expected [4], and hysteresis is absent. The same formula implies that as the layer thickness increases, the hysteresis loop becomes smaller (to be more precise, hysteresis decreases upon an increase in $N$ for a fixed $S_d$).

The results obtained in this and in the previous section show that, as in the case of temperature-induced variations of the pitch, the variations of the helix in a layer in weak fields (for strong anchoring) are universal by nature, which is manifested in that their form depends not on the separate parameters of the sample under investigation, but only on one dimensionless parameter $S_d$.

## 7. PITCH VARIATION IN THE RAPINI MODEL POTENTIAL

In the Rapini model (see, for example, [3, 20, 21]), the surface anchoring potential should be substituted into formulas (1) and (2) in the form $W_S(\varphi) = -(W/2)\cos^2\varphi$, the critical angle being $\varphi_c = \pi/4$. In the case of strong anchoring, the relation (16) determining the equilibrium value of the angle $\varphi(E)$ of deviation of the director on the surface in an external field assumes the form

$$4S_d[\varphi(E) - \varphi_0(p(E))] + \sin 2\varphi(E) = 0, \tag{20}$$

where parameter $S_d$ is defined as

$$S_d = K_{22}/dW. \tag{21}$$

Expression (13) for the barrier assumes the form

$$\frac{B(T)}{W} = 2S_d\left[\frac{\pi}{4}\operatorname{sgn}\varphi_0(p(E)) - \varphi_0(p(E))\right]^2 + \cos^2\varphi(p(E)) - \frac{(\sin^2 2\varphi)p(E)}{8S_d} - \frac{1}{2}. \tag{22}$$

Using formulas (16) and (20), we can also write expression (22) in the form

$$\frac{B(T)}{W} = 2S_d\left[\frac{\pi}{4}\operatorname{sgn}\varphi_0(p(E)) - \varphi_0(p(E))\right]$$

$$\times\left[\frac{\pi}{4}\operatorname{sgn}\varphi_0(p(E)) + \varphi(p(E)) - 2\varphi_0(p(E))\right] \quad (23)$$

$$+\cos^2\varphi(p(E)) - \frac{1}{2}.$$

The calculations based on formulas (20)–(23) allow us to determine the angle $\varphi(E)$ of rotation of the director on the surface of the layer in the presence of anchoring as a function of the angle $\varphi_0(E)$ of free rotation of the director on the layer surface, i.e., the angle of rotation in the absence of surface anchoring. In particular, we can determine the values of various quantities at jump points. For example, the angle of free rotation of the director at a jump point is given by

$$\varphi_0(E_j) = \varphi_c + \frac{1}{4S_d}, \quad (24)$$

and the value of the helix pitch corresponding to a jump in a bulk smectic $C^*$ is defined by the formula

$$p(E_j) = \frac{2d}{N - 1/2 - 1/(2\pi S_d)}. \quad (25)$$

It follows from relation (24) that the interesting case in which $\varphi(E_j) = \varphi_0(E_j) = 0$ after a jump, i.e., a new helical configuration with a different value of $N$ corresponds to the minimum of the free energy for the new value of $N$, corresponds to $S_d = 1/\pi$. In this case, the difference in the angles formed by the director on the surface with the direction of alignment before and after the jump amounts to $\pi/4$. The maximum value of this difference is equal to $\pi/2$ and is realized in the limit of vanishingly weak surface anchoring, i.e., in the limit of an infinitely large value of $S_d$.

The values of helical pitches in a layer for the sample immediately before a jump and after it can easily be determined from the above formulas. For example, when the field in a layer with $N$ half-turns of the helix increases, the values of the pitch before and after the jump are given, respectively, by the formulas

$$p_c = \frac{2d}{N - 1/2}, \quad p_j = \frac{2d}{N - 1 + 2\varphi_j/\pi}, \quad (26)$$

where $\varphi_j$ is the angle of deviation of the director after the jump.

As the field decreases, the corresponding values of the helical pitches in the layer are given by

$$p_c = \frac{2d}{N + 1/2}, \quad p_j = \frac{2d}{N + 1 + 2\varphi_j/\pi}. \quad (27)$$

The second relation in (19) for the field hysteresis of the pitch assumes the form

$$d\left(\frac{1}{p(E_{j+})} - \frac{1}{p(E_{j-})}\right) = -\frac{1}{2\pi S_d}. \quad (28)$$

The results of calculations of the pitch and the angle of deviation of the director from the direction of alignment in the strong-anchoring approximation formulated above (under the assumption that the directions of alignment coincide on both surfaces of the layer) for the Rapini anchoring potential will be presented in the next section.

## 8. RESULTS OF CALCULATIONS

In order to illustrate quantitatively the general analysis of the behavior of a helix in a layer of a smectic $C^*$ in the presence of surface anchoring forces, which was carried out in the previous sections, we consider the results of numerical calculations in a situation which simplifies general relations to the maximum possible extent. It was noted earlier that in the general case, the solution of the problem depends on a large number of parameters determining the dependence of free energy (1) not only on the magnitude of the applied field, but also on the mutual orientation of the directions of alignment on the surfaces of the layer and on the direction of the applied field. Consequently, in order to simplify the presentation of the results of calculations, it is expedient first to confine our analysis to the case with the minimum number of computational parameters and then, after determining general regularities, to consider a more complicated problem, introducing additional parameters. It was found that the simplest case is that in which the directions of alignment and of the surface anchoring force coincide on both surfaces of the layer, the number of turns in the layer thickness is quite large, and the surface anchoring is quite strong so that relation (15) holds in a large range of electric field values. It is also convenient to assume (although this assumption is not of a fundamental nature) that in zero field, the director on both surface is oriented along the direction of alignment and the pitch variation in an external field for a bulk smectic $C^*$ can be described by the analytic formula (5) from [13]:

$$p(E) = p_0/[1 - (E/E_0)^2]^{1/2}, \quad (29)$$

where $E_0 = 4\pi P_S/\varepsilon_a$. It should be noted that relation (29) is valid for the so-called intermediate (between dielectric and ferroelectric) mode of helix unwinding in the field, for which $(P_S p_0)^2 = \pi\varepsilon_a K_{22}$.

Under the limitations formulated above, the field dependences of the pitch in a layer and of the deviation of the director from the alignment directions were calculated using the Rapini model potential on the basis of the formulas derived in the previous section. We assumed that $N = 10$; i.e., five turns of a helix fit into the layer thickness in zero field, and $S_d = 1/2\pi$.

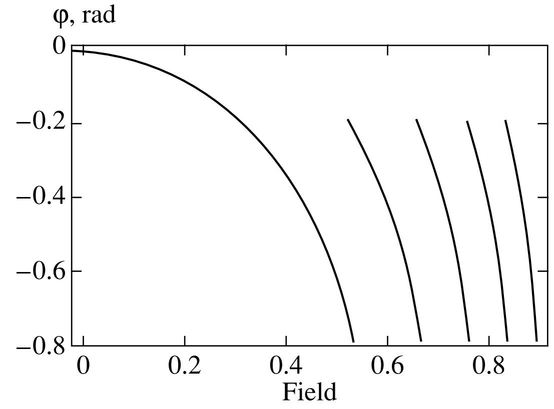

**Fig. 1.** Dependence of the angle of deviation of the director on the surface from the alignment direction on the field (reduced to its critical value for increasing field) ($S_d = 1/2\pi$; the initial number of half-turns in zero field is $N = 10$).

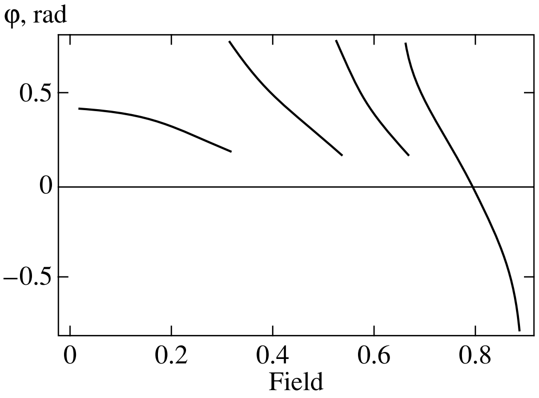

**Fig. 2.** Dependence of the angle of deviation of the director on the surface from the alignment direction on the field (reduced to its critical value for decreasing field) ($S_d = 1/2\pi$; the initial number of half-turns in zero field is $N = 10$).

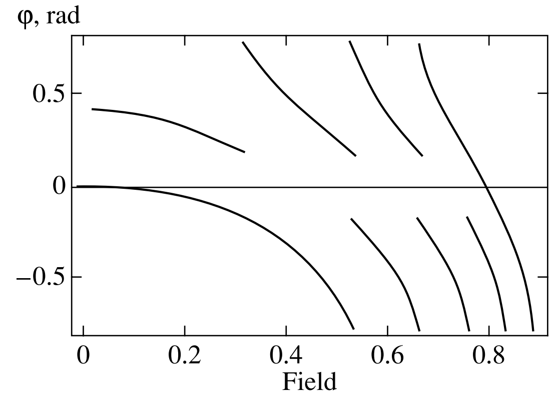

**Fig. 3.** Dependence of the angle of deviation of the director on the surface from the alignment direction on the field (reduced to its critical value) for both directions of field variation ($S_d = 1/2\pi$; the initial number of half-turns in zero field is $N = 10$).

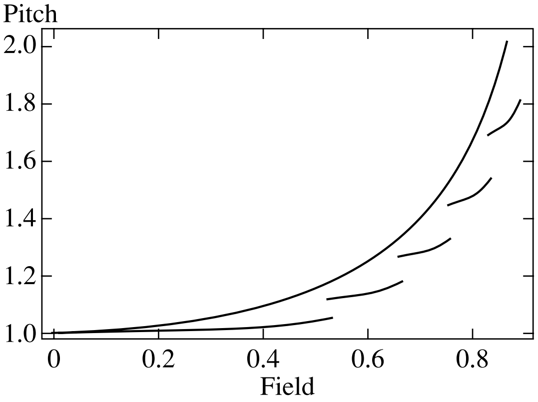

**Fig. 4.** Dependence of pitch (normalized to its initial value) on the field (reduced to its critical value) in an increasing field ($S_d = 1/2\pi$; the initial number of half-turns in zero field is $N = 10$). The solid curve depicts the corresponding dependence for a bulk smectic *C**, described by formula (29).

Figure 1 shows the field dependence of the angle of deviation of the director from the alignment direction for a field increasing from zero. After the attainment of the critical value, the angle of deviation undergoes a jump and a further increase in the field again leads to its continuous variation until the critical value of the angle is attained. It is known from the solution of the problem for a bulk smectic *C** [11–13, 16] that the helix pitch increases with the field; consequently, the angle of deviation of the director from the alignment direction becomes negative. Figure 2 shows a similar field dependence of the angle of deviation of the director, but in a decreasing field. As in an increasing field, the angle of deviation in the decreasing field experiences jumps after attaining the critical value. However, the angle of deviation in this case is positive except in the initial region of the decreasing field. Another characteristic feature of Fig. 2 is that in zero field the angle of deviation of the director does not assume the initial zero value from which its variation started in the increasing field (see Fig. 1). This is a consequence of the fact that the helical configuration in zero field in the present case has a number of half-turns over the layer thickness which is smaller by unity. This is a manifestation of bistability of the helical configuration in the field.

Figure 3 shows the field dependences of the angle of deviation of the director on the surface for both directions of field variation. This figure demonstrates the hysteresis of the jumps in the angle of deviation of the director (and pitch) in a layer in the presence of surface anchoring. The jump from a configuration with $N$ half-turns in a layer to a configuration with $N - 1$ half-turns occurs at a higher value of the field than the reverse transition from the configuration with $N - 1$ half-turns in a layer to that with $N$ half-turns taking place in the decreasing field.

Figure 4 shows the results of calculating the field dependence of the helix pitch in a layer with surface anchoring in an increasing field and the variation of the

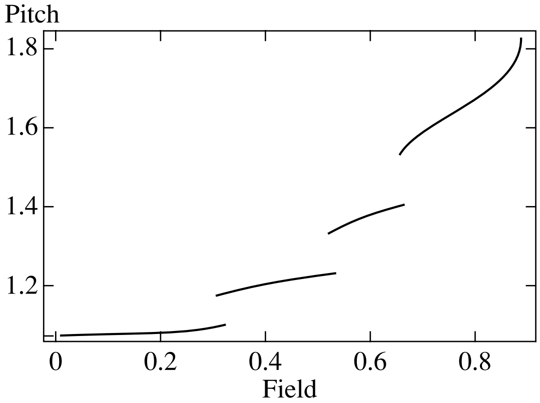

**Fig. 5.** Dependence of pitch (normalized to its initial value) on the field (reduced to its critical value) in a decreasing field ($S_d = 1/2\pi$; the initial number of half-turns in zero field is $N = 10$).

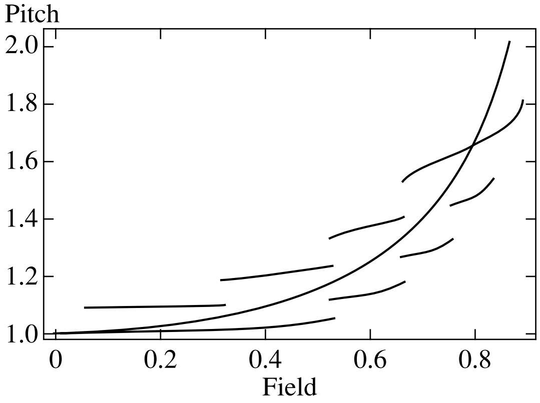

**Fig. 6.** Dependence of pitch (normalized to its initial value) on the field (reduced to its critical value) for both directions of field variation ($S_d = 1/2\pi$; the initial number of half-turns in zero field is $N = 10$). The solid curve depicts the corresponding dependence for a bulk smectic $C^*$.

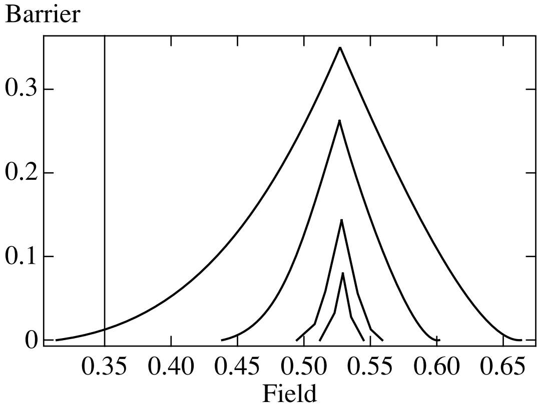

**Fig. 7.** Dependence of the height of the barrier (referred to the depth of anchoring potential) between configurations with $N = 9$ and 8 on the field (normalized to its critical value) in an increasing field (right branches of the curves) and of the barrier between configurations with $N = 8$ and 9 (left branches of the curves) in a decreasing field ($S_d = 1/2\pi$, $2/2\pi$, $5/2\pi$, and $10/2\pi$ (from top to bottom); the initial number of half-turns in a layer is $N = 10$).

pitch with the field in a bulk smectic $C^*$ for the values of parameters ensuring the applicability of formula (29). This figure demonstrates the pitch jumps in the layer as well as the fact the pitch in the layer in an increasing field remains smaller than the helix pitch in a bulk smectic $C^*$ for the same value of the field. Figure 5 shows the same dependence in a decreasing field. In this case, the pitch in the layer in the decreasing field remains larger than the helix pitch in a bulk smectic $C^*$ for the same value of the field (except in the initial region of field variation).

In Fig. 6, the results of calculation of the pitch variation are presented for both directions of field variation for a layer with anchoring and for a bulk smectic $C^*$. This figure demonstrates the jumpwise nature of pitch variations in the field for a layer (against the background of the continuous variation for a bulk crystal) and the hysteresis of pitch jumps as well as bistability of the helical configuration in a layer which were considered above.

Figure 7 shows the results of calculations of the barrier height between the helical configurations with the numbers of half-turns over the layer thickness differing by unity for various values of the parameter $S_d$ in the case of increasing (right slope of the curves) and decreasing (left slope) values of the applied field. The figure demonstrates a hysteresis for the transition between the configurations for opposite directions of the field variation as well as the dependence of the hysteresis width (the distance between the points corresponding to zero values of the barrier height for opposite directions of field variation) on the value of parameter $S_d$. The larger the value of this parameter, the smaller the hysteresis width.

## 9. CONCLUSIONS

The above analysis of the behavior of a helix of a smectic $C^*$ in a finite-thickness layer with surface anchoring forces in an external electric field revealed the characteristic features of the effect such as the existence of bistable states in the helical configurations of the smectic $C^*$, the jumps in the values of parameters of the smectic $C^*$ in a layer and their field hysteresis at the points of these jumps for opposite directions of field variation. Naturally, we did not touch upon many details of these dependences in specific experimental situations in view of the large number of the parameters of the problem in the general case. Our numerical calculations were made under the simplest assumptions concerning the parameters of a smectic $C^*$ layer. Nevertheless, the results presented in this article make it possible to predict the behavior of a smectic $C^*$ layer in an electric field in situations free of the limitations used also in our calculations.

For example, the inclusion of the dependence of the free energy of the layer on the direction of the applied field relative to the direction of the alignment on the surface leads to the dependence of the angle of deviation of the director from the preferred orientation at the jump points on the direction of the applied field. This means that, for example, that for the above-mentioned dependence, the angles of deviation of the director at the jump points for the field perpendicular to the direction of the alignment is smaller than for the field parallel to the direction of alignment (see Fig. 1). The inclusion of the field dependence of the critical angle in the same figure would lead to a change in the angles of deviation of the director at jump points with the field instead of remaining unchanged as in Fig. 1. The corresponding changes, which can be easily predicted qualitatively, will also be manifested in other figures.

It is also clear that the predominance of the dielectric or ferroelectric mechanisms of helix unwinding in the field (or, which is the same, the deformation of a simple helix by π- and 2π-solitons) would also change the pattern of the phenomenon analyzed above. For example, in the case of the ferroelectric mechanisms of deformation of a helix in the field, strictly opposite field orientation leads to different deformations of the helix of a smectic *C** in contrast to the dielectric mechanism in which the sign reversal of the field does not change the deformation of the helix. For this reason, the strong surface anchoring approximation used in the above calculations, which allows one to disregard the direction of the external field, “works” better when the dielectric mechanism of helix unwinding dominates.

As in the case of temperature-induced variations of the pitch [10], thermal fluctuations of director orientation will also affect the hysteresis of jumps (cause its decrease). The displacement of the jump point can be estimated, for example, with the help of Fig. 7, depicting the dependence of the height of the barrier between the helical configurations on the applied field. In the presence of fluctuations, a jump occurs not at the point of barrier vanishing, but at the point at which the barrier height becomes equal to temperature. However, it follows from general considerations that away from the jump points, the effect of fluctuations for the field hysteresis under investigation is weaker than in the case of the temperature hysteresis since the applied field suppresses orientational fluctuations. The last statement refers only to the range of parameters $S_d$ considered by us here and exceeding the critical value $S_{dc}$, for which the mechanism of the director slip on the surface through the anchoring potential barrier. For smaller values of the parameter $S_d$, the fluctuations of the director may play a certain role in discrete transitions of the helix between configurations with different numbers of turns fitting into the layer thickness. Consequently, in order to ensure the optimal conditions for the experimental observation of pitch jumps in a smectics *C** in a field in accordance with the mechanism of the director slip over the surface through the potential barrier, it is expedient to use the layers of liquid crystals for which parameter $S_d$ exceeds its critical value only slightly.

The peculiarities revealed in the variation of the helix in a smectic *C** in a layer are useful both for an analysis of the physics of liquid crystal state (e.g., for reconstructing the form of the surface anchoring potential from the results of measurements) and for applications (this especially refers to bistable states of helical configurations of a smectic *C**, jumps in the values of liquid crystal parameters in the layer, and their hysteresis. As in the case of temperature-induced pitch variations in the layer [4, 5], optical methods of investigations [22–24] appear the most suitable for studying the effects considered here.

Our analysis was carried out for static electric fields with an orientation perpendicular to the helical axis of a smectic *C**. However, a helix in a field exhibits a similar behavior for other orientations of the static field as well as for varying fields and a magnetic field. It should also be noted that we assumed that the only consequence of the external field applied to a smectic *C** is the deformation of its helix. However, it is well known [16] that other structural characteristic of liquid crystals (e.g., the tilt angle of molecules in a layer) may also vary in strong fields. However, an analysis of these effects is beyond the scope of this article.

## ACKNOWLEDGMENTS

This work was supported financially by the INTAS (grant no. 30234); one of the authors (E.K.) is also grateful to the Russian Foundation for Basic Research (project no. 00-02-17785) for partial financial support of the research.

## REFERENCES

1. *Proceedings of the 7th International Conference on Ferroelectric Liquid Crystals, Darmstadt University of Technology, Germany, 1999*, Ferroelectrics **243–246** (2000).
2. I. Musevic, R. Blinc, and B. Zeks, *The Physics of Ferroelectric and Antiferroelectric Liquid Crystals* (World Scientific, Singapore, 2000).
3. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals* (Clarendon Press, Oxford, 1993).
4. H. Zink and V. A. Belyakov, Mol. Cryst. Liq. Cryst. **265**, 445 (1995); Pis’ma Zh. Éksp. Teor. Fiz. **63**, 37 (1996) [JETP Lett. **63**, 43 (1996)].
5. H. Zink and V. A. Belyakov, Zh. Éksp. Teor. Fiz. **112**, 524 (1997) [JETP **85**, 285 (1997)]; Mol. Cryst. Liq. Cryst. **329**, 457 (1999).
6. T. Furukawa, T. Yamada, K. Ishikawa, *et al.*, Appl. Phys. B: Lasers Opt. **B60**, 485 (1995).
7. Z. Zhuang, Y. J. Kim, and J. S. Patel, Phys. Rev. Lett. **84**, 1168 (2000).
8. W. Greubel, Appl. Phys. Lett. **25**, 5 (1974).
9. G. S. Chilaya, Kristallografiya **45**, 944 (2000) [Crystallogr. Rep. **45**, 871 (2000)].
10. V. A. Belyakov and E. I. Kats, Zh. Éksp. Teor. Fiz. **118**, 560 (2000) [JETP **91**, 488 (2000)].

11. P. G. de Gennes, Solid State Commun. **6**, 123 (1968).
12. R. Dreher, Appl. Phys. Lett. **12**, 281 (1968).
13. V. A. Belyakov and V. E. Dmitrienko, Zh. Éksp. Teor. Fiz. **78**, 1568 (1980) [Sov. Phys. JETP **51**, 787 (1980)].
14. R. Dreher, Solid State Commun. **13**, 1571 (1973).
15. W. J. A. Goossens, J. Phys. (Paris) **43**, 1469 (1982).
16. S. A. Pikin, *Structural Transformations in Liquid Crystals* (Nauka, Moscow, 1981).
17. S. Chandrasekhar, *Liquid Crystals* (Cambridge Univ. Press, Cambridge, 1992, 2nd ed.).
18. P. O. Andreeva, V. K. Dolganov, R. Fouret, *et al.*, Phys. Rev. E **59**, 4143 (1999).
19. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982; Pergamon, New York, 1984).
20. L. M. Blinov, E. I. Kats, and A. A. Sonin, Usp. Fiz. Nauk **152**, 449 (1987) [Sov. Phys. Usp. **30**, 604 (1987)].
21. L. M. Blinov and V. G. Chigrinov, *Electrooptics Effects in Liquid Crystal Materials* (Springer-Verlag, New York, 1994), Chap. 3.
22. V. A. Belyakov and V. E. Dmitrienko, Sov. Sci. Rev., Sect. A **13**, 1 (1989).
23. V. A. Belyakov, *Diffraction Optics of Complex-Structured Periodic Media* (Nauka, Moscow, 1988; Springer-Verlag, New York, 1992).
24. V. A. Belyakov and V. E. Dmitrienko, in *Proceedings of the Second USA-USSR Symposium on Light Scattering in Condensed Matter* (Plenum, New York, 1979).

*Translated by N. Wadhwa*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 393–396.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 445–449.*
*Original Russian Text Copyright © 2001 by Orlov, Khvostantsev, Gromnitskaya, Stal'gorova.*

**SOLIDS**
**Structure**

# Pressure Effects on the Kinetic Properties and Phase Transitions in Lithium

**A. I. Orlov, L. G. Khvostantsev, E. L. Gromnitskaya*, and O. V. Stal'gorova**

*Vereshchagin Institute of High-Pressure Physics, Russian Academy of Sciences,*
*Troitsk, Moscow oblast, 142190 Russia*
*e-mail:grom@hppi.troitsk.ru*

Received May 4, 2000

**Abstract**—The pressure dependences of the electrical resistance and thermal electromotive force of lithium were measured at room temperature. The results substantiated the occurrence of a phase transition caused by increasing pressure (6.7 GPa). A phase transition was detected when pressure was decreased (6.4 GPa). Temperature effects on the pressures of these transitions were studied near room temperature. At pressures above 4 GPa, the pressure dependences of thermal electromotive force and of the velocity of ultrasonic shear waves in BCC lithium exhibited anomalies. The suggestion was made that applying pressure increased the role played by electron-phonon and phonon-phonon interactions in lithium. *© 2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

Much work, both theoretical and experimental, has been done to study the lightest alkali metal, lithium. Less attention has been given to pressure effects on the physical properties of lithium; in particular, data on the phase diagram are obviously insufficient. At room temperature, a phase transition in lithium was detected and interpreted as the BCC $\longrightarrow$ FCC transition [1]. Various authors reported different pressure values for this transition; the transition was only observed by various methods when pressure was increased (direct transition). Electrical resistance jump measurements gave 7.0 GPa [2, 3], the values obtained in X-ray diffraction studies were $6.9 \pm 0.4$ [1] and 7.5 GPa [4], whereas according to the Knight shift data [5], the transition pressure equaled 6.3 GPa [5].

It was shown in several works that the temperature of the martensite transition observed earlier at low temperatures near 77 K [6–8] increased in the pressure range 0–2 GPa [9–11]. The suggestion was made that this transition was similar in nature to the transition at room temperature and high pressures [9].

Theoretical calculations showed that the energies of all lithium phases with closely packed structures (BCC, FCC, HCP, and 9*R*) were close to each other [12, 13]. It is therefore not surprising that the phase transition at room temperature was accompanied by a volume change as small as 0.25% [1] (0.16% [4]), and at low temperatures, the transition did not involve any volume jump [9]. Ultrasonic studies of lithium showed that the phase transition at both low and room temperatures were preceded by a strong softening of shear characteristics [9, 14] related to changes in the low-frequency phonon spectrum region.

Studies of the kinetic characteristics of lithium such as electrical resistance and thermal electromotive force (emf) at pressures up to 9 GPa and a comparison of the results with the ultrasonic data on the velocity of transverse wave propagation allow us to gain insight into the state of the crystal lattice before the phase transition and can be used in theoretical work. Exact data on the phase transition pressure parameters at room temperature would provide a basis for further studies of the $p$–$T$ phase diagram of lithium at high temperatures.

## 2. EXPERIMENTAL

Electrical resistance and thermal emf were measured on a Toroid-type high-pressure apparatus [15], which could operate during increasing and decreasing pressure; in this respect, our procedure was different from those mentioned above [2, 3]. Another distinguishing feature of our measurements was virtually hydrostatic conditions and the use of comparatively large samples, whereas, for instance, in [2], measurements were performed on anvils and with very thin (about 0.3 mm) samples, which might distort the behavior of electrical resistance during the phase transition in lithium.

Pressure was created in a cell made of Teflon about 1 cm$^3$ by volume. The medium that transmitted pressure to the sample was liquid polyethylenesiloxane. This liquid proved to be as inert toward lithium as kerosene. A microscopic examination of the sample before and after applying pressure revealed no changes on the surface. An LE-1 lithium sample containing 99.0% Li was prepared in a neutral medium in the form of a cylinder 7 mm high and about 2 mm in diameter. In most of the experiments, pressure was determined from the known load dependence of pressure. This dependence was measured in several experiments using a two-section ampule

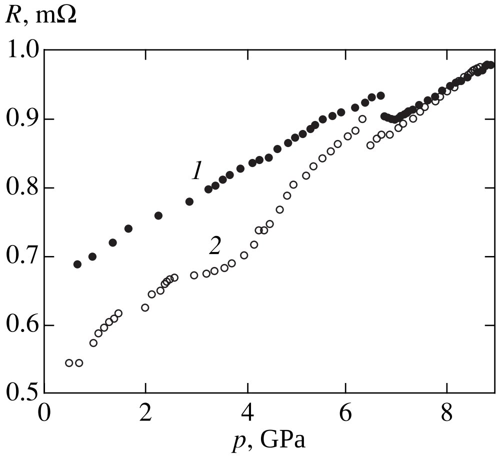

**Fig. 1.** Electrical resistance of lithium during (*1*) increasing and (*2*) decreasing pressure, $T = 293$ K.

[16], one section of which contained a manganin pressure gauge and a bismuth reference in a 4 : 1 methanol–ethanol mixture. The procedure that we used gave an absolute pressure measurement error not exceeding 0.1 GPa (the relative error of a single measurement in one experiment did not exceed 0.01 GPa).

Electrical resistance was measured by the four-point scheme.

In thermal emf measurements, the bottom of the sample was brought in contact with the copper lid of an ampule, which served as a good thermostat, and a plane heater was mounted on the top end of the sample. The temperature was measured by copper–constantan thermocouples 0.1 mm in diameter; the temperature difference (3–5°C) at the points of thermal emf measurements was maintained constant during pressure variations.

The velocity of ultrasonic shear wave propagation, $V_t(p)$, was measured in a chamber of the type "toroid with a plane bottom" by the pulsed ultrasonic method at 3–5 MHz frequencies [14]. To circumvent the difficulties arising from the large difference of the wave resistances of the chamber material (the chamber was made of a VK-6 metal-ceramic hard alloy) and lithium and from acoustic disturbances caused by the geometry of the working cell, the usual diameter of the chamber was increased to 16 mm and the sample was surrounded by a $2 \times 2$ mm$^2$ "belt-filter," which was a mixture of paraffin with VK-6 crumbs. A sample about 7 mm high and 16 mm in diameter was placed into a catlinite container and, at the ends, protected from contact with the material of the chamber by copper caps. Pressure was in each experiment determined from electrical resistance jumps of reference materials, namely, bismuth, thallium, and barium. Reference wires were situated vertically at the border between lithium and catlinite.

## 3. RESULTS AND DISCUSSION

The results obtained in studying the pressure dependence of the electrical resistance of lithium $R(p)$ were quite reproducible from one experiment to another. Some measurements were taken in a two-section ampule, that is, in the presence of the manganin pressure gauge and bismuth reference. The results are shown in Fig. 1. The electrical resistance of lithium was found to increase linearly as pressure grew to 6.7 GPa. The phase transition from the BCC to the FCC structure was accompanied by about a 5% decrease in electrical resistance. Further, in the high-pressure phase, resistance continued to increase. The very large direct transition-induced electrical resistance jump reported in [2] was, we believe, caused by a methodological error. In an anvil-type high-pressure chamber used in [2], samples have very small dimensions. When a two-point scheme of resistance measurements is used and the contacts have a strong tendency to oxidize, the instrumental effects are fairly difficult to take into account. Their subtraction from the total measured resistance can result in underestimating the fraction of resistance due to the sample proper and, therefore, in overestimating the relative decrease in resistance due to the phase transition. This methodological difficulty has been mentioned in [3].

The pressure dependences of lithium resistance at various temperatures allowed us to determine the parameters of the BCC ⟷ FCC transitions. At 20°C, the pressure of the direct transition in lithium was $p = 6.7 \pm 0.1$ GPa and the pressure of the onset of the reverse transition was $p = 6.4 \pm 0.1$ GPa. We obtained different $R(p)$ isotherms at 15 to 25°C, which allowed us to estimate the slope of the boundaries of the direct and reverse transitions in the $p$–$T$ diagram; this slope equaled 0.03 GPa/K. This result leads us to conclude that the spread of the literature data on the direct phase transition pressure is caused not only by different errors of different procedures and apparatus but also by the strong temperature dependence of the transition onset pressure. Extrapolating the phase transition boundary from room to low temperatures with the use of the obtained coefficient, 0.03 GPa/K, yields the martensite transition coordinates reported for lithium in [6, 9].

Thermal emf measurements under pressure, $S(p)$, were performed in a one-section ampule. A bismuth pressure reference was placed near the sample. The load dependence of pressure was determined preliminarily as described above, with the use of a manganin pressure gauge placed together with the bismuth reference in a hydrostatic mixture of alcohols (methanol–ethanol). Following [17], pressure-induced changes in the calibration of the copper-constantan thermocouple were assumed to be insignificant.

The thermal emf of lithium was measured with reference to the thermal emf of copper, which equaled 1.8 μV/K and insignificantly changed as pressure varied. The obtained pressure dependence of the absolute thermal emf of lithium is shown in Fig. 2. The thermal emf of BCC lithium increased as pressure grew up to the onset of the phase transition, and the phase transition caused a sharp decrease in its value. The $S(p)$ dependences measured while increasing and decreasing pressure coincided, which is once again evidence of high hydrostatic pressure conditions in our experiments. Some spread of the data at pressures up to $p = 2$ GPa can be explained by methodological difficulties of measurements at low pressures in a Toroid-type chamber. Nevertheless, extrapolating the pressure dependence of lithium thermal emf from pressures above 2 GPa to the atmospheric pressure gave $S_0 = 10.5$ μV/K, which equaled the handbook value for normal conditions [18]. An anomaly was regularly observed in the $S(p)$ dependence at $p > 4$ GPa during both increasing and decreasing pressure. A substantial anomaly in this pressure region was also observed in the nuclear magnetic resonance study [5] of the pressure dependence of the self-diffusion coefficient of lithium.

The results obtained in measuring the velocity of ultrasonic shear waves in lithium at room temperature and pressures up to 7 GPa, $V_t(p)$ [9], are shown in Fig. 3. The $V_t(p)$ dependence has some peculiarities. The velocity changed insignificantly in the initial region and remained virtually constant in the pressure range 1.5–2.5 GPa. Between 4 and 6.4 GPa, it decreased sharply and then began to increase; that is, above 6.4 GPa, a new, “harder,” phase was formed. It follows from the calculations performed in [14] that, before the BCC $\longrightarrow$ FCC transition, the lithium lattice loses stability toward shear in the {110} plane in the $[1\bar{1}0]$ direction under the action of pressure. This should result in a decrease in the $c' = (c_{11} - c_{12})/2$ elastic constant and manifest itself by anomalies of the $V_t(p)$ dependence for polycrystalline lithium. This is exactly what follows from the character of the pressure dependence of the velocity of ultrasonic shear waves obtained in our experiments. Note in addition that a decrease in $V_t(p)$ at $p > 4$ GPa correlates with anomalies of the $S(p)$ dependence in the same pressure range.

A comparison of the $R(p)$, $S(p)$, and $V_t(p)$ dependences shows that sharp changes in the kinetic and elastic physical properties of lithium, which are clearly caused by the phase transition, are observed at almost equal pressures. Some difference in the recorded phase transition pressures is likely to be caused by the temperature dependence of the transition and certain deviations from hydrostatic conditions in ultrasonic experiments.

The thermal emf of lithium and other alkali metals at pressures up to 0.3 GPa was very accurately measured in [19]. The thermal emf of lithium equaled $S_0 = 10.6$ μV/K at atmospheric pressure and slightly decreased as pressure increased (to $S = 10.5$ μV/K at $p = 0.3$ GPa). As mentioned, the accuracy of $S$ measurements below 2 GPa was fairly low in our experiments. Our data are therefore difficult to compare with those reported in [19]. It can only be noted that there are no discrepancies exceeding experimental errors.

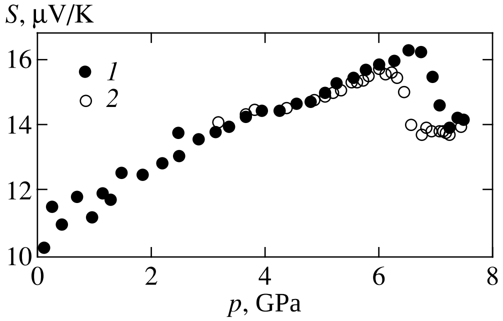

**Fig. 2.** Thermal emf of lithium during (*1*) increasing and (*2*) decreasing pressure at room temperature.

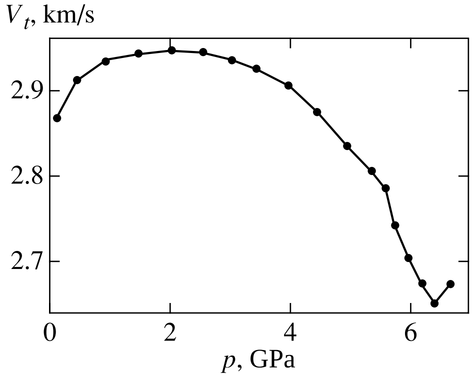

**Fig. 3.** Velocity of transverse ultrasonic waves in lithium at room temperature during increasing pressure.

The data of ultrasonic measurements [14] were used to obtain the pressure dependence of lithium compression at $p < 2.5$ GPa and room temperature. The dependence was described by the polynomial

$$\frac{V}{V_0} = 0.999 - 0.0792p + 0.00729p^2,$$

where $V_0$ is the volume at atmospheric pressure. Based on these results, we obtained the dependence of thermal emf

$$Z = (S - S_0)/S_0,$$

where $S_0$ is the thermal emf at atmospheric pressure, on the volume of lithium under compression to 2.5 GPa (Fig. 4). The $Z(V)$ dependence was constructed by averaging the experimental $S(p)$ values (see Fig. 2).

According to the theoretical calculations [19] based on the model of free electrons, $Z$ should decrease as pressure increases, which is at variance with the data on lithium even at small pressures. A sharp increase in $Z$ above 1.5 GPa and a substantial decrease in the velocity of ultrasonic shear wave propagation at pressures much bellow the transition pressure are evidence that the

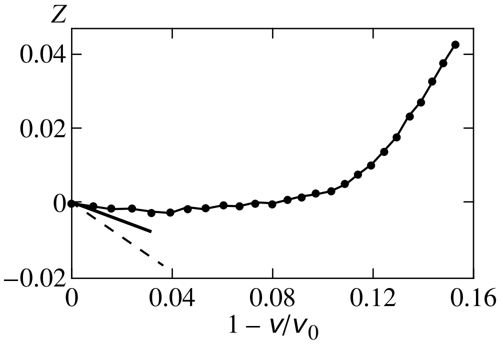

**Fig. 4.** Relative changes in lithium thermal emf versus changes in relative volume. Undotted solid line is the data from [19], and dashed line is the theoretical dependence obtained in [19] for the model of free electrons.

phase transition is preceded by strengthening of electron-phonon and phonon-phonon interactions in lithium. It follows that the model of free electrons is inapplicable to the kinetic properties of lithium.

To summarize, we determined the $p$–$T$ coordinates of the direct and reverse transitions in lithium close to room temperature with an error of 0.1 GPa. Extrapolating the boundary of the phase transition from room to low temperatures with the use of the obtained coefficient, 0.03 GPa/K, yields the martensite transition coordinates reported for lithium in [9]. Note that ultrasonic experiments at low temperatures show that extrapolating the martensite transition boundary from low to room temperatures yields a pressure of about 7 GPa. This gives more reasons to treat both transformations as one structural phase transition. To conclusively substantiate this suggestion, it would be interesting to perform measurements under pressure in the temperature range $T = 100$–$300$ K.

## REFERENCES

1. B. Olinger and J. W. Shaner, Science **219**, 1071 (1983).
2. R. A. Stager and H. G. Drickamer, Phys. Rev. **132**, 124 (1963).
3. T. N. Lin and K. J. Dunn, Phys. Rev. B **33**, 807 (1986).
4. M. Hanfland, I. Loa, K. Syassen, *et al.*, Solid State Commun. **112**, 123 (1999).
5. R. Bertani, M. Mali, J. Roos, and D. Brinkmann, J. Phys.: Condens. Matter **2**, 7911 (1990).
6. C. S. Barrett, Acta Crystallogr. **9**, 671 (1956).
7. A. W. Overhauser, Phys. Rev. Lett. **53**, 64 (1984).
8. W. Schwarz, O. Blascko, and I. Gorgas, Phys. Rev. B **44**, 6785 (1991).
9. O. V. Stal'gorova and E. L. Gromnitskaya, Fiz. Tverd. Tela (St. Petersburg) **37**, 1671 (1995) [Phys. Solid State **37**, 908 (1995)].
10. H. G. Smith, R. Berliner, J. D. Jorgensen, *et al.*, Phys. Rev. B **41**, 1231 (1990).
11. V. G. Vaks, M. I. Katsnelson, V. G. Koreshkov, *et al.*, J. Phys.: Condens. Matter **1**, 5319 (1989).
12. A. Y. Liu and L. Cohen, Phys. Rev. B **44**, 9678 (1991).
13. A. Y. Liu, A. A. Quong, J. K. Freericks, *et al.*, Phys. Rev. B **59**, 4028 (1999).
14. E. L. Gromnitskaya and O. V. Stal'gorova, Zh. Éksp. Teor. Fiz. **106**, 1453 (1994) [JETP **79**, 785 (1994)].
15. L. G. Khvostantsev, L. F. Vereshchagin, and A. P. Novikov, High Temp.–High Press. **9**, 637 (1977).
16. A. I. Orlov and L.G. Khvostantsev, Experiment in Geosciences **2** (2), 56 (1998).
17. F. P. Bundy, J. Appl. Phys. **32**, 483 (1961).
18. *Properties of Elements: Reference Book*, Ed. by M. E. Drits, P. B. Budberg, N. T. Kuznetsov, *et al.* (Metallurgiya, Moscow, 1997), Vol. 1, p. 37.
19. J. S. Dugdale and J. N. Mundy, Philos. Mag. **6**, 1463 (1961).

*Translated by V. Sipachev*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 397–409.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 450–463.*
*Original English Text Copyright © 2001 by Manakova.*

**SOLIDS**
**Electronic Properties**

# The Formation of Heavy-Fermion States in Non-Fermi-Liquid Impurity Systems[¶]

**L. A. Manakova**
*Kurchatov Institute Russian Research Centre, pl. Akademika Kurchatova 1, Moscow, 123182 Russia*
*e-mail: manakova@kurm.polyn.kiae.su*

Received July 6, 2000

**Abstract**—A mechanism for the occurrence of heavy-fermion states in non-Fermi-liquid (NFL) metals with *f*-shell impurities is proposed. The impurity with an unstable valence is suggested to have an energy spectrum consisting of a deep *f*-level and quasicontinuum states (narrow band) in resonance with the Fermi energy. Depending on the impurity concentration, the single-site NFL states are generated by the two-channel Kondo scattering for the low concentration (the Kondo regime) or by the screening interaction for a relatively high concentration (the X-ray-edge regime). It is shown that the NFL states are unstable against the scattering of the NFL excitations by electron states of the narrow band. This scattering generates additional narrow Fermi-liquid (FL) resonances at/near the Fermi level in the Kondo regime and in the X-ray-edge regime. The mixed-valence states are shown to be induced by new FL resonances. The mixed valence mechanism is local and is related to the instability of single-site NFL states. The FL resonances lead to the existence of additional energy scales and of pseudogaps near the Fermi level in the mixed-valence states. They also considerably narrow the region with a nearly integer valence. *© 2001 MAIK "Nauka/Interperiodica".*

## 1. INTRODUCTION

At present, intermetallic compounds with the *f*-shell atoms Ce or U are an important class of alloys in which non-Fermi-liquid (NFL) behavior is observed (see [1, 2] for a review). The anomalous temperature dependences of their linear specific heat, magnetic susceptibility, and resistivity strongly support the NFL scenario. The Ce and U ions carry magnetic dipole or electric quadrupole moments that interact with the spins and charges of the conduction electrons, thereby giving rise to the Kondo effect and the NFL behavior at low temperatures. The *f*-electron compounds of interest have been alloyed with nonmagnetic elements (with a few possible exceptions) [2, 3]. The thermodynamic measurements evidence in favor of the quadrupole two-channel Kondo model introduced in [4]. We note that according to photoemission spectra, the U-based compounds look much more like the mixed-valence ones (see references in [5, 6]). Recently [7], it was shown that the temperature behavior of the specific heat and magnetic susceptibility is governed by nonuniversal power-law dependences for a relatively high concentration of the *f*-shell atoms.

Taking the foregoing into account, it would be highly desirable to have the unified treatment involving the explanation of two essential facts:

(1) the coexistence of the single-ion two-channel Kondo effect and the mixed-valence state;

(2) the possibility of nonuniversal power-law energy dependences on the parameters.

It should be noted that the role of instabilities of the NFL states in forming the heavy-fermion (HF) states has not been completely clarified. At the same time, it is well known that the single-ion NFL state is unstable against any perturbation that eliminates the orbital or spin degeneracy of the impurity. Two instability mechanisms are presently known in the two-channel quadrupole and orbital Kondo model. In [4], the instability is induced by the Jahn–Teller distortions of the impurity site. The second mechanism [8] attributes the instability to the channel anisotropy. As shown in [9] and [10], there occurs a new physical realization of the two-channel quadrupole Kondo model and of the NFL state instability against the scattering generated by the tunnel process in the doped size-quantized structures. The physical reason of the instability is the existence of additional narrow Fermi-liquid (FL) resonances induced by tunneling.

For a metal containing orbitally degenerated deep impurity states, it was shown in [11] that the NFL state can be unstable against the scattering of the multiparticle excitations having different $z$ projections of the quadrupole moment.

In this paper, we propose a new mechanism for the occurrence of HF states in NFL metals with the *f*-shell impurities. We assume that a specific feature of atoms with an unstable valence is an energy spectrum that contains two unfilled shells: the orbitally degenerate deep *f*-level states and the atomic quasicontinuum states (narrow band) near the Fermi level. As shown below, the scattering of the NFL excitations by atomic quasicontinuum states, which is potential in its character, generates additional FL resonances near the Fermi level. Along with the NFL excitations, new FL resonances form an additional branch of heavy-fermion states with the characteristic energy that is much smaller than the

---

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

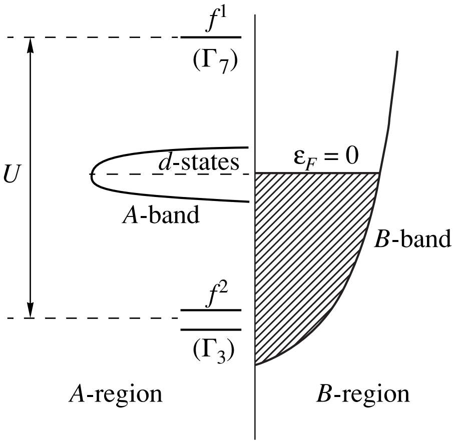

**Fig. 1.** The initial electron spectrum. $A$ and $B$ are the impurity region and the conduction band, respectively, $U$ is the Hubbard repulsion. The $A$-band consists of the impurity states in resonance with the Fermi level.

width of the NFL resonance (even in the case of the Kondo effect). New FL resonances generate the mixed-valence state. The heavy-fermion states have a local origin within the treatment proposed below.

In conclusion, we briefly discuss the temperature transitions within the proposed framework and the role of single-site NFL fluctuations in the "concentrated" heavy-fermion systems.

## 2. THE IMPURITY MODEL AND THE SCATTERING PROBLEM FOR AN INTERACTING SYSTEM

**2.1.** It is commonly known that an ion with unfilled $d$- or $f$-shells partially retains its atomic properties in a crystal. This is possible due to the presence of a centrifugal barrier separating the region $A$ in which the atomic forces act from the region $B$ where the lattice potential acts. The height of the barrier is comparable to other characteristic energies of the system, i.e., the Fermi energy and the interatomic interaction energy. The typical energy spectrum of lanthanide and actinide ions with an unstable valence seems to contain quasilocal deep $f$-levels together with the quasicontinuum states under the centrifugal barrier. The atomic quasicontinuum may be formed by the $d$-shell states being in resonance with the conduction band states at the Fermi level. A similar impurity model with a highly degenerate $f$-level was also considered in [12]. The initial electron spectrum before mixing is depicted in Fig. 1.

The Hamiltonian of the system is given by

$$H = H_A + H_B + H_{AB}, \tag{2.1}$$

where $H_B$ and $H_A$ are the Hamiltonians of the conduction band and of the impurity region. The Hamiltonian $H_{AB}$ describes the hybridization $H_h$ and the scattering $H_{sc}$ between electron states of the conduction band and the impurity region. The Hamiltonian of the impurity region is given by

$$H_A = H_0^f + H_0^d + H_U^f,$$

where $H_0^f$ is the Hamiltonian of the deep level, $H_U^f$ is the Hubbard repulsion, and $H_0^d$ is the Hamiltonian of the narrow $A$-band. In what follows, we do not consider the intraband interactions, assuming that they are weak compared to the interaction between the deep level and the band electrons. Therefore, $H_0^d$ is the Hamiltonian of the noninteracting $d$-electrons. It is also assumed that the $d$- and $f$-shells are not mixed in the impurity region. We start from the low-lying electron configuration of the isolated ion and then take the mixing with the conduction electrons into account in the spirit of the Anderson model with two unfilled shells.

We consider the situation where the deep level is a $\Gamma_3$ quadrupole (non-Kramers) doublet of the crystal field interacting with the $\Gamma_8$ quartet of the conduction electrons. However, we emphasize that the mechanism proposed here can be applied for all compounds in which the symmetry allows the local quartet of conduction states to couple to the two-fold degenerate level and an additional potential scattering of the multiparticle excitations exists.

For U-based compounds, the $\Gamma_3$ doublet in the $5f^2$ configuration is formed as a result of splitting the multiplet with the total moment $J = 4$ with the cubic crystal field. The $\Gamma_3$ doublet has an electric quadrupole moment and no magnetic dipole moment. The quantum numbers of the $\Gamma_3$ level electron are the numbers of lines $\mu$ for the irreducible representation of the point group $\mu_{\Gamma_3} = \pm 1$. The two quantum number values $\mu = \pm 1$ correspond to the projections of the quadrupole moment on the $z$ axis, i.e., $Q_{zz} = \pm 8$.

The multiparticle configuration of the unfilled shell is denoted by $n$; $|n; \mu\rangle$, where $n$ indicates the number of electrons and $\mu$ is the set of quantum numbers characterizing the configuration.

For relatively large values of the Hubbard repulsion in the absence of hybridization, the ground state configuration of the ion $U^{4+}$ is the singly occupied $\Gamma_3$ doublet with the electron configurations $|1; +1\rangle$ and $|1; -1\rangle$ and the energy $E_f$. The electron creation operators and electron numbers correspond to the singly occupied states:

$$\begin{gathered} f^+_{\mu = +1} = |1; +1\rangle\langle 0; 0|; \quad f^+_{\mu = -1} = |1; -1\rangle\langle 0; 0|; \\ n_{f\mu} = f^+_\mu f_\mu; \quad \sum_\mu n_{f\mu} = 1; \\ H_0^f = \sum_{n,\mu} E_{n,\mu} |n; \mu\rangle\langle n, \mu| \equiv \sum_\mu E_{f\mu} f^+_\mu f_\mu, \\ H_U^f = \sum_{\mu\mu'} U_{\mu\mu'} n_{f\mu} n_{f\mu'} (1 - \delta_{\mu\mu'}). \end{gathered} \tag{2.2}$$

The $\Gamma_3$ states are hybridized with the partial conduction band waves having the total angular momentum $j = 5/2$. Taking the splitting of the $j = 5/2$ multiplet by the cubic crystal field into account amounts to the transition from the angular momentum representation to the irreducible representations of the point group of the crystal. The latter representation has the quartet $\Gamma_8$ that can be hybridized with the $\Gamma_3$ doublet. The $\Gamma_8$ quartet possesses two groups of the states: $\Gamma_8^{(+)}$, $\Gamma_8^{(-)}$ with $\Gamma_8^{(+)} = |\Gamma_8; 2\rangle, |\Gamma_8; 1\rangle$ and $\Gamma_8^{(-)} = |\Gamma_8; -2\rangle, |\Gamma_8; -1\rangle$. The groups $\Gamma_8^{(+)}$ and $\Gamma_8^{(-)}$ correspond to different signs of $j_z$. Different signs of $j_z$ correspond to different sings of the spin projection $\sigma_z$. In addition, the states $|\Gamma_8; \pm 2\rangle$ and $|\Gamma_8; \pm 1\rangle$ have the respective $z$ components of the quadrupole moment $Q_{zz} = \pm 8$.

In other words, the quartet $\Gamma_8$ of partial waves decomposes into the tensor product $\Gamma_3 \otimes \Gamma_7$. It is therefore described by a combination of the "orbital" ($\Gamma_3$) and the "spin" (magnetic) ($\Gamma_7$) indices.

The partial states of the conduction electrons mixed with the $\Gamma_3$ doublet can therefore be classified by the quantum numbers $|\varepsilon; \gamma, \alpha\rangle$, where $\varepsilon = v_F k - \varepsilon_F$, with $k$ being the wave vector modulus and $\varepsilon_F$ being the Fermi energy. In what follows, we choose the position of the Fermi level as zero. The quantum number $\gamma = 2, 1 \equiv \mu$ corresponds to the two values of the quadrupole moment within the groups $\Gamma_8^{(+)}$ and $\Gamma_8^{(-)}$; the magnetic quantum numbers $\alpha = \pm$ distinguish the respective groups $\Gamma_8^{(+)}$ and $\Gamma_8^{(-)}$. The operators $a^+_{B\mu\alpha}(\varepsilon)$ describe the states $|\varepsilon; \mu, \alpha\rangle$ in the $B$-band.

In terms of these states, the hybridization Hamiltonian

$$H_h = \sum_{\mathbf{k}\sigma n}\sum_{\mu_{n+1}\mu_n}(V^f_{\mu_{n+1}\mu_n\sigma}(\mathbf{k})a^+_{B\sigma}(\mathbf{k})|n+1; \mu_{n+1}\rangle\langle n; \mu_n| + \text{H.c.}),$$

where $a^+_{B\sigma}(\mathbf{k})$ creates the conduction band electron with the spin $\sigma$ and the wave vector $\mathbf{k}$, can be written as

$$H_h = \sum_{\mu\alpha}\int_{-\infty}^{+\infty} d\varepsilon \rho_{0B}(\varepsilon)(V^f_{\mu\alpha}(\varepsilon)a^+_{B\mu\alpha}(\varepsilon)f_\mu + \text{H.c.}). \quad (2.3)$$

Here, $\rho_{0B}(\varepsilon)$ is the density of states (DOS) in the $B$-continuum, the terms with $V^f_{\mu\mu'\alpha}(\varepsilon)$, $\mu \neq \mu'$ are neglected because of the cubic symmetry, and the matrix elements $V^f_{\mu\mu\alpha}(\varepsilon)$ are denoted by $V^f_{\mu\alpha}(\varepsilon)$.

In finding the interaction Hamiltonian in what follows, it is significant that, because of the band state symmetry, the hybridization matrix elements $V^f_{\mu\alpha}(\varepsilon)$ are nonzero for both components of the $\Gamma_3$ doublet with $\mu = \pm 1$. This means that the matrix elements $V^f_{\mu_{n+1}\mu_n\sigma}(\mathbf{k})$ are spatially nonlocal. We additionally assume the hybridization matrix elements to be independent of the sign of the $z$ component $j_z$ of the total momentum $j$, i.e., $V^f_{\mu\alpha} \equiv V^f_\mu$.

The Hamiltonian $H_{AB}$ in Eq. (2.1) also involves the scattering between electron states of the $A$- and $B$-bands. In terms of the partial states, the scattering Hamiltonian is given by

$$\begin{aligned}H_{sc} &= \sum_{\mu\alpha}\int_{-\infty}^{+\infty} d\varepsilon\rho_{0A}(\varepsilon) \\ &\times \int_{-\infty}^{+\infty} d\varepsilon'\rho_{0B}(\varepsilon')T^{AB}_{\mu\mu}(\varepsilon, \varepsilon')a^+_{A\alpha\mu}(\varepsilon)a_{B\alpha\mu}(\varepsilon'),\end{aligned} \quad (2.4)$$

where the operators $a_{A\mu\alpha}(\varepsilon)$ describe the states in the atomic continuum (A). The scattering with $\mu \neq \mu'$ is absent because of the cubic symmetry. We assume that the scattering matrix elements, as well as the hybridization ones, are independent of the quantum number $\alpha$.

In defining the NFL states, it is important to account for the splitting of the $f$ doublet ground state due to a local deviation from the cubic symmetry at the impurity site. In the Hamiltonian, the splitting is described by the term

$$H_\Delta = \Delta\hat{\tau}^z_f. \quad (2.5)$$

Because the Hubbard repulsion $U$ is the largest parameter in the problem, it is convenient first to take the effective interaction induced by $U$ into account and then to use the multiparticle states as a basis for solving the scattering problem. As shown below, the system described by the Hamiltonian $H$ in (2.1) has two physical mechanisms generating singularities at/near the Fermi level. The Hubbard repulsion $U$ generates the effective interaction between conduction electrons and the deep level. This interaction induces an NFL resonance at the Fermi level in the $B$-band. The scattering of the multiparticle excitations in the conduction band by the electron states of the $A$-band generated by $H_{sc}$ results in the formation of additional FL resonances near the Fermi level.

**2.2.** In the system with the Hamiltonian $H$, the excitations are completely described by the Green function

$$\hat{G}_{f\mu}(z) = \langle f_\mu|(z - \hat{H})^{-1}|f_\mu\rangle.$$

Because the energy $U$ is dominant, it is essential to properly treat correlations on site. To calculate $\hat{G}_{f\mu}(z)$, we use the method of the equations of motion [19] that correctly accounts for these on-site correlations. This gives

$$\hat{G}_{f\mu}(z) = \hat{G}^0_{f\mu}(z)\frac{1-\Sigma_{A\mu}(z)\Sigma_{B\mu}(z)}{\hat{D}^{AB}_{\mu}(z)}, \tag{2.6}$$

where

$$\hat{G}^0_{f\mu}(z) = [z-\varepsilon_f-\hat{\Sigma}^h_{B\mu}(z)]^{-1}$$

is the Green function of the interacting system without scattering; we then have

$$\hat{D}^{AB}_{\mu}(z) = 1-\Sigma^{sc}_{A\mu}(z)\hat{W}_{B\mu}(z),$$

$$\hat{W}_{B\mu}(z) = \Sigma^{sc}_{B\mu}(z)+\Sigma^{sc}_{B\mu}(z)\Sigma^{sc}_{\mu B}(z)\hat{G}^0_{f\mu}(z).$$

Equation (2.6) implies that the full Green function $\hat{G}_{f\mu}(z)$ has features of two types. The function $\hat{G}^0_{f_r\mu}(z)$ describes the contributions of the multiparticle resonances at the Fermi level due to the interaction between the conduction electrons and the deep level. The second factor in Eq. (2.6) is generated by the scattering of the multiparticle excitations via the atomic quasicontinuum states. The scattering results in additional singularities, namely, simple poles near the Fermi level. The pole positions are determined by the equation $\hat{D}^{AB}_{\mu}(z) = 0$. The self-energy functions $\Sigma^h_{B\mu}(z)$ and $\Sigma^{sc}_{\nu\mu}(z)$ with $\nu = A$, $B$ are expressed as spectral expansions of multiparticle Green functions of the $A$- and $B$-bands,

$$\begin{aligned}\Sigma^{sc}_{\nu\mu}(z) &= \left|T^{\nu}_{\mu}(0)\right|^2\sum_p\frac{1}{z-\varepsilon_{\nu}(p)}\\ &= \left|T^{\nu}_{\mu}(0)\right|^2\int d\varepsilon\frac{\rho_{\nu}(\varepsilon)f(\varepsilon)}{z-\varepsilon},\\ \Sigma^h_{B\mu}(z) &= \left|V^f_{\mu}(0)\right|^2\int d\varepsilon\frac{\rho_B(\varepsilon)f(\varepsilon)}{z-\varepsilon},\end{aligned} \tag{2.7}$$

where $\varepsilon_{\nu}(p)$ is the excitation spectrum at the Fermi level, $\rho_{\nu}(\varepsilon)$ is the DOS corresponding to this spectrum, and $f(\varepsilon)$ is the Fermi function. In Eqs. (2.7), it is assumed that $V^f_{\mu}(\varepsilon) \approx V^f_{\mu}(0)$ and the scattering matrix elements are separable:

$$T^{AB}_{\mu\mu}(\varepsilon, \varepsilon') \approx T_{AB}(0, 0) = T^A_{\mu}(0)T^B_{\mu}(0),$$

where $T^A_{\mu}(0)$ is dimensionless.

Without the interaction, we have $\rho_{\nu}(\varepsilon) = \rho_{0\nu}$ and $\hat{G}^0_{f\mu}(z) = [z - \bar{\varepsilon}_{f\mu} - i\gamma_{f\mu}]^{-1}$, where $\bar{\varepsilon}_{f\mu}$ is the energy of the deep level renormalized by the hybridization and $\gamma_{f\mu}$ is the width of this level. In this case, both $\hat{G}^0_{f\mu}(z)$ and $\Sigma^0_{\nu\mu}(z)$ have no singularities near the Fermi level.

In the interacting system as $U \longrightarrow \infty$, we are interested in the case where the dominant effect of the interaction is the generation of a multiparticle resonance (the $f_r$-level) near the Fermi level. The Green function $\hat{G}^0_{f_r\mu}(z)$ of this resonance must then be inserted in Eq. (2.6). The multiparticle peaks in $\hat{G}^0_{f_r\mu}(z)$ at the Fermi level determine the properties of the DOS $\rho_B(\varepsilon)$ and of the self-energy functions $\Sigma^{sc,\,h}_{B\mu}(z)$.

To obtain the density of states at the Fermi level in the interacting system, the following consideration can be used. It is known [19] that the exact Green function of the conduction electrons in the impurity Anderson model is given by (in our notation)

$$\begin{aligned}G_{\mu\alpha}(k, k'; z) &= \delta_{\mu\mu'}\delta_{\alpha\alpha'}\delta(k-k')G_{0\mu\alpha}(k, k'; z)\\ &+ G_{0\mu\alpha}(k; z)V^{f*}_{\mu\alpha}(k)G^{(0)}_{f\mu\mu}(z)V^f_{\mu\alpha}(k')G_{0\mu\alpha}(k; z),\end{aligned} \tag{2.8}$$

where $G_{0\mu\alpha}(k; z)$ is the Green function of noninteracting electrons (in accordance with the definition given above, the variables $k$ and $\varepsilon$ are identical). Because of the symmetry properties, the function $G^{(0)}_{f\mu\mu}(z)$ can have only diagonal components. The Green function of an impurity state $G^{(0)}_{f\mu\mu}(z)$ involves all the interactions induced by the Hubbard repulsion $U$. Near the Fermi level, the multiparticle resonance Green function $\hat{G}^0_{f_r\mu}(z)$ must be inserted in Eq. (2.8). Thus, the DOS of multiparticle excitations at the Fermi level takes the form

$$\rho_B(\varepsilon)-\rho_{0B}(\varepsilon) = -\frac{1}{\pi}A_{\rho}\mathrm{ImSp}\hat{G}^0_{f_r\mu}(\varepsilon), \quad \varepsilon > 0,$$

where $A_{\rho} \propto \gamma_B\rho_{0B}$, $\gamma_B \equiv \sum_{\mu}\gamma_{B\mu}$, and $\gamma_{B\mu} \equiv \left|V^f_{\mu}\right|^2\rho_{0B}$.

With the foregoing taken into account, the complete solution of the scattering problem requires determination of the main interaction and calculation of the Green function $\hat{G}^0_{f_r\mu}(z)$.

## 3. THE INTERACTION HAMILTONIAN AND THE NFL STATE

**3.1.** To derive the effective interaction between the deep $f$-doublet and the conduction electrons, we suppose that for relatively large values of the Hubbard repulsion, the ground state configuration of the ion $U^{4+}$ is the singly occupied $\Gamma_3$ doublet with the electron configurations $|1; +1\rangle$, $|1; -1\rangle$ and the energy $E_f$. Taking virtual transitions into the excited states with the energies $E_2 = 2E_f + U$ into account and using either the projection operator techniques or the Schriffer–Wolff transformation for the Hamiltonian $H_U^f + H_0^f + H_h$, we obtain the standard expression

$$H_{\text{int}} = \sum_{\mu\mu'\alpha\alpha'} \iint d\varepsilon d\varepsilon' \rho_{0B}(\varepsilon)\rho_{0B}(\varepsilon') \times V_{\mu\mu'}(\varepsilon, \varepsilon') a^{+}_{B\alpha\mu}(\varepsilon) a_{B\alpha'\mu'}(\varepsilon') f^{+}_{\mu} f_{\mu'}. \tag{3.1}$$

As $U_{\mu\mu'} \longrightarrow \infty$, the matrix elements in Eq. (3.1) become

$$V_{\mu\mu'}(\varepsilon, \varepsilon') \sim \frac{V^{f*}_{\mu\alpha}(\varepsilon) V^{f}_{\mu'}(\varepsilon')}{\varepsilon_f}, \quad \varepsilon_F - E_f \equiv \varepsilon_f.$$

The doubly degenerate *f*-level containing one electron can be conveniently described in terms of the pseudospin variable $\hat{\tau}_f$. The projections of the pseudospin operator $\hat{\tau}_f$ on the coordinate axes coincide with the components of the quadrupole moment tensor. The projection $\hat{\tau}^z_f \sim Q_{zz}$ on the *z* axis has two values corresponding to the occupation of the different orbitals of the doublet. The operator $\hat{\tau}^x_f \propto J_x^2 - J_y^2$ inverts the pseudospin, and we can therefore write

$$\hat{\tau}^i_f = \sum_{\mu\mu' = \pm 1} f^{+}_{\mu} \sigma^i_{\mu\mu'} f_{\mu'},$$

where $\sigma^i$ are the Pauli matrices.

The index $\alpha = \pm$ is magnetic, and therefore, it cannot change under the scattering by the electric quadrupole moment of the impurity nonmagnetic $\Gamma_3$ doublet described by (2.4). In other words, for Hamiltonian (2.4) to possess the time reversal property, the quantum number $\alpha$ must be conserved during the scattering. The scattering processes change only the states belonging to the same group ($\Gamma_8^{(+)}$ or $\Gamma_8^{(-)}$), and these states form a representation for the pseudospin $\hat{\tau}_f = 1/2$.

The time reversal symmetry therefore guarantees the transfer from Hamiltonian (2.4) to the two-channel quadrupole exchange Hamiltonian with the channel index $\alpha$,

$$\begin{gathered} H_{\text{int}} = \sum_{\mu\mu'\alpha} \sum_{i = x, y, z} \iint d\varepsilon d\varepsilon' \rho_{0B}(\varepsilon)\rho_{0B}(\varepsilon') \\ \times V_i(\varepsilon, \varepsilon') a^{+}_{B\alpha\mu}(\varepsilon) \sigma^i_{\mu\mu'} a_{B\alpha\mu'}(\varepsilon') \hat{\tau}^i_f, \\ V_i(\varepsilon, \varepsilon') \sigma^i_{\mu\mu'} \equiv V_{\mu\mu'}(\varepsilon, \varepsilon'). \end{gathered} \tag{3.2}$$

Because the hybridization matrix elements are complex in general, Eqs. (3.1) and (3.2) contain the term involving $\hat{\tau}^y_f$ along with the term involving $\hat{\tau}^x_f$. We are interested in the case where the dominant effect of the interaction is the generation of a multiparticle resonance at the Fermi level. The Green function corresponding to this resonance can be calculated using the bosonization method by reducing the Hamiltonian v to the resonance-level model proposed in [14]. To reduce the Hamiltonian $H_0 = H_{00} + H_{\text{int}} + H_\Delta$ with the two-channel exchange in Eq. (3.2) to the resonance-level model, it is convenient to rewrite $H_0$ as

$$\begin{aligned} H_0 = {} & i v_F \sum_{\mu\mu'\alpha} \int_{-\infty}^{+\infty} \psi^{+}_{\mu\alpha}(x) \partial_x \psi_{\mu\alpha}(x) \\ & + \frac{1}{2} \sum_{\mu\alpha} \sum_{i = x, y, z} V_i \psi^{+}_{\mu\alpha}(0) \hat{\sigma}^i_{\mu\mu'} \psi_{\mu'\alpha}(0) \hat{\tau}^i_f + \Delta \hat{\tau}^z_f, \end{aligned} \tag{3.3}$$

where

$$\psi_{\mu\alpha}(x) = \int_{-\infty}^{+\infty} dk e^{ikx} a_{B\mu\alpha}(k)$$

and

$$\psi_{\mu\alpha}(0) = \psi_{\mu\alpha}(x = 0).$$

The bosonic representation of the fermion fields $\psi_{\mu\alpha}(x)$ takes the form

$$\begin{gathered} \psi_{\mu\alpha}(x) = \hat{\eta}_{\mu\alpha} \frac{e^{-i\Phi_{\mu\alpha}(x)}}{(2\pi a)^{1/2}}, \quad \hat{\eta}^2_{\mu\alpha} = 1, \\ \Phi_{\mu\alpha}(x) = (\pi)^{1/2} \left[ \int_{-\infty}^{x} dx' P_{\mu\alpha}(x') + \varphi_{\mu\alpha}(x) \right], \end{gathered} \tag{3.4}$$

where $\varphi_{\mu\alpha}(x)$ is the boson field, $P_{\mu\alpha}(x')$ is the canonically conjugate momentum, $[\varphi_{\mu\alpha}(x), P_{\mu'\alpha'}(x')] = i\delta(x - x')\delta_{\mu\mu'}\delta_{\alpha\alpha'}$, and $a$ is the lattice constant. The operators $\hat{\eta}_{\mu\alpha}$ ensure the anticommutation relations between different species of fermions. The boson fields $\varphi_{\mu\alpha}(x)$ and $P_{\mu\alpha}(x)$ can be rewritten in terms of the collective variables that are introduced by means of the canonical transformation of $\varphi_{\mu\alpha}(x)$ and $P_{\mu\alpha}(x)$:

$$\begin{gathered} \varphi_{c, f} = \frac{1}{2}[(\varphi_{11} + \varphi_{12}) \pm (\varphi_{21} + \varphi_{22})], \\ \varphi_{s, (sf)} = \frac{1}{2}[(\varphi_{11} - \varphi_{12}) \pm (\varphi_{21} - \varphi_{22})]. \end{gathered} \tag{3.5}$$

Similar expressions can be written for the conjugate fields $P_{\mu\alpha}(x)$, $\mu, \alpha = 1, 2$. The Fourier components of the boson fields $k^{1/2}\varphi_l(k)$ correspond to the charge (*c*), flavor (*f*), pseudospin (*s*), and mixed (flavor-quadrupole, sf) density operators $\rho_l(k)$. The flavor is generated by the channel index $\alpha$.

In terms of the collective bosonic variables, the spinless fermion collective fields are given by

$$\psi_l(x) = \frac{e^{-i\Phi_l(x)}}{(2\pi a)^{1/2}}, \quad l = c, f, s, (sf). \tag{3.6}$$

The Hamiltonian $H_0$ can be represented as a sum of four terms corresponding to the four spinless fermion collective channels. The charge and flavor channels are not coupled to the impurity pseudospin. The other channels give the following terms in the Hamiltonian $H_0 = H_{00} + H_{\rm int} + H_\Delta$:

$$\begin{aligned} H_{00} &= i v_F \sum_{l = s, (sf)} \int_{-\infty}^{+\infty} dx \psi_l^+(x) \partial_x \psi_l(x), \\ H_{\rm int} + H_\Delta &= \frac{V_x}{(2a\pi)^{1/2}} [\psi_{sf}^+(0) + \psi_{sf}(0)] \hat{\tau}_f^x \\ &+ \tilde{V}_z \psi_s^+(0) \psi_s(0) \hat{\tau}_f^z + \Delta \hat{\tau}_f^z, \quad \tilde{V}_z \equiv 2(V_z - \pi v_F). \end{aligned} \tag{3.7}$$

The Hamiltonian in Eq. (3.7) corresponds to the resonance-level model that yields a multiparticle resonance (the $f_r$ level) at the Fermi level. The $f_r$ level can be described in terms of the fermion operators $d^+$ and $d$ coupled to the pseudospin operator $\hat{\tau}_f$ via the Majorana representation: $d^+ = \hat{\tau}_f^+ \hat{\eta}$, $\hat{\tau}_f^z = d^+d_-(1/2)$, where $\hat{\eta}$ is the Majorana (real) fermion operator such that $\hat{\eta}^2 = 1$. The Green function $\hat{G}_{f_r}^0(z)$ of the resonance level contains the anomalous components $\propto\langle dd\rangle$ and $\propto\langle d^+d^+\rangle$ in addition to the normal components $\propto\langle dd^+\rangle$ because the number of fermions is not conserved in the models described by Eq. (3.7).

**3.2.** It is known [15, 16] that the two-channel model described by Eqs. (3.3) and (3.7) has two regions with essentially different physical properties depending on the relation between $T_K$ and $\Delta$, where $T_K$ is the exponential Kondo temperature.

We consider the region of the parameters where the Kondo physics plays the key role. This case is referred to as the "Kondo regime" in what follows. It occurs under the condition

$$T_K \gg \Delta. \tag{3.8}$$

In this case, the model described by (3.7) renormalizes to the strong coupling limit [15, 16]. In this limit, the quantities $\Gamma_K = \pi\rho_{0B}V_x^2$ and $\Delta$ renormalize to $T_K$ and $\Delta^2/T_K$, respectively. The fixed point lies on the line $\tilde{V}_z = 0$ [8] (the Emery–Kivelson line), and the screening interaction is not essential for small energies. The quantity $T_K$ is defined on the Emery–Kivelson line and depends on $V_x$ only. For this reason, the parameters $T_K$ and $\Delta$ are independent. The NFL state is generated by the impurity degrees of freedom that are not hybridized with the conduction electrons [14, 17]. Near the Fermi level at $T = 0$, the Green function becomes

$$\hat{G}_{f_r}^{(0)}(z) = \pm\left[\frac{\hat{\sigma}_0 - \hat{\sigma}_x}{z - \Sigma_K(z)} + \frac{\hat{\sigma}_0 + \hat{\sigma}_x}{z}\right], \tag{3.9}$$

where $\Sigma_K(z)$ is the self-energy part determined by the hybridization term in Eq. (3.7) and $\pm$ corresponds to $\mathrm{Re}(z) \gtrless 0$. As usual, the exponential pole at $|z| \propto T_K$ in the first term of $\hat{G}_{f_r}^0(z)$ has the exponentially small residue $Z_K \propto \exp(-\varepsilon_{\Gamma_3}/\gamma_B)$, $T_K \propto \varepsilon_F Z_K$.

On the other hand, under the conditions

$$T_K \ll \Delta, \quad \tilde{V}_z \gg V_{x,y}, \tag{3.10}$$

the model does not renormalize to the strong coupling limit (or equivalently, to the fixed point at $\tilde{V}_z = 0$) for low temperatures because of a very weak renormalization of $\Delta$ [15]. In this case, the NFL state is generated by the screening interaction in Eq. (3.7) and by the nonhybridized impurity degrees of freedom. This mechanism is referred to as the "X-ray-edge regime" in what follows. In this case, the hybridization occurring in the *sf*-channel can be treated as a perturbation of the ground state obtained at $V_x = 0$. At $V_x = 0$, the problem is solved exactly. To obtain the Green function $\hat{G}_{f_r}^0(z)$ at $V_x = 0$, we use the technique that was previously applied to the well-known problem of the X-ray absorption in metals.

We first diagonalize the Hamiltonian $H_{00}^s + H_s + H_\Delta$ in (3.7) at $V_{x,y} = 0$. For this, we introduce the boson operators $b_{sk} = k^{-1/2}\rho_s(k)$ and $b_{sk}^+ = k^{-1/2}\rho_s(-k)$, where

$$\begin{aligned} \rho_s(k) &= \frac{1}{N^{1/2}} \sum_{q=0}^{k_D - k} \psi_s^+(q) \psi_s(q+k), \\ \rho_s(-k) &= \frac{1}{N^{1/2}} \sum_{q=k}^{k_D} \psi_s^+(q) \psi_s(q-k), \quad k \ge 0, \end{aligned} \tag{3.11}$$

are density operators, $\psi_s(k)$ are Fourier components of the fields $\psi_s(x)$, and the cutoff occurs at $k_D \sim a^{-1}$. Using the operators $b_{sk}$ and $b_{sk}^+$ we write the Hamiltonian as

$$\begin{aligned} H_{00}^s + H_s + \Delta\hat{\tau}_f^z &= v_F \sum_{k>0} k b_{sk}^+ b_{sk} + \tilde{V}_z \left(d^+d - \frac{1}{2}\right) \\ &\times \sum_{k>0} \left(\frac{k}{N}\right)^{1/2} (b_{sk}^+ + b_{sk}) + \Delta\hat{\tau}_f^z. \end{aligned} \tag{3.12}$$

This is diagonalized by the canonical transformation

$$U_B = \exp\left(\tilde{V}_z \rho_{0B}\left(d^+d - \frac{1}{2}\right) \sum_{k>0} (kN)^{-1/2} (b_{sk} - b_{sk}^+)\right).$$

Under this operation, the Hamiltonian $H_{00}^{s} + H_s + \Delta\hat{\tau}_f^z$ is transformed to

$$\tilde{H}_s = v_F \sum_{k>0} k \tilde{b}_{sk}^{+} \tilde{b}_{sk} + \tilde{\Delta}\left(\tilde{d}^{+}\tilde{d} - \frac{1}{2}\right), \qquad (3.13)$$

where

$$\tilde{d}^{+} = U_B d^{+} U_B^{-1} \equiv U_{0B} d^{+},$$

$$\tilde{b}_{sk}^{+} = U_B b_{sk}^{+} U_B^{-1} = b_{sk}^{+} + \frac{\rho_{0B}\tilde{V}_z}{(kN)^{1/2}} d^{+} d,$$

$$U_{0B} = \exp\left(\tilde{V}_z \rho_{0B} \sum_{k>0} \frac{b_{sk} - b_{sk}^{+}}{(kN)^{1/2}}\right),$$

$\tilde{\Delta} = \Delta - \varepsilon_U$, and $\varepsilon_U = \tilde{V}_z^2 \rho_{0B}$ is the “polaron shift”.

Equation (3.13) allows us to find the Green function of the resonance level,

$$\hat{G}_{f_r}^{(0)}(t) = \hat{G}_{f_r}^{(00)}(t) \langle U_{0B}^{+}(t) U_{0B}(0) \rangle_D, \qquad (3.14)$$

where $U_{0B}(t)$ is derived from $U_{0B}(0)$ by the substitution $b_{sk} \longrightarrow b_{sk} e^{i\varepsilon_k t}$. In Eq. (3.14), $\langle \ldots \rangle_D$ denotes averaging over the states of the diagonalized Hamiltonian $H_{00}^{s} + H_s$ and $\hat{G}_{f_r}^{(00)}(t)$ is the Green function with the $s$-channel interaction disregarded. The averaging is performed in the standard way using the relations

$$e^{\hat{A}} e^{\hat{B}} = e^{\hat{A} + \hat{B} + (1/2)[\hat{A}, \hat{B}]},$$

$$\langle e^{[F(b^{+}, b)]} \rangle = e^{(1/2)\langle F^2(b^{+}, b) \rangle},$$

where $F$ is an arbitrary linear combination of boson operators. As a result, we find that at large times $\varepsilon_F t \gg 1$, the function in Eq. (3.14) is given by

$$\hat{G}_{f_r}^{(0)}(t) \sim \hat{G}_{f_r}^{(00)}(t) t^{-\alpha_s}, \qquad (3.15)$$

where $\alpha_s = (\delta_s/\pi)^2$ and $\delta_s$ is the phase shift for the scattering described by $H_s$ in the pseudospin channel.

At $V_{x,y} = 0$, we use Eq. (3.15) with $\hat{G}_{f_r}^{(00)}(t) \propto e^{-i\tilde{\Delta}t}$ to obtain the known expression for the Green function

$$G_{f_r}^{(0)}(z) = \frac{A_{(\pm)}\Gamma(1-\alpha_s)}{z - \tilde{\Delta}} \left(\frac{z - \tilde{\Delta}}{W}\right)^{\alpha_s}, \qquad (3.16)$$

where $A_{(+)} = -1$ and $A_{(-)} = (-1)^{-\alpha_s}$ for $\mathrm{Re}(z - \tilde{\Delta}) \gtrless 0$, respectively, $\Gamma(x)$ is the gamma function, and $W$ is the cutoff parameter of the order of the conduction band width.

We next recall (e.g., from [23]) that including the hybridization $V_x$ as a perturbation in the X-ray-edge Hamiltonian, we recover the previous X-ray-edge results with the energy shifted as $i\omega \longrightarrow i\omega + i\Gamma_K \operatorname{sgn}\omega$, $\Gamma_K = \pi\rho_{0F}V_x^2$, in the resonance level Green function $\hat{G}_{f_r}$. Within the framework of two-channel model (3.5), the width due to the hybridization appears only for half the impurity degrees of freedom $\tau_{\Gamma_3}^{x}$ hybridized with the conduction *sf*-channel.

The same result can be obtained by writing Hamiltonian (3.5) in terms of the hybridized states and then considering the screening interaction for these states. One can readily show that additional interactions induced by the transition to the new basis are proportional to $\tilde{V}_z (V_x/W)$ and are therefore much smaller than the screening interaction. In the new basis, Hamiltonian (3.7) is reduced under condition (3.10) to a Hamiltonian of the X-ray-edge type. In the present case, the hybridization in Eq. (3.7) gives the level width related to half the degrees of freedom of the impurities hybridized with the conduction electrons.

Using Eqs. (3.14) and (3.15) with

$$\hat{G}_{f_r}^{(00)} \sim (\hat{\sigma}_0 - \hat{\sigma}_x) e^{-i(\tilde{\Delta} - i\Gamma_K)t} + (\hat{\sigma}_0 + \hat{\sigma}_x) e^{-i\tilde{\Delta}t},$$

we thus obtain $\hat{G}_{f_r}^{(0)}(z)$ in the energy representation,

$$\hat{G}_{f_r}^{(0)}(z) = A_{(\pm)}\Gamma(1-\alpha_s) \times \left[\frac{\hat{\sigma}_0 - \hat{\sigma}_x}{z - \tilde{\Delta} + i\Gamma_K}\left(\frac{z - \tilde{\Delta} + i\Gamma_K}{W}\right)^{\alpha_s} + \frac{\hat{\sigma}_0 + \hat{\sigma}_x}{z - \tilde{\Delta}}\left(\frac{z - \tilde{\Delta}}{W}\right)^{\alpha_s}\right], \qquad (3.17)$$

where

$$\Gamma_K \sim W \frac{\gamma_{B1}\gamma_{B2}}{\varepsilon_f^2}.$$

Because we calculate the retarded Green function in Eq. (3.17), we must have $\mathrm{Im}z < 0$. If the radial parts of the wave functions entering the matrix elements $V_\mu^f$ are independent of $\mu$, we readily obtain

$$\gamma_{B1} \sim \gamma_{B2} = \gamma_B, \quad \Gamma_K \sim W\frac{\gamma_B^2}{\varepsilon_f^2}. \qquad (3.18)$$

The power-law dependence occurs in Eq. (3.17) under conditions (3.10).

It follows from (3.17) that the multiparticle NFL resonance at the Fermi level is generated by the mixed flavor-quadrupole (*sf*) mode. The interactions in the pseudospin channels having the screening character lead to the effective broadening of the resonance level. The second term in Eq. (3.17) is due to the impurity degrees of freedom that are not hybridized with the conduction electrons.

In concluding this section, we write the expression for the DOS $\rho_B(\varepsilon)$ near the Fermi level. The multiparticle resonances at the Fermi level are described by the Green functions in Eqs. (3.9), (3.17), and (3.16). These Green functions must be inserted in Eq. (2.8), after which $\rho_B(\varepsilon)$ is derived. In particular, inserting Eq. (3.17) in Eq. (2.8), we find the DOS in the X-ray-edge regime:

$$\rho_B(\varepsilon) - \rho_{0B}(\varepsilon) = -\frac{1}{\pi} A_\rho \mathrm{Im}\mathrm{Sp}\hat{G}^{(0)}_{f_r}(\varepsilon)$$

$$= A_\rho \sum_{i = 1, 2} \frac{\sin\left[(1-\alpha_s)\arctan\dfrac{\Gamma_i}{\varepsilon - \tilde{\Delta}}\right]}{W^{\alpha_s}[(\varepsilon - \tilde{\Delta})^2 + \Gamma_i^2]^{(1-\alpha_s)/2}}, \tag{3.19}$$

$$\varepsilon > 0,$$

where $A_\rho \sim \gamma_B \rho_{0B}$. The widths $\Gamma_1 = \delta \longrightarrow 0^+$ and $\Gamma_2 = \Gamma_K$ correspond to the two contributions to the Green function $\hat{G}^{(0)}_{f_r}$ in Eq. (3.17).

In the Kondo regime, the DOS is determined by function (3.9).

## 4. THE FL RESONANCES NEAR THE FERMI LEVEL

**4.1.** The scattering of the multiparticle excitations due to the term $H_{sc}$ results in simple poles near the Fermi level in the complete Green function $\hat{G}_{f\mu}(z)$ in Eq. (2.6). The poles correspond to new Fermi-liquid resonances. The positions $z_r^{(\pm)} = \varepsilon_r^{(\pm)} - i\gamma_r^{(\pm)}$ of the poles are determined by the equation

$$D_\mu^{AB}(z_r^{(\pm)}) = 1 - \Sigma^{sc}_{A\mu}(z_r^{(\pm)}) W_{B\mu}(z_r^{(\pm)}) = 0. \tag{4.1}$$

Because this equation is the same for all terms of the matrix $\hat{D}_\mu^{AB}$, the matrix indices are omitted in Eq. (4.1).

The expression for the Green function (2.6) near the FL resonance with the energy $z_{r\mu}$ becomes

$$\hat{G}_{f\mu}(\varepsilon) = F_r^{-1} \frac{z_{r\mu}\hat{G}^{(0)}_{f_r}(z_{r\mu})}{\varepsilon - \tilde{\Delta} - z_{r\mu}}, \tag{4.2}$$

where we expanded the denominator in Eq. (2.6) near the resonance energy as $D(\varepsilon) = D'(z_{r\mu})(\varepsilon - \tilde{\Delta} - z_{r\mu})$, where $D'(z_{r\mu}) \approx F_r/z_{r\mu}$ (with the indices of the denominator omitted at the moment) and $F_r$ is a function of the parameters of the order of unity. The energy dependences of $\Sigma^{hsc}_{\nu\mu}(z)$ in (2.6) are determined by the DOS $\rho_\nu(\varepsilon)$. In the model under consideration, the function $\Sigma^{sc}_{A\mu}(z)$ has no features at the Fermi level, which allows us to write

$$\mathrm{Re}\Sigma^{sc}_{A\mu}(0) \approx \rho_{0A}(0), \quad \mathrm{Im}\Sigma^{sc}_{A\mu}(0) = 0. \tag{4.3}$$

The self-energy functions $\Sigma^{sc}_{B\mu}(z)$ have the features corresponding to the NFL peaks in the DOS $\rho_B(\varepsilon)$.

In the Kondo regime, the main singular term appears in $\Sigma^{sc}_{B\mu}(z)$ because of the $\delta$-like contribution to the spectral function induced by the second term in the Green function (3.9) as $z \longrightarrow \varepsilon + i0^+$. In other words, this singular term is due to the impurity degrees of freedom that are not hybridized with the conduction electrons. The self-energy $\Sigma^{sc}_{B\mu}(z)$ takes the form

$$\Sigma^{sc}_{B\mu}(z) \approx \mathrm{const} + V^f_\mu T^{B*}_\mu \frac{\gamma_B \rho_{0B}}{z}. \tag{4.4}$$

In the X-ray-edge regime, using the density of states in Eq. (3.19), we obtain the contribution of the resonance levels to the self-energy function at zero temperature,

$$\Sigma^{sc}_{B\mu}(z) \approx A_{B\mu}(\gamma_B \gamma_{B\mu} \rho_{0B}) \left(\frac{T^{B*}_\mu}{V^{f*}_\mu}\right) \times \left(\frac{W}{z - \tilde{\Delta} + i\Gamma_K}\right)^{1-\alpha_s} (-1)^{\alpha_s - 1}, \tag{4.5}$$

where $|A_{B\mu}| \sim 1$.

In the Kondo regime, inserting (3.9) and (4.4) in (4.1) and taking the most singular term $\propto 1/z^3$ in $W_{B\mu}(z)$ into account, we readily obtain two resonances above and below the Fermi level that occur due to the scattering of the nonhybridized impurity degrees of freedom. The energies of these resonances are determined by

$$\frac{|z_r^{(\pm)}|}{W} = A_r(\gamma_{A\mu}\rho_{0A})^{1/3}(\gamma_B \rho_{0B}), \tag{4.6}$$

where $A_r \sim 1$ and $\gamma_{A\mu} = |T^B_\mu|^2 \rho_{0A}$. The resonance width above the Fermi level is much smaller than the resonance width below the Fermi level. The former width is determined by the terms in Eq. (4.1), which are much smaller than the leading singular term $\propto 1/z^3$. Therefore, the pseudogap exists near the Fermi level for $|\varepsilon^{(-)}| \gg \gamma_r^{(-)}$ and for $|\varepsilon^{(-)}| \ll \gamma_r^{(-)}$.

In addition, Eq. (4.1) has two solutions above and below the Fermi level with $|z + i\gamma_K| \ll \gamma_K$. For this reason, the shape and the width of the Kondo peak change weakly at the Fermi level. In particular, the width of the Kondo peak has a small additional term $\sim T_K(\gamma_{A\mu}\rho_{0A})(\gamma_B\rho_{0B}) \ll T_K$ due to the scattering.

The qualitative picture of the DOS in the Kondo regime near the Fermi level is shown in Fig. 2. We see that the FL resonances generate both the additional energy scale $\gamma_r \ll T_K$ and the pseudogap near the Fermi level.

We thus obtained the essential result that the scattering of the nonhybridized impurity degrees of freedom

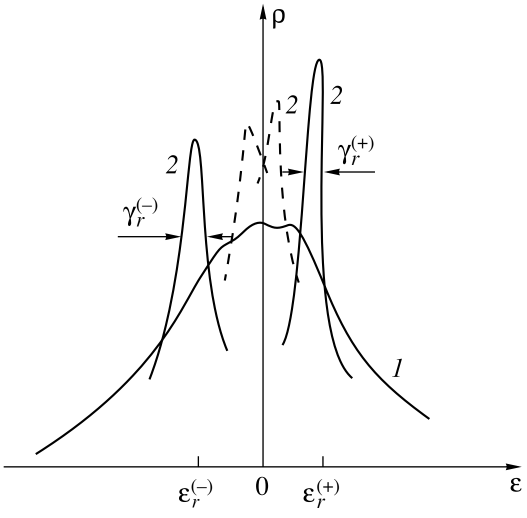

**Fig. 2.** The Kondo resonance (curve *1*) and new FL resonances (curves *2*) show the respective FL resonances with $|\varepsilon_r^{\pm}| \gg \gamma_r^{\pm}$ and $|\varepsilon_r^{\pm}| \ll \gamma_r^{\pm}$.

by the electron states of the narrow band leads to the existence of new resonances near the Fermi level.

According to the experimental data [2], there exists a concentration region where the Kondo energy $T_K$ increases exponentially with decreasing the impurity concentration. At the same time, the hybridization matrix elements and, consequently, the widths $\gamma_B$ and $\gamma_{A\mu}$ remain approximately constant in this region. We can therefore expect that the condition $|z_r| \ll T_K$ is satisfied at a sufficiently low concentration of the impurity atoms.

**4.2.** Using expressions (2.7) and (3.17), it is easy to verify that in the X-ray-edge regime, Eq. (4.1) possesses solutions of two types with their energies satisfying the respective conditions

$$|z_r^{(\pm)}| \ll \Gamma_K, \quad \text{the narrow resonances,}$$

$$|z_r^{(\pm)} + i\Gamma_K| \ll \Gamma_K, \quad \text{the “wide” resonances.}$$

For simplicity, we here used the condition

$$\tilde{\Delta} \ll \Gamma_K, z_r^{(\pm)}.$$

The signs “±” correspond to the resonances above and below the Fermi level. For $|\varepsilon_r^{(\pm)}| \ll \gamma_r^{(\pm)}, |\gamma_r^{(\pm)} - \Gamma_K|$, the widths of the FL resonances are determined by

$$\begin{aligned}\frac{\gamma_r^{(\pm)}}{W} &= A_1(\gamma_{A\mu}\gamma_{B\mu}\rho_{0B}^2)^{1/(1-\alpha_s)} \\ &\times (\gamma_\beta\rho_{0B})^{2/(1-\alpha_s)}\left(\frac{\varepsilon_f}{\gamma_{B\mu}}\right)^4, \quad \gamma_r^{(\pm)} \ll \Gamma_K,\end{aligned} \tag{4.7}$$

$$\begin{aligned}\frac{|\gamma_r^{(\pm)} - \Gamma_K|}{W} &= A_2(\gamma_{A\mu}\gamma_{B\mu}\rho_{0B}^2)^{1/3(1-\alpha_s)} \\ &\times (\gamma_\beta\rho_{0B})^{2/3(1-\alpha_s)}, \quad |\gamma_r^{(\pm)} - \Gamma_K| \ll \Gamma_K,\end{aligned} \tag{4.8}$$

where $A_{1,2} \sim 1$. In this case, the FL resonances merge into a single weakly split resonance at the Fermi level (Fig. 3b). For $|\varepsilon_r^{(\pm)}| \gg \gamma_r^{(\pm)}, |\gamma_r^{(\pm)} - \Gamma_K|$, the energies $|\varepsilon_r^{(\pm)}|$ are determined by the expressions in the right-hand sides of Eqs. (4.7) and (4.8) and by $\gamma_r^{(\pm)} = |\varepsilon_r^{(\pm)}|\sin\varphi$ with $\varphi \ll 1$. In this case, pairs of the FL resonances appear above and below the Fermi level (Fig. 3a). Pairs of the FL resonances can exist because the Green function $\hat{G}_{f\mu}(z)$ has two branches above and below the Fermi level. For $|\varepsilon_r^{(\pm)}| \gg \gamma_r^{(\pm)}$, there are well-determined pseudogaps near the Fermi level in the case of the narrow resonances.

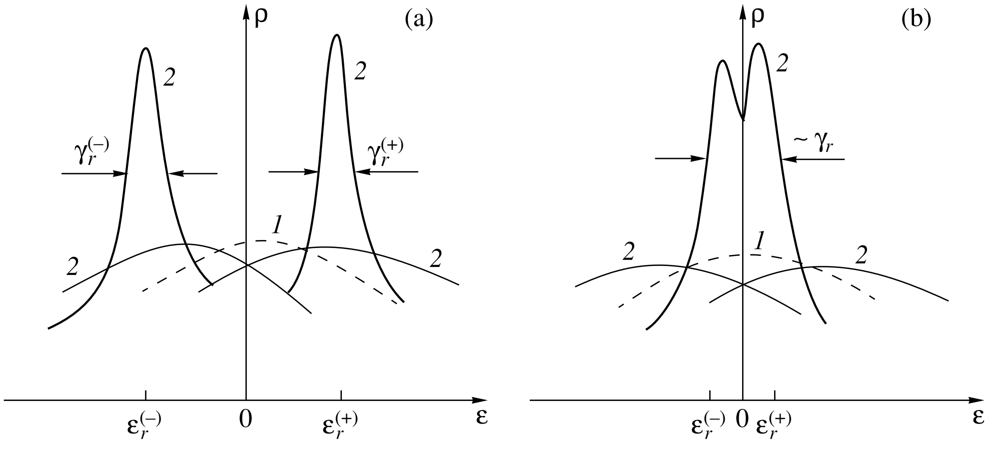

**Fig. 3.** The NFL (curves *1*) and FL (curves *2*) resonances in the X-ray-edge regime: (a) the narrow resonances for $|\varepsilon_r^{\pm}| \gg \gamma_r^{\pm}$; (b) these resonances for $|\varepsilon_r^{\pm}| \ll \gamma_r^{\pm}$.

Two types of FL resonances correspond to the existence of the hybridized and nonhybridized impurity degrees of freedom. In particular, the narrow resonances, which determine a new small energy scale near the Fermi level, are generated by the interband scattering of the nonhybridized impurity degrees of freedom. In other words, the narrow resonances result from broadening and displacement of the zero-width term in the spectral function $\hat{G}^{(0)}_{f_r}$ (see the second term in Eq. (3.17)) due to the interband scattering.

Equations (4.7) and (4.8) imply that the FL resonances exist for the deep level ($\varepsilon_f \gg \gamma_B$) under the condition

$$\frac{\gamma_{A\mu}}{\gamma_{B\mu}} \ll (\gamma_{B\mu}\rho_{0B})^{2-6\alpha_s}\left(\frac{W}{\varepsilon_f}\right)^{6(1-\alpha_s)}, \tag{4.9}$$

which is the same for the resonances of both types.

Condition (4.9) is satisfied for all values of $\alpha_s$ in the following cases. First, for $\gamma_{A\mu} \gg \gamma_B$ and sufficiently "shallow" $f$-levels such that

$$\gamma_B \ll \varepsilon_f \ll W\left(\frac{\gamma_B}{W}\right)^{(1-3\alpha_s)/3(1-\alpha_s)}, \tag{4.10}$$

and second, for $\gamma_{A\mu} \ll \gamma_B$ and $\varepsilon_f \sim W$.

On the other hand, the widths of the NFL resonance and, correspondingly, the characteristic binding energy of the collective states forming the NFL resonance can be estimated as

$$\varepsilon_K \sim \Gamma_K\left(\frac{\varepsilon_F}{\Gamma_K}\right)^{\alpha_s}. \tag{4.11}$$

This estimate is derived from the NFL DOS in Eq. (3.19). As $\alpha_s$ increases, the binding energy $\varepsilon_K$ also increases.

The FL resonance can appear if the collective states defined in Eqs. (3.6) and (3.11) decay. Taking the foregoing into account, we must bear in mind that the decay of collective states becomes more difficult as $\alpha_s$ increases. Therefore, the structure of the FL resonances near the Fermi level essentially depends on the magnitude of the parameter $\alpha_s$ that describes the scattering in the quadrupole (pseudospin) channel. From the imaginary part of Eq. (4.1), we readily find that the narrow resonances exist for $\alpha_s \le 3/5$. For $1/7 < \alpha_s \le 1/3$, the narrow resonances appear above and below the Fermi level. For $\alpha_s > 5/7$, FL resonances are absent.

In addition to condition (4.9), we thus find that the narrow FL resonances can exist when the pseudospin channel interaction $\tilde{V}_z$ is not very strong.

In the limiting case where $V_{xy} = 0$, the FL resonance exists above the Fermi level for $\alpha_s \le 1/3$. Its energy is determined by

$$\frac{\left|\gamma_r^{(\pm)}\right|}{W} \sim (\gamma_{A\mu}\rho_{0A})^{1/(1-\alpha_s)} \times (\gamma_B\rho_{0B})^{1/(1-\alpha_s)} \ll \gamma_B. \tag{4.12}$$

In the X-ray-edge regime, the narrow FL resonances provide peaks in the DOS with the widths much smaller than those of the NFL resonance (see Fig. 3). Thus, their existence allows us to obtain a new mechanism for the appearance of the small energy scale.

We also mention that as shown in [11], the model without the continuum in the impurity region does not give narrow FL resonances and therefore does not lead to the small energy scale. The "wide" resonances above and below the Fermi level and a local state above the Fermi level have been obtained in this model. Additional mechanisms are required for broadening local states.

At the same time, the existence of the narrow FL resonances leads to the appearance of pseudogaps near the Fermi level in the X-ray-edge regime. The pseudogap occurs under the Fermi level for a single narrow FL resonance at $1/3 < \alpha_s \le 3/5$. At $\alpha_s \le 1/3$ for the split FL resonances, the pseudogap also splits into two branches above and below the Fermi level. The pseudogaps are well determined for $|\varepsilon_r| \gg \gamma_r$. The minimum value of the DOS inside the pseudogaps is of the order of the magnitude of the "wide" resonances. The maximum widths of the pseudogaps are of the order $\left|\varepsilon_r^{(\pm)}\right|$ and are determined by the expression in the right-hand side of Eq. (4.7).

The conditions required for the appearance of pseudogaps are identical to those for the existence of the narrow FL resonances.

## 5. THE MIXED-VALENCE AND NEARLY INTEGER STATES

The criterion that enables us to choose between the two types of states involves the partial $f$-component $\rho_{f_r}(0)$ of the DOS at the Fermi level and the DOS $\rho_f(\varepsilon_{f\mu})$ at the deep level. For $\rho_{f_r}(0) \gg \rho_f(\varepsilon_{f\mu})$, the charged excitations play the key role at the Fermi level, while the opposite inequality means that their role is negligible. The former case corresponds to the mixed-valence state, and the latter case leads to the state with a nearly integer valence.

The Green function $G^{(0)}_{f\mu}(z)$ for $|z|$ close to the energy $\varepsilon_{f\mu}$ of the deep level can be represented as

$$G^{(0)}_{f\mu}(z) \approx \frac{Z_{f\mu}}{z - \varepsilon_{f\mu}}, \tag{5.1}$$

where $Z_{f\mu} \sim 1$ is the residue at the pole $z = \varepsilon_{f\mu}$. The energy $\varepsilon_{f\mu}$ renormalized by hybridization is determined by the equation

$$\varepsilon_{f\mu} = \varepsilon_f + \Sigma^{(0)}_{B\mu}(\varepsilon_{f\mu}) \equiv \bar{\varepsilon}_f + i\gamma_{f\mu}.$$

The maximum value of the DOS at the deep level can therefore be estimated as

$$\rho_f(\varepsilon_{f\mu}) \sim \rho_{0B}\left(\frac{\varepsilon_F}{\gamma_B}\right). \tag{5.2}$$

We now verify our criterion for the Kondo resonance. It is well known [19, 20] that in this case, the density of charged states is small at the Fermi level. Using the “resonance level” formalism, one can see this from the small residue $Z_K$ that determines the pole contribution to the Green function at $|z|$ close to the Fermi energy,

$$\begin{gathered} G_K(z) \approx \frac{Z_K}{z - E_K}, \\ \rho_K(0) \sim \frac{Z_K}{\gamma_K} \sim \rho_{0B} \ll \rho_f(\varepsilon_{f\mu}), \end{gathered} \tag{5.3}$$

where $E_K \sim i\gamma_K$ and $\gamma_K \sim T_K$. In accordance with our criterion, the inequality corresponds to a small contribution of the charged excitations at the Fermi level.

However, for new FL resonances with the widths $\gamma_r^{\pm}$ in Eq. (4.6), the following inequality holds:

$$\rho^{FL}_{f_r}(0) \sim \frac{W}{\gamma_r}\rho_{0B} \gg \rho_f(\varepsilon_{f\mu}). \tag{5.4}$$

Therefore, additional FL resonances lead to the existence of a mixed-valence state in the Kondo regime.

In the X-ray-edge regime, the NFL resonance is generated by the flavor-quadrupole and the quadrupole (pseudospin) modes that have a charge due to the quadrupole contribution. The component $\rho^{NFL}_{f_r}(0) = -(1/\pi)\mathrm{Im}\,\mathrm{Sp}\,\hat{G}^{(0)}_{f_r}(0)$ is then estimated as

$$\begin{aligned} \rho^{NFL}_{f_r}(0) &\sim \rho_{0B}\left(\frac{W}{\Gamma_K}\right)^{1-\alpha_s} \\ &\sim \rho_{0B}\left(\frac{\bar{\varepsilon}_f}{\gamma_B}\right)^{2(1-\alpha_s)}. \end{aligned} \tag{5.5}$$

For the narrow FL resonances, using expression (2.6) for the Green function $G_{f_r\mu}(z)$, we readily arrive at the estimate

$$\begin{aligned} \rho^{FL}_{f_r}(0) &= -\frac{1}{\pi}\sum_{\mu}\mathrm{Im}\,\mathrm{Sp}\,\hat{G}_{f\mu}(0) \\ &\sim \rho_{0B}\left(\frac{W}{\gamma_r}\right)^{(1-\alpha_s)}. \end{aligned} \tag{5.6}$$

Assuming $\bar{\varepsilon}_f \sim W$ and comparing (5.2) with (5.5) and (5.6), we find

$$\begin{gathered} \rho^{NFL}_{f_r}(0) \gg \rho_f(\varepsilon_{f\mu}) \ \text{ at } \ \alpha_s < \frac{1}{2}, \\ \rho^{NFL}_{f_r}(0) \ll \rho_f(\varepsilon_{f\mu}) \ \text{ at } \ \alpha_s > \frac{1}{2}, \end{gathered} \tag{5.7}$$

and also the inequality

$$\rho^{FL}_{f_r}(0) \gg \rho^{NFL}_{f_r}(0), \quad \rho_f(\varepsilon_{f\mu}), \tag{5.8}$$

which holds for all values of the parameters at which FL resonances exist.

It is interesting to note that under the conditions $\gamma_{A\mu} \gg \gamma_B$ and

$$\begin{aligned} W\left(\frac{\gamma_B}{W}\right)^{(1-2\alpha_s)/2(1-\alpha_s)} &\ll \varepsilon_f \\ \ll W\left(\frac{\gamma_B}{W}\right)^{(1-3\alpha_s)/3(1-\alpha_s)}&, \end{aligned} \tag{5.9}$$

the mixed-valence state and FL resonances exist simultaneously for all values of $\alpha_s$.

Inequalities (5.7)–(5.8) imply that, first, the state with a nearly integer valence can be realized only when FL resonances are absent and the parameter $\alpha_s$ is sufficiently large. Second, two types of the mixed-valence states are generated in our system.

The NFL mixed-valence state occurs for $\alpha_s < 1/2$ if FL resonances are absent.

In the limiting case where $V^{\Lambda}_{x,y} = 0$, the mixed-valence state exists only owing to the additional FL resonance.

The FL mixed-valence states are generated by the instability of the NFL state against the interband scattering. These states are formed under the same conditions that are necessary for the existence of FL resonances at the Fermi level. The type of the FL mixed-valence state depends on the type of the FL resonance (narrow or “wide”) that can be realized for a given set of parameters.

As shown above, narrow FL resonances exist for all values $\alpha_s < 1/2$. Thus, the main features of the FL mixed-valence state are the appearance of a small energy scale and the formation of pseudogaps.

The transitions between the NFL and FL mixed-valence states are characterized by changing the valence from one noninteger value to another. Taking the foregoing into account, we conclude that condition (4.9) alone is necessary for the transitions between two mixed-valence states.

When condition (4.9) is not satisfied, the direct transition between the NFL mixed-valence state and the state with a nearly integer valence occurs at $\alpha_s \approx 1/2$.

Apparently, the most realistic way to generate the transitions experimentally is to change the lattice parameter by doping [3]. This leads to changing the hybridization between conduction electrons and the $\Gamma_3$ level that enters the interaction matrix elements and the widths $\gamma_B$. We can thus obtain a series of transitions, which are considered in detail elsewhere.

## 6. CONCLUDING REMARKS

**6.1.** The above results allow us to understand the mechanisms of two important properties of HF NFL metals.

(1) The single-site two-channel Kondo effect and the mixed-valence state coexist because of additional FL resonances at/near the Fermi level. The scattering of the nonhybridized impurity degrees of freedom by the narrow *A*-band electrons generates these resonances. Therefore, two energy scales $T_K$ and $\gamma_r$ exist at the Fermi level. The FL resonance with the width $\gamma_r$ corresponds to the local mixed-valence state.

(2) There are two possible energy dependence types in a system with the two-channel quadrupole exchange interaction. In the Kondo regime ($T_K \gg \Delta$), one obtains the known universal energy dependences [14, 17, 22] because the Green function in Eq. (3.9) has a single energy scale $T_K$.

In the X-ray-edge regime ($T_K \ll \Delta$), nonuniversal power-law energy dependences must occur in accordance with the form of the Green function in Eqs. (3.16) and (3.17).

It follows from the experimental data [2] that the increase of the impurity concentration $x$ in the U-compounds results in (a) decreasing $T_K(x)$, (b) increasing the concentration of the impurity atoms by a noticeable value $\Delta$, and (c) increasing the anisotropy of the exchange parameters. Therefore, increasing the impurity concentration must enable crossing over from the Kondo regime with the universal energy dependences to the X-ray-edge regime with nonuniversal energy dependences.

As shown above, characteristic features of the NFL compounds with *f*-shell impurities are the different types of the mixed-valence states with the NFL and FL excitation spectra and the fact that the heavy-fermion-state type depends on the interaction parameter $\alpha_s$. In the other words, this parameter determines the role of the charge and spin excitations in the formation of heavy fermions.

Small energy scales and the pseudogaps are induced by the narrow FL resonances. Therefore, the instability of the NFL state provides a new physical mechanism for the small energy scale. Unlike in previous works [6, 21], this mechanism is especially appropriate for impurities with an unstable valence.

Thus, the instability of the NFL state induced by the interband scattering of multiparticle excitations considerably changes the mechanisms of the formation of heavy-fermion states.

**6.2.** We now briefly consider the features of the temperature dependences within the framework of the mechanism proposed in the present paper. The energy dependences of the Green functions (2.6), (3.9), (3.16), and (3.17) imply that new types of the temperature transitions (crossovers) occur in the system. When new FL resonances generated by scattering are not formed, a transition occurs from the universal temperature dependences of the physical quantities in the Kondo regime to nonuniversal power-law dependences in the X-ray-edge regime. The characteristic temperature of this crossover is $T_{c1} \sim \Delta$. In particular, the logarithmic dependence of the linear specific heat $C/T \propto \ln(T_K/T)$ must be transformed into the power-law dependence $C/T \propto T^{-1+\alpha_s}$. The former dependence was calculated in [14, 17] using expression (3.9) within the framework of the two-channel Kondo model. The power-law dependences follow from Eqs. (3.16) and (3.18) for the Green functions in the X-ray-edge regime. As mentioned in this section, the condition $T_K \ll \Delta$ can be realized at a relatively high concentration of the *f*-shell impurities. The power-law dependences of $C/T$ observed in $U_xY_{1-x}Pd_3$ at $x = 0.2$ in [7] can therefore be generated by the mechanism discussed here. We recall that historically, the alloys $U_xY_{1-x}Pd_3$ were the first systems where the NFL behavior induced by the two-channel quadrupole Kondo model was observed [3, 5].

In the two-channel quadrupole Kondo model, the magnetic susceptibility is known [2] to have the van Vleck contribution between the $\Gamma_3$ ground state and the first excited crystalline electric field. The van Vleck susceptibility is described by the temperature dependence $\chi \sim \chi_0 - \alpha(T/T_K)^{1/2}$. According to the experimental data [7], this dependence is also transformed into a power-law one as the impurity concentration increases.

The quadrupole susceptibility $\chi_Q$ has the logarithmic divergence $\propto \ln(T_K/T)$ in the Kondo regime. It is experimentally determined from the nonlinear magnetic susceptibility $\chi_3$ [24]. Correspondingly, $\chi_Q$ and $\chi_3$ must exhibit the same crossover as the specific heat.

We emphasize that the crossover discussed here corresponds to the transition between the state with a nearly integer valence and the mixed-valence state.

The existence of the FL resonances generated by the scattering of the NFL excitations results in crossovers between the FL and NFL temperature dependences within both the Kondo regime and the X-ray-edge regime. The characteristic temperatures of these crossovers are $T_{c2} \sim \gamma_r$, where $\gamma_r$ are the widths of the FL resonances determined in Eqs. (4.6), (4.7). We note that the low-temperature transition to the FL state usually occurs at $T \sim \Delta^2/T_K$ in the two-channel Kondo model [3, 25]. The maximum value of the linear specific heat is equal to $(C/T)^{\max} \sim T_K/\Delta^2$.

Within the framework of our mechanism, it must be $(C/T)^{\max} \sim \gamma_r^{-1}$ for $\gamma_r \gg \Delta^2/T_K$. It is possible that the additional small scale $\gamma_r$ enters the scaling dependences in the FL–NFL transition region. The appearance of a new small energy scale is observed in the low-temperature scaling law of resistivity in [24]. In the X-ray-edge regions, the crossover at $T \sim T_{c2}$ corresponds to the transition between the FL and NFL mixed-valence states.

The temperature transitions between FL mixed-valence states of the different origins were considered in [26].

**6.3.** The above results are obtained for single-ion NFL effects. We now show that these effects can also be considerable in “concentrated” systems.

The ground state of these systems significantly depends on the competition between the intersite interaction, i.e., the indirect exchange of the RKKY type for pseudospins, and the on-site Kondo scattering leading to the screening of the quadrupole impurity moment by conduction electrons. The characteristic energy for the two-channel on-site Kondo scattering is determined by expression (4.11). The characteristic energy scale of the RKKY interaction is

$$\varepsilon_{RKKY} \sim c_i\left(\frac{V_{ex}^2}{\varepsilon_F}\right) \sim c_i\Gamma_K, \tag{6.1}$$

where $c_i$ is the concentration of the interacting atoms. In concentrated systems, i.e., at $c_i \sim 1$, the energies $\varepsilon_K$ and $\varepsilon_{RKKY}$ are such that

$$\varepsilon_K \gg \varepsilon_{RKKY} \quad \text{for} \quad \alpha_s \neq 0. \tag{6.2}$$

This implies that single-ion NFL effects can be very important even when the two-channel impurities form a sublattice.

The analysis presented here enables us to qualitatively understand two important aspects of the problem for the concentrated systems: the dependence of the HF properties on doping and physical reasons that can satisfactorily explain a number of properties of the concentrated systems within the framework of the single-ion quadrupole Kondo model.

This work was supported by the Russian Foundation for Basic Research and the International Association INTAS (grant no. 97-11066).

## REFERENCES

1. H. V. Lohneysen, Physica B (Amsterdam) **206-207**, 101 (1995).
2. M. B. Maple, R. P. Dickey, J. Herrmann, *et al.*, J. Phys.: Condens. Matter **8**, 9773 (1996).
3. S. L. Seaman and M. B. Maple, Physica B (Amsterdam) **199-200**, 396 (1994).
4. D. L. Cox, Phys. Rev. Lett. **59**, 1240 (1987).
5. H. Amitsuka and T. Sakakibara, J. Phys. Soc. Jpn. **63**, 736 (1994).
6. M. Koga and D. L. Cox, Phys. Rev. Lett. **82**, 2575 (1999); A. Schiller, F. B. Anders, and D. L. Cox, Phys. Rev. Lett. **81**, 3235 (1998).
7. M. C. de Andrade, R. Chau, R. P. Dickey, *et al.*, Phys. Rev. Lett. **81**, 5620 (1998).
8. M. Fabrizio, A. O. Gogolin, and Ph. Nozieres, Phys. Rev. Lett. **74**, 4503 (1995).
9. L. A. Manakova, Pis’ma Zh. Éksp. Teor. Fiz. **67**, 1009 (1998) [JETP Lett. **67**, 1069 (1998)].
10. L. A. Manakova, Zh. Éksp. Teor. Fiz. **114**, 1466 (1998) [Sov. Phys. JETP **87**, 796 (1998)].
11. L. A. Manakova, Pis’ma Zh. Éksp. Teor. Fiz. **69**, 721 (1999) [JETP Lett. **69**, 772 (1999)].
12. K. A. Kikoin, Physica B (Amsterdam) **163**, 343 (1990).
13. A. M. Tsvelik, J. Phys.: Condens. Matter **2**, 2833 (1990).
14. V. J. Emery and S. Kivelson, Phys. Rev. B **46**, 10812 (1992).
15. J. L. Black, K. Vladar, and A. Zawadowski, Phys. Rev. B **26**, 1559 (1982).
16. K. Vladar and A. Zawadowski, Phys. Rev. B **28**, 1546 (1983).
17. A. M. Sengupta and A. Georges, Phys. Rev. B **49**, 1020 (1994).
18. K. D. Schotte and U. Schotte, Phys. Rev. **182**, 479 (1969); P. Nozieres and C. T. de Dominicis, Phys. Rev. **178**, 1097 (1969).
19. A. C. Hewson, *The Kondo Problem to Heavy Fermions* (Cambridge Univ. Press, Cambridge, 1993).
20. N. Kawakami and A. Okiji, Phys. Rev. B **42**, 2383 (1990).
21. Yu. Kagan and N. V. Prokof’ev, Zh. Éksp. Teor. Fiz. **93**, 366 (1987) [Sov. Phys. JETP **66**, 211 (1987)].
22. A. W. W. Ludwig and J. Affleck, Phys. Rev. Lett. **67**, 3160 (1991).
23. P. Schlottmann, Phys. Rev. B **22**, 622 (1980).
24. F. Aliev, S. Vieira, R. Villar, and V. V. Moshchalkov, J. Phys.: Condens. Matter **8**, 9807 (1996).
25. P. D. Sacramento and P. Schlottmann, Phys. Rev. B **43**, 13294 (1991).
26. K. A. Kikoin and S. Burkatovskii, J. Mosc. Phys. Soc. **3**, 139 (1993).

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 410–414.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 464–469.*
*Original Russian Text Copyright © 2001 by Peregudov.*

## SOLIDS
## Electronic Properties

# Excitation of Standing Electron Waves in Fast-Ion Tracks

**V. N. Peregudov***
*Institute of Nuclear Fusion, Kurchatov Institute Russian Research Centre,*
*pl. Akademika Kurchatova 1, Moscow, 123182 Russia*
**e-mail: vlad@dni.polyn.kiae.su.*

Received April 9, 2001

**Abstract**—The specific case of slow-electron diffraction in fast-ion tracks is considered. The excitation condition for standing waves of δ electrons is a strong screening of the Coulomb interaction, for which most of the δ electrons are knocked out in radial directions relative to the track axis. In that case, an appreciable number of δ electrons with a wavelength of the order of the interatomic distance experience multiple backscatterings between the atomic chains located near the track axis. The lifetime of standing electron waves is estimated from diffractometric measurements of the decrease in atomic density on the track axis. The mobility of crystal atoms, their displacements, and the expenditure of energy on deformation are also estimated. *© 2001 MAIK “Nauka/Interperiodica”.*

Intensive studies of ionic-irradiation effects on the physical properties of widely used materials have provided insight into the many physical phenomena that occur in fast-ion tracks, such as anomalous defect formation [1, 2], alloy amorphization [3, 4], the so-called anisotropic expansion at low temperatures, and creep in the directions normal to ion beams [5–8]. There are also several theoretical studies in which models of a thermal flash [9–12], Coulomb explosion [13–15], soft phonon modes [16], structural modification [17], and the local electric field [18] generated by δ electrons in tracks were considered. The wide use of semiconductor crystals as detectors of high-energy ions stimulates investigations into their structural degradation, which is accompanied by a degradation of electrophysical properties. The peculiar features of defect formation in silicon irradiated with xenon and krypton ions with energies of 5.6 and 0.21 GeV, respectively, were studied in [19] by two-crystal X-ray diffractometry. It follows from the experimental data that the atomic density decreases approximately by 10% in Kr-ion tracks at large radiation doses. Chelyadinskii *et al.* [19] attribute the appearance of a second peak in the plot of interplanar spacing $\Delta d$ against depth to the ions falling into “old” tracks and having path lengths that are larger than ordinary ones by 10%. More informative experimental data can be obtained by using currently available methods of structural diagnostics. These methods are based on three-crystal X-ray diffractometry, which makes it possible to separate diffuse scattering and to demonstrate the peculiar behavior of dilatations in fast-ion tracks. At high energies, the deceleration of ions by electrons is much larger than their deceleration by nuclei, and it determines the behavior of several physical processes in tracks. Large energy release into an electron subsystem gives rise to intense fluxes of δ electrons with an initially distinct angular anisotropy.

In this paper, we consider the peculiar behavior of collective electron excitation in crystals irradiated by fast ions. Our approach is based on a microscopic examination of the interaction between those crystal atoms that are simultaneously excited by a fast impinging ion. While crossing the crystal planes along the normal, such an ion simultaneously knocks δ electrons out of several atoms in each crystal plane located in the immediate vicinity of the track axis. For a screened Coulomb interaction with a screening radius of the order of the Bohr radius, the knocking of δ electrons out of atoms that are not in the immediate vicinity of the track axis may be disregarded. Below, we consider the simplest case where the track axis passes between two chains of crystal atoms. In this case, the number of interacting atoms closest to the track axis and excited simultaneously is $N = 2$, and the scattering system becomes simplest. The number of δ electrons emitted by each interacting atom per unit time and per unit solid angle in direction θ relative to the track axis is

$$\phi(z, \mu) = \frac{V_1}{\pi\varepsilon_m A N}\frac{dE_1}{dz}\frac{\mu}{(\mu^2 + \xi^2)^2}, \qquad (1)$$

where $V_1$ and $dE_1/dz$ are the ion velocity and the ion energy losses for electron excitation, respectively; $m = \cos\theta$ gives the emission direction of the δ electrons relative to the track axis; $\varepsilon_m = 4mE_1/M_1$ is the maximum energy of the knocked-out δ electrons at fast-ion energy $E_1$; $\xi^2 = \hbar^2/2mr_{sc}^2\varepsilon_m$, $r_{sc}$ is the screening radius; $m$ and $M_1$ are the electron and fast-ion masses, respectively; $N$ is the number of interacting atoms excited simultaneously; and $A$ is the normalization factor,

$$A = \ln(1 + \xi^{-2}) - (1 + \xi^2)^{-1}.$$

1063-7761/01/9302-0410$21.00 © 2001 MAIK “Nauka/Interperiodica”

The energy of the δ electrons emitted at angle θ is $\varepsilon(\mu) = \varepsilon_m\mu^2$. The stopping power of a Kr ion in silicon is estimated to be $dE_1/dz \approx 8$ keV nm$^{-1}$. It follows from Eq. (1) that the maximum emission intensity of the δ electrons for $\xi \ll 1$ corresponds to nearly radial directions with $\mu \approx \xi/\sqrt{3}$. The direction-integrated emission intensity is estimated from Eq. (1) to be

$$\phi_{\text{int}} = \frac{r_{sc}E_m f'(z/L)}{NA\xi L\hbar}. \tag{2}$$

Here, the function $f(z/L) = E_1(z)/E_m$ gives the decrease in the energy of the fast impinging ions at depth $z$ along the track axis; and $E_m$ and $L$ are the initial energy and path length of the fast ions, respectively. For a silicon crystal irradiated with krypton ions with an energy of 210 MeV, the path length is $L \approx 30$ µm. If $N = 2$ and the screening radius is approximately equal to the Bohr radius, $r_{sc} \approx a_B = \hbar^2/me^2$, then $\phi_{\text{int}} \approx 1 \times 10^{18}$ s$^{-1}$. Since the mean interplanar spacing along the [111] axis in silicon is $d \approx 0.15$ nm, the mean excitation time of the atoms in a separate (111) crystal plane is $t_{ex} \approx d/V_1 \approx 7 \times 10^{-18}$ s. This corresponds to the knocking of about seven ($Z_* \approx 7$) δ electrons out of each interacting atom. Thus, after the passage of the fast ions, the opposite atoms of the two chains under consideration irradiate each other with δ electrons. This interaction between the atomic chains, which is switched on by a passing ion, causes their mutual repulsion over the lifetime of the excited state of δ electrons. Below, we estimate the number of δ electrons that provides such an interaction between the atomic chains.

To this end, the scattering pattern of these electrons should be considered. First, note that the energy of these δ electrons is

$$\varepsilon[\text{Ry}] \approx \frac{\hbar^2}{6mr_{sc}^2} \approx \frac{1}{6}\left(\frac{a_B}{r_{sc}}\right)^2$$

and corresponds to an electron wavelength $\lambda \approx 2\pi\sqrt{3}\,r_{sc}$. The electron transport mean free path $l_{tr}$, which is determined by elastic collisions with crystal atoms, for a screened Coulomb interaction is given by

$$\frac{1}{l_{tr}} = \pi n_0\left(\frac{Z_0e^2}{\varepsilon}\right)^2\int_0^1\frac{\eta^2 d\eta}{(\eta^2+\xi^2/\mu^2)^2}, \tag{3}$$

where $n_0$ is the atomic density of the target material and $Z_0$ is the number of electrons in the target ionized atom. It is easy to see that the maximum of the integrand in Eq. (3) at $\eta = \sin(\gamma/2) = 1$ corresponds to $\xi^2/\mu^2 \approx 3$, suggesting that the scattering angles $\gamma \approx \pi$ (head-on collisions) mainly contribute to elastic scattering and that the δ electrons under consideration can conserve energy for subsequent elastic collisions and can experience multiple backscatterings. This means that in the case of an approximate equality between the transport mean free path, half the electron wavelength, and the distance between the atomic chains, each pair of interacting atoms produces a standing electron wave in which the scattered δ electrons under consideration are involved. At the same time, the interacting atoms suffer several tens of elastic collisions, causing the mutual repulsion of the atomic chains under consideration. A gradual increase in the distance between the chains results in the violation of diffraction conditions for electrons and determines the lifetime of such a collective excited state.

The energy dependence of the electron transport mean free path (3) is

$$l_{tr} = l_0 g(\varepsilon), \tag{4}$$

where

$$l_0 = \frac{(a_B/r_{sc})^4}{2\pi Z_0^2 n_0 a_B^2},$$

and $g$ is a universal function that depends on parameter $\varsigma = 2mr_{sc}^2\varepsilon/\hbar^2$ as follows:

$$g(\varsigma) = \varsigma\left(\arctan\sqrt{\varsigma} - \frac{\sqrt{\varsigma}}{1+\varsigma}\right)^{-1}.$$

It has a minimum $g_m \approx 3.42$ at $\varsigma \approx 0.68$. For the δ electrons with energy $\varepsilon_w \approx \hbar^2/6mr_{sc}^2$ under consideration, $g$ is close to the minimum and is $g(1/3) \approx 3.8$.

The excitation condition for standing electron waves is

$$l_{tr} \approx \frac{k\lambda}{2} = k\pi\sqrt{3}r_{sc}. \tag{5}$$

This is the Bragg condition for backscattering. It is satisfied for $k = 1$ at the screening radius

$$r_{sc} = a_B\left(\frac{g(1/3)}{2\pi^2\sqrt{3}Z_0^2 n_0 a_B^3}\right)^{1/5}. \tag{6}$$

For a silicon crystal ($n_0 = 5 \times 10^{22}$ cm$^{-3}$ and $Z_0 \approx 7$), Eq. (6) gives the screening radius $r_{sc} \approx 0.8a_B$. For standing electron waves to be excited, the distance $l$ between the atomic chains must also be a multiple of half the wavelength of a δ electron:

$$l \approx k\lambda/2 = k\pi\sqrt{3}r_{sc}. \tag{7}$$

Equation (7) gives the screening radius at $k = 1$, which is also $r_{sc} \approx 0.8a_B$. The simultaneous satisfaction of conditions (5) and (7) yields the relation

$$Z_0^2 n_0 a_B^3\left(\frac{l}{a_B}\right)^5 \approx 4.5\pi^3 g\left(\frac{1}{3}\right) \tag{8}$$

for the parameters of a scattering system composed of two atomic chains. In our case, these parameters are: atomic density $n_0$, distance $l$ between the chains of interacting atoms, and the degree of ionization $Z_0$ of these atoms. If the number of atomic chains involved in the scattering of δ electrons is large, the number of parameters of the scattering system increases. At specified $n_0$ and $l$, relation (8) gives the degree of ionization of the interacting atoms at which they provide the backscattering of δ electrons and conserve their energy:

$$Z_0 \approx 3\pi\sqrt{\frac{\pi g(1/3)}{2n_0 a_B^3}}\left(\frac{a_B}{l}\right)^{5/2}. \qquad (9)$$

For silicon with $n_0 = 5 \times 10^{22}$ cm$^{-3}$ and $l \approx 0.23$ nm, Eq. (9) gives the number of electrons, $Z_0 \approx 7$, that remain in the interacting atoms after their ionization by fast ions and that scatter the knocked-out δ electrons. The number $N_w$ of δ electrons that each interacting atom donates to the generation of standing electron waves is determined by the ratio $\sqrt{\sigma_{bs}}/2\pi l$, where $\sigma_{bs}$ is the backscattering cross section. Estimates yield

$$\sigma_{bs} \approx \frac{9\pi}{32}Z_0^2\left(\frac{r_{sc}^2}{a_B}\right)^2,$$

which gives $\sqrt{\sigma_{bs}} \approx 0.2$ nm at $Z_0 \approx 7$ and $r_{sc} \approx 0.8a_B$. Therefore, the expression for the number of electrons in the state of a standing wave in the ranges of wavelengths and distances under consideration is

$$N_w \approx \frac{Z_*\sqrt{\sigma_{bs}}}{2\pi l} \approx \frac{1}{8\pi^2\sqrt{2\pi}}\frac{Z_0 Z_* l}{a_B},$$

which has a maximum at $Z_0 \approx Z_* \approx Z/2$ ($Z$ is the atomic number of the interacting atom). At $Z_0 \approx Z_* \approx 7$ and scattering length $l \approx 0.23$ nm, we obtain $N_w \approx 1$ for silicon. $N_w$ has the following physical meaning: this is the initial population of the state of a standing electron wave under consideration.

The parameters $l$, $Z_0$, and $n_0$ of the scattering medium given by relation (8) provide the required scattering power of the interacting atoms to maintain a standing wave at $k = 1$. The states at $k = 2, 3, 4, \dots$ can also be excited. However, their populations are considerably lower because of the small number of δ electrons with a shorter wavelength in the initial energy spectrum. The states of a standing wave with $k \geq 2$ can also be populated by δ electrons with energy $\varepsilon_w$ using atomic chains whose separation is equal to or is a multiple of $kl$.

As the interacting atoms are displaced during the scattering and recombination of δ electrons, the parameters $l$ and $Z_0$ of the scattering medium increase and condition (8) is violated. This causes the standing electron waves to be damped out. Thus, the lifetime of standing waves is determined by two damping processes and can be represented as

$$\tau^{-1} = \tau_p^{-1} + \tau_c^{-1},$$

where $\tau_c$ and $\tau_p$ are the lifetimes attributable to the displacement of interacting atoms and to the recombination of δ electrons, respectively. The decay of the quasi-steady state of a standing wave may be considered as an exponential decrease in its population: $N_w(t) \propto N_w\exp(-t/\tau)$. The recombination time in semiconductor crystals is known to be typically $\tau_p > 10^{-10}$ s. The displacement time of interacting atoms is much shorter, $\tau_c \ll \tau_p$. Therefore, the lifetime of a standing wave is determined by atomic displacement, $\tau \approx \tau_c$, and can be estimated from experimental data in [19].

For the lifetime of a standing wave to be estimated, the displacement of interacting atoms under multiple collisions of δ electrons must be considered. This displacement can be represented as resulting from the diffusion of a heavy gas in a light gas. The displacement velocity $V(t)$ of atoms can be represented as the product of their mobility by the force $F(t)$ determined by the momentum transferred to the atom from δ electrons during head-on collisions per unit time. If the oscillation frequency in a standing wave, $\omega \approx \pi v_w/l$, is not equal to the natural oscillation frequency of crystal atoms $\omega_0$, then it will suffice to take into account only the weakening of the force exerted on atoms due to a reduction in population $N_w$. In that case,

$$F(t) \approx 2mv_w N_w\frac{v_w}{l}\exp\left(-\frac{t}{\tau}\right)$$

and the equation of motion for the interacting atoms is

$$V(t) \approx \frac{ds}{dt} \approx \frac{4b}{l}\varepsilon_w N_w\exp\left(-\frac{t}{\tau}\right), \qquad (10)$$

where $N_w$ is the initial number of δ electrons (per atom) in the state of a standing wave under consideration, and $\varepsilon_w = mv_w^2/2 \approx \pi^2\hbar^2/2ml^2$ is the energy of such an electron. From Eq. (10), we derive the time dependence of atomic displacement, $s(t) \approx u[1 - \exp(-t/\tau)]$, and the total displacement

$$u \approx \frac{4\tau b N_w \varepsilon_w}{l}. \qquad (11)$$

Disregarding the atomic displacement along the track axis, we find the decrease in atomic density on the track axis, $\Delta n/n_0$, to be determined by a lengthening of the distance between the interacting atomic chains:

$$\begin{aligned}\frac{\Delta n}{n_0} &\approx \frac{2u}{l} \approx \frac{4\pi^2 N_w\hbar^2\tau^2}{Mml^4} \\ &= 4\pi^2 N_w\frac{m}{M}\left(\frac{\tau\,\mathrm{Ry}}{\hbar}\right)^2\left(\frac{a_B}{l}\right)^4,\end{aligned} \qquad (12)$$

where $\mathrm{Ry} = e^2/a_B \approx 27.2$ eV. Here, the mobility of the interacting atoms is taken to be $b \approx \tau/M$, where $M$ is their mass. It follows from (12) that the lifetime of a standing wave can be estimated from the measured reduction in atomic density on the track axis, $(\Delta n/n_0)_{\exp}$, using

$$\tau \approx \frac{\hbar}{2\pi \mathrm{Ry}}\sqrt{\frac{M}{mN_w}\left(\frac{\Delta n}{n_0}\right)_{\exp}}\left(\frac{l}{a_B}\right)^2. \qquad (13)$$

Assuming the frequency of standing waves, $\omega \approx \pi v_w/l \approx 2 \times 10^{16}$ s$^{-1}$, for silicon irradiated with krypton ions with energy $E_m = 210$ MeV to be much higher than the natural oscillation frequencies $\omega_0$ of crystal atoms, we use formula (13) to estimate the lifetime of a standing wave. The lifetime estimated for $(\Delta n/n_0)_{\exp} \approx 0.1$, $M = 28M_n$, $l = 0.23$ nm, and $N_w = 1$ is $\tau \approx 5.2 \times 10^{-15}$ s. The characteristic energy of the quasi-steady-state level for a standing wave is equal to the energy of the δ electrons, $\varepsilon_w \approx 7.5$ eV, in this wave for silicon, while the energy width of this level is $\Delta\varepsilon_w \approx \hbar/\tau \approx 0.13$ eV. It follows from (13) that the lifetime of the δ electrons with a wavelength $\lambda \approx 2l$ under consideration is proportional to the square of the wavelength. This means that the standing waves generated by longer wavelength δ electrons on more widely separated atomic chains can exist longer.

Given that the mass of the interacting atoms is $M = 28M_n \approx 4.67 \times 10^{-23}$ g, their mobility follows from our estimate of the lifetime for a standing electron wave: $b \approx \tau/M \approx 1 \times 10^8$ s g$^{-1}$. The δ electrons under consideration move between atoms at a velocity of $v_w \approx \pi\hbar/ml \approx 1.7 \times 10^8$ cm s$^{-1}$ and traverse a distance of $v_w\tau \approx 8.8$ nm in the lifetime of the standing wave. This means that at $\lambda \approx 0.23$ nm, each chain atom suffers $v\tau/l \approx 40$ head-on collisions with δ electrons in the lifetime of the standing wave. In this time, the atoms are displaced by $u \approx 0.5l\Delta n/n_0 \approx 0.01$ nm.

The mean displacement velocity of the interacting atoms in time $\tau$ is $\langle V\rangle \approx u/\tau \approx 2 \times 10^5$ cm s$^{-1}$ and exceeds the velocity imparted by the atom in a single collision by almost two orders of magnitude, $V_1 = 2mv_w/M \approx 6.6 \times 10^3$ cm s$^{-1}$. The mean velocity $\langle V\rangle$ has the same order of magnitude as the initial atomic velocity, which, according to (10), is $V_0 \approx 4b\varepsilon_w N_w/l \approx 2.3 \times 10^5$ cm s$^{-1}$. The motion of an atom in the force field of a damped standing wave gradually slows down from $V_0$ to zero.

The work done by standing electron waves to displace the interacting atoms can be estimated from the formula

$$A_d \approx \int_0^\infty F(t)V(t)dt \approx N_w\varepsilon_w\left(\frac{\Delta n}{n_0}\right)_{\exp},$$

which yields $A_d \approx 0.75$ eV per atom for silicon. The ratio $A_d/N_w\varepsilon_w \approx 0.1$ means that the standing wave spends 10% of its energy on crystal deformation in the track region. In the case where the silicon (111) surface is irradiated with krypton ions with energy 210 MeV, about $4 \times 10^5$ silicon atoms are contained in the two interacting chains at a path length $L \approx 30$ μm if the mean interatomic distance along the track axis is $d \approx 0.15$ nm. The deformation energy of these two chains ($\simeq$300 keV) then accounts for about 0.14% of the energy of the impinging krypton ion. The excitation of several standing waves can lead to a higher expenditure of fast-ion energy on deformation.

In the right-hand part of Eq. (10), we discarded the terms that describe viscosity and the elastic force restoring the displaced atoms to their original position, because we consider the case where the electron pressure acts much faster than the elastic and friction forces. An experiment [19] and an estimate of the tension between two atoms, $2n_0A_d \approx 7.5 \times 10^{22}$eV cm$^{-3}$, which exceeds the shear modulus $G \approx 10^{10}$ N m$^{-2} \approx 6 \times 10^{22}$ eV cm$^{-3}$, suggest that the displacement of the interacting atoms is irreversible. Including the restoring force $M\omega_0^2 s$ and the friction force $M\kappa\, ds/dt$ in the right-hand part of Eq. (10) can bring the atoms to their original position $s = 0$ only when $\beta = (\omega_0\tau)^2/(1 + \kappa\tau) \sim 1$ ($k$ is the coefficient of viscosity). In this case, the atomic displacement changes in accordance with

$$s = \frac{u(e^{-\beta t/\tau} - e^{-t/\tau})}{1 + \kappa\tau - \omega_0^2\tau^2},$$

the interacting atom executes an aperiodic oscillation, and the breaking point is not reached. For $\beta \ll 1$, the atom is displaced so rapidly that the energy transferred from δ electrons has time to be spent on the deformation, which exceeds the shear modulus and becomes irreversible as a result. In the latter case, the atom no longer returns to its original position, and, according to (11), the displacement is $u \approx 0.5l\Delta n/n_0$.

In conclusion, it should be emphasized that the standing electron waves excited in fast-ion tracks can displace atomic chains. This mechanism must lead to the generation of dislocations, which can then be stabilized or move through the crystal. In this paper, we have considered in detail the excitation of a linearly polarized wave between linear atomic chains. In general, a superposition of standing waves corresponding to different distances between the atomic chains can be excited. The excitation of a complete spectrum of standing electron waves can give rise to peculiar cigar-shaped dilatations with an appreciable reduction in atomic density on the track axis.

## ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 00-02-17693 and 01-02-16279).

## REFERENCES

1. A. Dunlop and D. Lesueur, Radiat. Eff. Defects Solids **126**, 132 (1993).

2. A. Iwase and T. Iwata, Nucl. Instrum. Methods Phys. Res. B **90**, 332 (1994).
3. A. Benyagoub and L. Thome, Phys. Rev. B **38**, 10205 (1988).
4. A. Audouard, E. Balanzad, S. Bouffard, *et al.*, Phys. Rev. Lett. **65**, 875 (1990).
5. S. Klaumunzer, C. Li, and G. Schumacher, Appl. Phys. Lett. **51**, 97 (1987).
6. S. Klaumunzer, C. Li, S. Loffler, *et al.*, Radiat. Eff. Defects Solids **108**, 131 (1989).
7. A. I. Ryazanov, A. E. Volkov, and S. Klaumunzer, Phys. Rev. B **51**, 12107 (1995).
8. H. Trinkaus and A. I. Ryazanov, Phys. Rev. Lett. **74**, 5072 (1995).
9. M. I. Kaganov, I. M. Lifshits, and L. V. Tanatarov, Zh. Éksp. Teor. Fiz. **31**, 232 (1956) [Sov. Phys. JETP **4**, 173 (1957)].
10. I. M. Lifshits, M. I. Kaganov, and L. V. Tanatarov, At. Énerg. **6**, 391 (1959).
11. T. Talemonde, C. Dufour, and E. Paumier, Phys. Rev. B **46**, 14362 (1992).
12. K. Yasui, Nucl. Instrum. Methods Phys. Res. B **90**, 409 (1994).
13. R. L. Fleischer, P. B. Price, and R. M. Walker, J. Appl. Phys. **36**, 3645 (1965).
14. S. Klaumunzer, Ming-dong Hou, and G. Schumacher, Phys. Rev. Lett. **57**, 850 (1986).
15. D. Lesueur and A. Dunlop, Radiat. Eff. Defects Solids **126**, 163 (1993).
16. A. Dunlop, P. Legrand, D. Lesueur, *et al.*, Europhys. Lett. **15**, 765 (1991).
17. A. E. Volkov and V. A. Borodin, in *Proceedings of the 1st International Congress on Radiation Physics; High Current Electronics and Modification of Materials, Tomsk, 2000*, Ed. by D. Vaisburg, Vol. 1, p. 231.
18. E. V. Metelkin and A. I. Ryazanov, Zh. Éksp. Teor. Fiz. **117**, 420 (2000) [JETP **90**, 370 (2000)].
19. A. R. Chelyadinskiĭ, V. S. Varichenko, and A. M. Zaĭtsev, Fiz. Tverd. Tela (St. Petersburg) **40**, 1627 (1998) [Phys. Solid State **40**, 1478 (1998)].

*Translated by V. Astakhov*

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 415–423.*
*From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 470–479.*
*Original English Text Copyright © 2001 by Kagan, Kugel, Khomskii.*

**SOLIDS**
**Electronic Properties**

# Phase Separation in Systems with Charge Ordering[¶]

**M. Yu. Kagan[a,] *, K. I. Kugel[b], and D. I. Khomskii[c]**
[a]*Kapitza Institute for Physical Problems, Russian Academy of Sciences, Moscow, 117334 Russia*
**e-mail: kagan@kapitza.ras.ru*
[b]*Institute of Theoretical and Applied Electrodynamics, Russian Academy of Sciences, Moscow, 127412 Russia*
[c]*Laboratory of Applied and Solid State Physics, Materials Science Center,*
*University of Groningen, 9747AG, Groningen, the Netherlands*

Received December 25, 2000

**Abstract**—A simple model of charge ordering is considered. It is explicitly shown that at any deviation from half-filling ($n \neq 1/2$), the system is unstable with respect to the phase separation into the charge ordered regions with $n = 1/2$ and the metallic regions with a smaller electron or hole density. A possible structure of this phase-separated state (metallic droplets in a charge ordered matrix) is discussed. The model is extended to account for the strong Hund-rule onsite coupling and the weaker intersite antiferromagnetic exchange. The analysis of this extended model allows us to determine the magnetic structure of the phase-separated state and to reveal the characteristic features of the manganites and other substances with charge ordering. *© 2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

The problem of charge ordering in magnetic oxides has attracted the attention of theorists since the discovery of the Verwey transition in magnetite at the end of the 1930s [1]. An early theoretical description of this phenomenon was given, e.g., in [2]. This problem was recently reexamined in a number of papers in connection with the colossal magnetoresistance in manganites, see, e.g., [3–5]. Mechanisms stabilizing the charge ordered state can be different: the Coulomb repulsion of charge carriers (the energy minimization requires keeping the carriers as far away as possible, similarly to the Wigner crystallization) or the electron-lattice interaction leading to the effective repulsion of electrons at the nearest-neighbor sites. In all cases, charge ordering can arise in mixed-valence systems if the electron bandwidth is sufficiently small. In the opposite case, the large electron kinetic energy stabilizes the homogeneous metallic state. In real materials, in contrast to the Wigner crystallization, the underlying lattice periodicity determines the preferential types of charge ordering. Thus, in the simplest bipartite lattice, to which belongs the colossal magnetoresistance manganites of the type $R_{1-x}A_xMnO_3$ (where R = La, Pr and A = Ca, Sr) or layered manganites $R_{2-x}A_xMnO_4$, $R_{2-2x}A_{1+2x}Mn_2O_7$, the optimum conditions for the formation of the charge ordered state exist for the doping $x = 1/2$. At this value of $x$, the concentrations of $Mn^{3+}$ and $Mn^{4+}$ are equal and the simple checkerboard arrangement is possible. The most remarkable experimental fact here is that even at $x \neq 1/2$ (in the underdoped manganites with $x < 1/2$), only the simplest version of charge ordering is experimentally observed with the alternating checkerboard structure of the occupied and empty sites in the basal plane [6]. In other words, this structure corresponds to the doubling of the unit cell, whereas more complicated structures with a longer period (or even incommensurate structures) do not actually appear in this case.

A natural question then arises as to how the extra or missing electrons can be redistributed for an arbitrary doping level such that the superstructure remains the same as for $x = 1/2$? To answer this question, the experimentalists introduced the concept of the incipient charge ordered state corresponding to the distortion of a long-range charge ordering by microscopic metallic clusters [7]. In fact, the existence of this state implies a kind of phase separation. We note that the phase separation scenario in manganites is very popular presently [8–15]. There is a growing evidence nowadays suggesting that an interplay between the charge ordering and the tendency toward phase separation plays an essential role in the physics of materials with colossal magnetoresistance.

In this paper, we consider a simple model allowing us to clarify the situation at an arbitrary doping. The model includes both the Coulomb repulsion of electrons on the neighboring sites and the magnetic interactions responsible for the magnetic ordering in manganites. After demonstrating the instability of the system toward phase separation in certain doping ranges, we consider the simplest form of the phase separation—the formation of metallic droplets in the insulating matrix. We estimate parameters of such droplets and construct the phase diagram illustrating the interplay between charge ordering, magnetic ordering, and phase separation.

---

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

We note that the charge ordering mechanism considered below (the Coulomb repulsion) is not the only one. The electron-lattice interaction can also play an important role (see, e.g., [16]). In application to manganites, one must also take the orbital and magnetic interactions into account [4, 16, 17]. These may be important, in particular, in explaining the fact that the charge ordering in half-doped perovskite manganites is a checkerboard one only in the basal plane, but it is "in-phase" along the *c*-direction. However, the nature of this charge ordering is not clear yet and presents a separate problem: it is not evident that the dominant mechanism is indeed given by the magnetic interactions responsible for this stacking of *ab*-planes in [16]. We also emphasize that the charge ordering is often observed in manganites at higher temperatures than the magnetic ordering, and one must seek a model that does not heavily rely on magnetic interactions. In contrast to magnetic interactions, the Coulomb interaction is one of the important factors that is always present in the systems under consideration. Moreover, it has a universal nature and does not critically depend on specific features of a particular system. Consequently, our treatment can also be applied to other systems with charge ordering such as magnetite $Fe_3O_4$ [1], cobaltites [18], nickelates [19], etc.

## 2. THE SIMPLEST MODEL FOR CHARGE ORDERING

We consider a simple lattice model for charge ordering,

$$\hat{H} = -t\sum_{\langle i,j\rangle} c_i^+ c_j + V\sum_{\langle i,j\rangle} n_i n_j - \mu\sum_i n_i, \tag{1}$$

where $t$ is the hopping integral, $V$ is the nearest-neighbor Coulomb interaction (a similar *nn* repulsion can also be obtained via the interaction with the breathing-type optical phonons), $\mu$ is the chemical potential, and $c_i^+$ and $c_j$ are one-electron creation and annihilation operators, $n_i = c_i^+ c_i$. The symbol $\langle i, j\rangle$ denotes the summation over the nearest-neighbor sites. Here, we omit spin and orbital indices for simplicity. As mentioned in the Introduction, the spin and orbital effects play an important role in the formation of the real structure in specific compounds; in this section, however, we emphasize the most robust effects related to the nearest-neighbor Coulomb repulsion. The magnetic effects are discussed in Section 5. We also assume that the double occupancy does not occur in this model because of the strong onsite repulsion between electrons.

Hamiltonian (1) explicitly accounts for the correlation effect that is most important for the formation of charge ordering, namely, the electron repulsion on neighboring sites. The long-range part of the Coulomb interaction only leads to the renormalization of the bandwidth W and does not significantly affect the properties of the uniform charge ordered state. However, it can produce a qualitative effect on the structure of the phase-separated state (see the discussion in the beginning of Section 4).

The models of type (1) with the *nn* repulsion responsible for the charge ordering are the most popular in describing this phenomenon, see, e.g., [2, 3, 5, 20] and references therein. Hamiltonian (1) captures the main physical effects; if necessary, one can add some extra terms to it, which we do in Section 5.

In the main part of this paper, we always speak about electrons. However, in application to real manganites, we mostly have in mind less than half-doped (underdoped) systems of the type $R_{1-x}A_xMnO_3$ with $x < 1/2$. For a real system, one must therefore substitute holes for our electrons. All the theoretical treatment definitely remains the same (from the very beginning, we could define the $c$ and $c^+$ operators in (1) as the operators of holes); we hope that this does not lead to any misunderstanding.

In what follows, we consider the simplest case of square (2D) or cubic (3D) lattices, where the simple two-sublattice ordering occurs for $x = 1/2$. As mentioned in the Introduction, this is the case in layered manganites, whereas in 3D perovskite manganites, this ordering occurs only in the basal plane (the ordering is "in-phase" along the $c$ direction). A more complicated model is apparently needed to account for this behavior.

For $n = 1/2$, model (1) was analyzed in many papers; we follow the treatment in [2]. As mentioned above, the Coulomb repulsion (the second term in (1)) stabilizes the charge ordering in the form of a checkerboard arrangement of the occupied and empty sites, whereas the first term (band energy) opposes this tendency. At arbitrary values of the electron density $n$, we first consider a homogeneous charge ordered solution and use the same ansatz as in [2], namely,

$$n_i = n[1 + (-1)^i\tau]. \tag{2}$$

This expression implies doubling the lattice periodicity, with the local densities

$$n_1 = n(1+\tau), \quad n_2 = n(1-\tau)$$

at the neighboring sites. We note that at $n = 1/2$ for a general form of the electron dispersion without nesting, the charge ordered state exists only for a sufficiently strong repulsion $V > 2t$ [2]. The order parameter is $\tau < 1$ for finite $V/2t$, and the ordering is not complete in general; i.e., an average electron density $n_i$ differs from zero or one even at $T = 0$.

We use the coupled Green function approach as in [2], which yields

$$\begin{cases} (E+\mu)G_1 - t_k G_2 - zVn(1-\tau)G_1 = \dfrac{1}{2\pi}, \\ (E+\mu)G_2 - t_k G_1 - zVn(1+\tau)G_2 = 0, \end{cases} \tag{3}$$

where $G_1$ and $G_2$ are the Fourier transforms of the normal lattice Green functions

$$G_{il} = \langle\langle c_i c_l^+ \rangle\rangle$$

for the sites $i$ and $l$ belonging respectively to the same sublattice or to different sublattices, $z$ is the number of nearest neighbors, and $t_k$ is the Fourier transform of the hopping matrix element. In deriving (3), we performed a mean-field decoupling and replaced the averages $\langle c_i^+ c_i \rangle$ by the onsite densities $n_i$ in Eq. (2). The solution of Eqs. (3) leads to the following spectrum:

$$E + \mu = Vnz \pm \sqrt{(Vn\tau z)^2 + t_k^2} = Vnz \pm \omega_k. \tag{4}$$

The spectrum defined by (4) resembles the spectrum of superconductor and, hence, the first term under the square root is analogous to the superconducting gap squared. In other words, we can introduce the charge ordered gap by the formula

$$\Delta = Vn\tau z.$$

It depends on the density not only explicitly, but also via the density dependence of $\tau$.

We thus obtain

$$\omega_k = \sqrt{\Delta^2 + t_k^2}. \tag{5}$$

We note a substantial difference between the spectrum of charge ordered state (5) and the superconducting state, namely, the chemical potential does not enter under the square root in (5) for $n \neq 1/2$, which is in contrast to the spectrum of superconductor, where

$$\omega_k = \sqrt{(t_k - \mu)^2 + \Delta^2}.$$

We can then find the Green functions

$$\begin{cases} G_1 = \dfrac{A_k}{E + \mu - Vnz - \omega_k + i0} \\ \quad + \dfrac{B_k}{E + \mu - Vnz + \omega_k + i0} \\ G_2 = \dfrac{t_k}{2\omega_k}\dfrac{1}{2\pi}\Big[\dfrac{1}{E + \mu - Vnz - \omega_k + i0} \\ \quad - \dfrac{1}{E + \mu - Vnz + \omega_k + i0}\Big], \end{cases} \tag{6}$$

where

$$A_k = \frac{1}{4\pi}\left(1 - \frac{\Delta}{\omega_k}\right), \quad B_k = \frac{1}{4\pi}\left(1 + \frac{\Delta}{\omega_k}\right). \tag{7}$$

After the standard Wick transformation

$$E + i0 \longrightarrow iE$$

in the expression for $G_1$, we find the densities in the following form:

$$\begin{aligned} n_1 &= n(1 + \tau) \\ &= \int\Big[\left(1 - \frac{\Delta}{\omega_k}\right) f_F(\omega_k - \mu + Vnz) \\ &\quad + \left(1 + \frac{\Delta}{\omega_k}\right) f_F(-\omega_k - \mu + Vnz)\Big]\frac{d\mathbf{k}}{2\Omega_{BZ}}, \\ n_2 &= n(1 - \tau) \\ &= \int\Big[\left(1 + \frac{\Delta}{\omega_k}\right) f_F(\omega_k - \mu + Vnz) \\ &\quad + \left(1 - \frac{\Delta}{\omega_k}\right) f_F(-\omega_k - \mu + Vnz)\Big]\frac{d\mathbf{k}}{2\Omega_{BZ}}, \end{aligned} \tag{8}$$

where

$$f_F(y) = \frac{1}{e^{y/T} + 1}$$

is the Fermi distribution function and $\Omega_{BZ}$ is the volume of the first Brillouin zone.

Adding and subtracting the two equations for $n_1$ and $n_2$, we obtain the resulting system of equations for $n$ and $\mu$:

$$\begin{aligned} 1 &= Vz\int\frac{1}{\omega_k}[f_F(-\omega_k - \mu + Vnz) \\ &\quad - f_F(\omega_k - \mu + Vnz)]\frac{d\mathbf{k}}{2\Omega_{BZ}}, \\ n &= \int[f_F(-\omega_k - \mu + Vnz) \\ &\quad + f_F(\omega_k - \mu + Vnz)]\frac{d\mathbf{k}}{2\Omega_{BZ}}. \end{aligned} \tag{9}$$

For low temperatures ($T \longrightarrow 0$) and $n \leq 1/2$, it is reasonable to assume that $\mu - Vnz$ is negative. Therefore,

$$f_F(\omega_k - \mu + Vnz) = 0$$

and

$$f_F(-\omega_k - \mu + Vnz) = \theta(-\omega_k - \mu + Vnz)$$

is the step function.

It is easy to see that for $n = 1/2$, the system of equations (9) yields identical results for all

$$-\Delta \leq \mu - Vnz \leq \Delta.$$

From this point of view, $n = 1/2$ is the indifferent equilibrium point. For infinitely small deviations from $n = 1/2$, that is, for densities $n = 1/2 - 0$, the chemical potential must be defined as

$$\mu = -\Delta + \frac{Vz}{2} = \frac{Vz}{2}(1 - \tau).$$

If we consider the strong coupling case $V \gg 2t$ and assume a constant density of states inside the band, we have

$$\tau = 1 - \frac{2W^2}{3V^2z^2}$$

for a simple cubic lattice and, therefore,

$$\mu = \frac{W^2}{3Vz}, \tag{10}$$

where $W = 2zt$ is the bandwidth. We note that for the density $n = 1/2$, the charge ordered gap $\Delta$ appears for an arbitrary interaction strength $V$. This is due to the existence of nesting in our simple model. In the weak coupling case $V \ll 2t$ and with perfect nesting, we have

$$\Delta \propto W \exp\left\{-\frac{W}{Vz}\right\}$$

and $\tau$ is exponentially small. For $Vz \gg W$ or, accordingly, for $V \gg 2t$, it follows that $\Delta \approx Vz/2$ and $\tau \longrightarrow 1$. As mentioned above, for a general form of the electron dispersion without nesting, the charge ordering exists only if the interaction strength $V$ exceeds a certain critical value of the order of the bandwidth $W$ [2]. In what follows, we restrict ourselves to the physically more instructive strong-coupling case $V \gg 2t$.

For the constant density of states (flat band), the integrals in (9) can be taken explicitly and the system of equations (9) can be easily solved for arbitrary $n$. We note, however, that in the strong-coupling case $V \gg 2t$ and for small density deviations from 1/2 ($\delta \ll 1$), the results are not very sensitive to the form of the electron dispersion. That is why we do not need to solve the system of equations (9) exactly.

We now consider the case where $n = 1/2 - \delta$, with $\delta \ll 1$ being the density deviation from 1/2. In this case, $\mu = \mu(\delta, \tau)$ and we have two coupled equations for $\mu$ and $\tau$. As a result,

$$\begin{aligned} \mu(\delta) &\approx Vnz(1-\tau) - \frac{4W^2}{Vz}\delta^2 \\ &\approx \frac{W^2}{3Vz} + \frac{4W^2}{3Vz}\delta + O(\delta^2). \end{aligned} \tag{11}$$

The energy of the charge ordered state is therefore given by

$$E_{CO}(\delta) = E_{CO}(0) - \frac{W^2}{3Vz}\delta - \frac{2W^2}{3Vz}\delta^2 + O(\delta^3), \tag{12}$$

where

$$E_{CO}(0) = -\frac{W^2}{6Vz}$$

is the energy precisely corresponding to the density $n = 1/2$ and $|E_{CO}(0)| \ll W$. At the same time, the charge ordered gap $\Delta$ is given by

$$\Delta \approx \frac{Vz}{2}\left[1 - 2\delta - \frac{2W^2}{3V^2z^2}(1 + 4\delta)\right]. \tag{13}$$

The dependence of the chemical potential $\mu$ and the total energy $E$ on $\delta$ in Eqs. (11) and (12) actually stems from this linear decrease of the energy gap $\Delta$ with the deviation from half-filling.

For $n > 1/2$, the energy of the charge ordered state starts to increase rapidly due to a large contribution of the Coulomb repulsion (the upper Verwey band is partially filled for $n > 1/2$). Note that for $n > 1/2$, contrary to the case where $n < 1/2$, each extra electron put into the checkerboard charge ordered state necessarily has occupied nearest-neighbor sites, increasing the total energy by $Vz|\delta|$. For $|\delta| = n - 1/2 > 0$, we then have

$$\begin{aligned} E_{CO}(\delta) &= E_{CO}(0) \\ &+ \left(Vz - \frac{W^2}{3Vz}\right)|\delta| - \frac{2W^2}{3Vz}\delta^2 + O(\delta^3). \end{aligned} \tag{14}$$

Accordingly, the chemical potential is given by

$$\mu(\delta) = Vz - \frac{W^2}{3Vz} - \frac{4W^2}{3Vz}|\delta| + O(\delta^2). \tag{15}$$

It undergoes a jump equal to $Vz$ as $\tau \longrightarrow 1$. We note that the gap $\Delta$ is symmetric for $n > 1/2$ and is given by

$$\Delta \approx \frac{Vz}{2}\left[1 - 2|\delta| - \frac{2W^2}{3V^2z^2}(1 + 4|\delta|)\right].$$

We could make the entire picture symmetric with respect to $n = 1/2$ by shifting all the one-electron energy levels and the chemical potential by $Vz/2$, i.e., defining

$$\mu' = \mu - Vz/2.$$

In terms of $\mu'$, Eqs. (11) and (15) can be written as

$$\mu' = -\frac{Vz}{2} + \frac{W^2}{3Vz} + \frac{4W^2}{3Vz}\delta, \quad n < \frac{1}{2},$$

$$\mu' = \frac{Vz}{2} - \frac{W^2}{3Vz} - \frac{4W^2}{3Vz}|\delta|, \quad n > \frac{1}{2}.$$

Similarly to the situation in semiconductors, we have $\mu' = 0$ precisely at the point $n = 1/2$, which means that the chemical potential lies in the middle of the band gap (see Fig. 1). At densities $n = 1/2 - 0$, the chemical potential $\mu' = -Vz/2$ coincides with the upper edge of the filled Verwey band.

## 3. PHASE SEPARATION

We now check the stability of the charge ordered state. At the densities close to $n = 1/2$, the dependence of energy on the charge density has the form illustrated in Fig. 2. This figure clearly indicates a possible instability of the charge ordered state. Indeed, the most remarkable implication of Eqs. (11)–(15) is that the compressibility $\kappa$ of the homogeneous charge ordered system is negative for the densities different from 1/2,

$$\frac{1}{\kappa} \propto \frac{d\mu}{dn} = -\frac{d\mu}{d\delta} = \frac{d^2E}{d\delta^2} = -\frac{4W^2}{3Vz} < 0, \tag{16}$$

where $\delta = 1/2 - n$. This is a manifestation of the tendency toward the phase separation characteristic of the charge ordered system with $\delta \neq 0$. The presence of a kink in $E_{CO}(\delta)$ (cf. Eqs. (12) and (14)) implies that one of the states into which the system might separate would correspond to the checkerboard charge ordered state with $n = 1/2$, whereas the other would have a certain density $n'$ smaller or larger than 1/2. This conclusion resembles that in [4] (see also [10, 14]), although the detailed physical mechanism is different. The possibility of phase separation in model (1) away from half-filling was also reported earlier in [12] for the infinite-dimensional case. In what follows, we focus our attention on the situation with $n < 1/2$ (underdoped manganites); the case where $n > 1/2$ apparently has certain special properties—the existence of stripe phases, etc. [13], the detailed origin of which is not yet clear.

It is easy to understand the physics of the phase separation in our case. As follows from (13), the charge ordered gap decreases linearly with deviation from half-filling. Correspondingly, the energy of the homogeneous charge ordered state rapidly increases, and it is more favorable to “extract” extra holes from the charge ordered state, putting them into one part of the sample, while creating the “pure” checkerboard charge ordered state in the other part. The energy loss due to this redistribution of holes is overcompensated by the gain provided by a better charge ordered.

However, the hole-rich regions would not be completely “empty,” similarly to pores (clusters of vacancies) in crystals: we can gain extra energy by “dissolving” a certain amount of electrons there. In doing this, we decrease the band energy of the electrons due to their delocalization. Thus, this second phase would be a metallic one. The simplest state of this kind is a homogeneous metal with the electron concentration $n_m$. This concentration, as well as the relative volume of the metallic and charge ordered phases, can be easily calculated by minimizing the total energy of the system. The energy of the metallic part of the sample $E_m$ is given by

$$E_m = -tzn_m + ct(n_m)^{5/3} + V(n_m)^2, \tag{17}$$

where $c$ is a constant.

Minimizing (17) with respect to $n_m$, we find the equilibrium electron density in the metallic phase. For the strong coupling case $V > zt$, we obtain (neglecting a relatively small correction provided by the term with $(n_m)^{5/3}$)

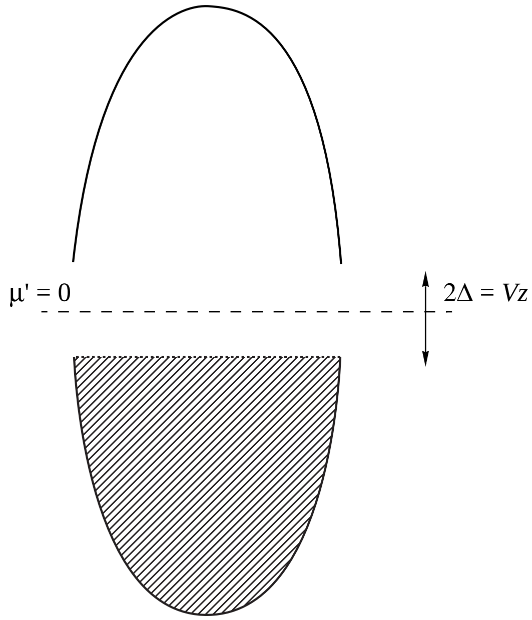

**Fig. 1.** Band structure of model (1) at $n = 1/2$. The lower Verwey band is completely filled. The upper Verwey band is empty. Chemical potential $\mu' = 0$ lies in the middle of the band gap with the width $2\Delta$.

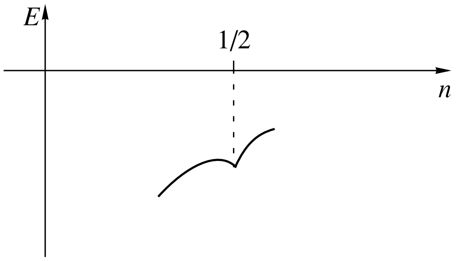

**Fig. 2.** Energy of the charge ordered state versus charge density for $n \longrightarrow 1/2$.

$$n_{m0} \approx tz/2V. \tag{18}$$

In accordance with this simple treatment, the system with $n_{m0} < n < 1/2$ would therefore undergo the phase separation into the charge ordered phase with $n = 1/2$ and the metallic phase with $n = n_{m0}$. For arbitrary $n$, the relative volumes $V_m$ and $V_{CO}$ of these phases can be found from the Maxwell construction,

$$\frac{V_m}{V_{CO}} = \frac{1/2 - n}{n - n_{m0}}, \tag{19}$$

which implies that the metallic phase occupies the part $V_m$ of the total volume $V$ given by

$$\frac{V_m}{V} = \frac{1/2 - n}{1/2 - n_{m0}}. \tag{20}$$

The metallic phase would occupy the entire sample when the total electron density $n$ is less than $n_{m0}$.

## 4. AN EXAMPLE: THE PHASE SEPARATED STATE WITH METALLIC DROPLETS

As argued above, the system with a short-range repulsion described by Eq. (1) is unstable with respect to the phase separation for $n$ close to but different from 1/2. The long-range Coulomb forces would, however, prevent the full phase separation into large regions containing all extra holes and the pure $n = 1/2$ charge ordered region. We can avoid this energy loss by forming, instead of one big metallic phase with many electrons, finite metallic clusters with smaller number of electrons. The limiting case would be a set of spherical droplets, each containing one electron. This state is similar to magnetic polarons ("ferrons") considered in the problem of phase separation for doped magnetic insulators [8, 14, 11].

We now estimate the characteristic parameters of these droplets. The main purpose of this treatment is to demonstrate that the energy of the state constructed in this way is lower than the energy of the homogeneous state, even if we treat these droplets rather crudely and do not optimize all their properties. In particular, we make the simplest assumption that the droplets have sharp boundaries and that the charge ordered state existing outside these droplets is not modified in their vicinity. This state can be treated as a variational one: optimizing the structure of the droplet boundary can only decrease its energy.

The energy (per unit volume) of the droplet state with the concentration of droplets $n_d$ can be written in total analogy with the ferron energy in the double-exchange model (see [14, 11]). This yields

$$\begin{aligned} E_{\text{droplet}} &= -tn_d\left(z - \frac{\pi^2 a^2}{R^2}\right) \\ &- \frac{W^2}{6Vz}\left[1 - n_d\frac{4}{3}\pi\left(\frac{R}{a}\right)^3\right], \end{aligned} \tag{21}$$

where $a$ is the lattice constant and $R$ is the droplet radius. The first term in (21) corresponds to the kinetic energy gain of the electron delocalization inside the metallic droplets and the second term describes the charge ordered energy in the remaining insulating part of the sample.

Minimization of the energy in (21) with respect to $R$ gives

$$\frac{R}{a} \approx \left(\frac{2V}{t}\right)^{1/5}. \tag{22}$$

The critical concentration $n_{dc}$ corresponds to the configuration where metallic droplets start to overlap, i.e., where the volume of the charge ordered phase (the second term in (21)) tends to zero. Hence,

$$n_{dc} = \frac{3}{4\pi}\left(\frac{a}{R}\right)^3 \propto \left(\frac{t}{V}\right)^{3/5}. \tag{23}$$

Actually, one should include the surface energy contribution to the total energy of the droplet. The surface energy must be of the order of $W^2R^2/V$. For large droplets, this contribution is small compared to the term $\propto R^3$ in (21); it would also be reduced for a "soft" droplet boundary. It is easy to show that even in the worst case of a small droplet (of the order of several lattice constants) with a sharp boundary, $R/a$ acquires the factor $1 - 0.2(t/2V)^{1/5}$ related to the surface contribution. Thus, the corrections related to the surface would not exceed about 20% of the bulk value. That is why we ignore this term below.

Comparing (12) with (21) and (22), we see that for the deviations from half-filling

$$0 < \delta \le \delta_c = 1/2 - n_{dc},$$

the energy of the phase-separated state is always lower than the energy of the homogeneous charge ordered state. The energy of the droplet state (21) with the radius given by (22) is also lower than the energy of the fully phase separated state obtained by the Maxwell construction from the homogeneous metallic state (17). Correspondingly, the critical concentration $n_{dc}$ in Eq. (23) is larger than $n_{m0}$ in Eq. (18). There is no contradiction here: in the droplet state that we constructed, the electrons are confined to spheres of the radius $R$, and even when these droplets start to overlap at $n = n_{dc}$, occupying the entire sample, the electrons, by construction, are still confined within their own spheres and avoid each other. In other words, a certain degree of charge ordered correlations is still present in our droplet state, decreasing the repulsion and hence the total energy.

Thus, the energy of the phase separated state with the droplets corresponds to the global minima of the energy for all $0 < \delta \le \delta_c$. This justifies our conclusion about the phase separation into the charge ordered state with $n = 1/2$ and a metallic state with small spherical droplets.

The situation encountered here resembles that of a partially filled strongly interacting Hubbard model, with the charge ordered state corresponding to the antiferromagnetic state of the latter and with the nearest-neighbor interaction $V$ playing the role of Hubbard's $U$. In both cases, the kinetic energy of doped carriers tends to destroy this "antiferro," or charge ordering, by first "spoiling" it in their vicinity and eventually leading to the formation of the metallic state (Nagaoka ferromagnetism). In the Hubbard model, we also face the situation with the phase separation at a sufficiently small doping [21].

We also note that for $n > 1/2$, the compressibility of the charge ordered state is again negative,

$$\frac{1}{\kappa} = -\frac{4W^2}{3Vz} < 0,$$

and has the same value as for $n < 1/2$. As a result, it is again more favorable to create a phase-separated state for these densities. However, as already mentioned, the nature of the second phase with $n > 1/2$ is not quite clear at present, and therefore, we do not consider this case here.

## 5. AN EXTENDED MODEL

We can now extend the model discussed in the previous sections by taking the essential magnetic interactions into account. In manganites, in addition to the conduction electrons in the $e_g$ bands, there also exist practically localized $t_{2g}$ electrons, which we now include in our consideration. The corresponding Hamiltonian is given by

$$\begin{aligned}\hat{H} = -t\sum_{\langle i,j\rangle,\sigma} c^{+}_{i\sigma}c_{j\sigma} + V\sum_{\langle i,j\rangle} n_i n_j - J_H\sum_i \mathbf{S}_i\boldsymbol{\sigma}_i \\ + J\sum_{\langle i,j\rangle}\mathbf{S}_i\mathbf{S}_j - \mu\sum_i n_i.\end{aligned} \tag{24}$$

In comparison to (1), the additional terms here correspond to the strong Hund-rule onsite coupling $J_H$ between the localized spins $\mathbf{S}$ and the spins of conduction electrons $\sigma$, and a relatively weak Heisenberg antiferromagnetic (AFM) exchange $J$ between neighboring local spins. In real manganites, the AFM ordering of the zigzag (CE) type in the charge ordered phase is determined not only by the exchange of the localized $t_{2g}$ electrons but to a large extent by the charge- and orbitally ordered $e_g$ electrons themselves. For simplicity, we ignore this factor and assume the superexchange interaction to be the same in the charge ordering and in the metallic phases.

It is physically reasonable to consider this model in the limit

$$J_H S > V > W > JS^2.$$

In the absence of the Coulomb term, this is exactly the conventional double-exchange model (see, e.g., [8, 14]). As is usually assumed in the theory of the double exchange (that is, in the theory where $J_H \gg W$), the main role of the itinerant electrons is to form a parallel arrangement of local spins. The exchange-correlation effects of the itinerant electrons themselves are not very important here and can be included in the renormalization of the effective bandwidth.

We note that the absence of doubly occupied sites in (24) is guaranteed by the large Hund term. It also favors the metallicity in the system, because the effective bandwidth depends on the magnetic order in our problem. The estimate for the critical concentration is therefore different from the one in (23). Similarly to [14], the metallic droplets are ferromagnetic (FM) because of the double exchange. The energy of one such droplet is given by

$$\begin{aligned}E = -t\left(z - \frac{\pi^2 a^2}{R^2}\right) - \frac{W^2}{6Vz}\left[1 - \frac{4}{3}\pi\left(\frac{R}{a}\right)^3\right] \\ + zJS^2\frac{4}{3}\pi\left(\frac{R}{a}\right)^3 - zJS^2\left[1 - \frac{4}{3}\pi\left(\frac{R}{a}\right)^3\right].\end{aligned} \tag{25}$$

The last two terms in (25) describe the loss of the Heisenberg AFM exchange energy inside the FM metallic droplets and the gain of this energy in the AFM insulating part of the sample, respectively. The minimization with respect to the droplet radius (as in (21)) yields

$$\frac{R}{a} \propto \left(\frac{t}{V} + \frac{JS^2}{t}\right)^{-1/5}. \tag{26}$$

We note that at $t/V \ll JS^2/t$, Eq. (26) gives the same estimate for the radius of a FM metallic droplet, as in [8, 14]

$$\frac{R}{a} \sim \left(\frac{t}{JS^2}\right)^{1/5}.$$

In the opposite limit, where $t/V \gg JS^2/t$, we reproduce the same result as in (22):

$$\frac{R}{a} \sim \left(\frac{V}{t}\right)^{1/5}.$$

Finally, the critical concentration $n_c$ is estimated as

$$n_c \propto \left(\frac{t}{V} + \frac{JS^2}{t}\right)^{3/5}. \tag{27}$$

As a result, also taking the tendency to the phase separation at very small values of $n$ into account [8–11, 14], we arrive at the following phase diagram for the extended model (cf. [11]):

1. At $0 < n < (JS^2/t)^{3/5}$, it corresponds to the phase separation into an FM metal with $n = n' > 0$ embedded into the AFM insulating matrix ($n = 0$). To minimize the Coulomb energy, it may again be favorable to split this metallic region into droplets with the concentration $n'$ and the average radius given by Eq. (26) with $t/V = 0$, each containing one electron and kept apart from one another.

2. At $(JS^2/t)^{3/5} < n < (t/V + JS^2/t)^{3/5} < 1/2$, the system is an FM metal. Of course, we need a window of parameters to satisfy the inequality in the right-hand side. In actual manganites, where $t/V \sim 1/2$–$1/3$ and $JS^2/t \sim 0.1$, these conditions imposed on $n$ are not necessarily satis-

fied. Experiments suggest that this window is present for $La_{1-x}Ca_xMnO_3$, but it is definitely absent for $Pr_{1-x}Ca_xMnO_3$ [11].

3. Finally, at $(t/V + JS^2/t)^{3/5} < n < 1/2$, we have the phase separation in the form of FM metallic droplets inside the AFM charge ordered matrix.

This phase diagram is in a good qualitative agreement with many available experimental results for real manganites [22–25], in particular with the observation of the small-scale phase separation close to the doping 0.5 [26]. We also note that our phase diagram has certain similarities to the phase diagram obtained in [27, 28] for the problem of spontaneous ferromagnetism in doped excitonic insulators.

## 6. CONCLUSIONS

In summary, we have shown that the narrow-band system that has the checkerboard charge ordering at $n = 1/2$ (corresponding to the doping $x = 0.5$) is unstable toward phase separation away from half-filling ($n \neq 1/2$). The system separates into regions with the ideal charge ordering ($n = 1/2$) and other regions where extra electrons or holes are trapped. The simplest form of these metallic regions could be spherical metallic droplets embedded into the charge ordered insulating matrix. Simple considerations allow estimation of the size of these droplets and the critical concentration, or doping $x_c = 1/2 - \delta_c$, at which the metallic phase occupies the entire sample and the charge ordered phase disappears. Taking into account the magnetic interactions does not change these conclusions but somewhat modifies the characteristic parameters of the metallic droplets.

The long-range Coulomb interaction may also modify the results, but we do not expect any qualitative changes. For realistic values of the parameters, the size of metallic droplets is still microscopic (about 10 Å) and the excess charge contained in them is rather small.

The picture obtained corresponds rather well to the known properties of 3D and layered manganites close to (less than) half-doping, $x \leq 1/2$. The percolation picture of transport properties emerging from this treatment is confirmed by the results reported in [7, 15, 22, 24–26]; moreover, the coexistence of ferromagnetic reflections and those of the CE type magnetic structure typical of the charge ordered state at $x = 0.5$ were observed by the neutron scattering [29]. Thus, the general behavior of the underdoped manganites ($x \leq 0.5$) is in a good qualitative agreement with our results.

Our treatment also leads to the same tendency to phase separation (instability of the homogeneous charge ordered phase) for the overdoped mode $x > 0.5$. It is still not clear what would be the second phase in this case. Therefore, we did not concentrate our attention on this case.

Our treatment is also applicable to other systems with charge ordering, such as cobaltites [18] and nickelates [19]. It would be interesting to study them for charge carrier concentrations different from the commensurate "checkerboard" one.

A number of important problems still remain unresolved (the origin of the "in-phase" ordering along the *c*-direction in perovskite manganites, the detailed description of inhomogeneous states in the overdoped mode $x > 1/2$, and the behavior at finite temperatures). Nevertheless, in spite of the simplifications introduced, our model seems to capture the essential physics underlying the interplay between phase separation and charge ordering in transition metal oxides.

## ACKNOWLEDGMENTS

We are grateful to N.M. Plakida and M.S. Mar'enko for stimulating discussions. D.Kh. expresses gratitude to S.-W. Cheong and Y. Moritomo for fruitful discussions of the experimental aspects of the problem. The work was supported by INTAS (grants no. 97-0963 and 97-11954), the Russian Foundation for Basic Research (project nos. 00-02-16255 and 00-15-96570), and by the Russian–Dutch Program for Scientific Cooperation funded by the Netherlands Organization for Scientific Research (NWO). M.Yu.K. acknowledges the support of the Russian President Program (grant no. 96-15-9694). The work of D.Kh. was also supported by the Netherlands Foundation for the Fundamental Research of Matter (FOM) and by the European network OXSEN.

## REFERENCES

1. E. Verwey, Nature **144**, 327 (1939); E. Verwey and P. W. Haayman, Physica (Amsterdam) **8**, 979 (1941).
2. D. I. Khomskii, Preprint No. 105 (Lebedev Institute of Physics, USSR Academy of Sciences, 1969).
3. T. Mutou and H. Kontani, Phys. Rev. Lett. **83**, 3685 (1999).
4. J. van den Brink, G. Khaliullin, and D. Khomskii, Phys. Rev. Lett. **83**, 5118 (1999).
5. G. Jackeli, N. B. Perkins, and N. M. Plakida, Phys. Rev. B **62**, 372 (2000).
6. Z. Jirák, S. Krupička, Z. Šimša, *et al.*, J. Magn. Magn. Mater. **53**, 153 (1985).
7. A. Arulraj, A. Biswas, A. K. Raychaudhuri, *et al.*, Phys. Rev. B **56**, R8115 (1998); M. Uehara, S. Mori, C. H. Chen, and S.-W. Cheong, Nature **399**, 560 (1999).
8. E. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996) [Phys. Usp. **39**, 781 (1996)].
9. A. Moreo, S. Yunoki, and E. Dagotto, Science **283**, 2034 (1999).
10. D. Arovas and F. Guinea, Phys. Rev. B **58**, 9150 (1998).
11. D. I. Khomskii, Physica B (Amsterdam) **280**, 325 (2000).
12. G. S. Uhrig and R. Vlamink, Phys. Rev. Lett. **71**, 271 (1993).
13. S. Mori, C. H. Chen, and S.-W. Cheong, Nature **392**, 473 (1998).

14. M. Yu. Kagan, D. I. Khomskii, and M. V. Mostovoy, Eur. Phys. J. B **12**, 217 (1999).
15. L. P. Gor'kov and V. Z. Kresin, Pis'ma Zh. Éksp. Teor. Fiz. **67**, 934 (1998) [JETP Lett. **67**, 985 (1998)].
16. S. Yunoki, T. Hotta, and E. Dagotto, Phys. Rev. Lett. **84**, 3714 (2000).
17. I. V. Solov'iev and K. Terakura, Phys. Rev. Lett. **83**, 2825 (1999).
18. Y. Moritomo, M. Takeo, X. J. Liu, *et al.*, Phys. Rev. B **58**, R13334 (1998).
19. J. A. Alonso, J. L. García-Múñoz, M. T. Fernández-Díaz, *et al.*, Phys. Rev. Lett. **82**, 3871 (1999).
20. P. Pietig, R. Bulla, and S. Blawid, Phys. Rev. Lett. **82**, 4046 (1999).
21. P. B. Visscher, Phys. Rev. B **10**, 943 (1974).
22. N. A. Babushkina, L. M. Belova, A. N. Taldenkov, *et al.*, J. Phys.: Condens. Matter **11**, 5865 (1999).
23. M. Hennion, F. Moussa, G. Biotteau, *et al.*, Phys. Rev. Lett. **81**, 1957 (1998).
24. G. Allodi, R. De Renzi, G. Guidi, *et al.*, Phys. Rev. B **56**, 6036 (1997).
25. I. F. Voloshin, A. V. Kalinov, S. E. Savel'ev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 157 (2000) [JETP Lett. **71**, 106 (2000)].
26. Y. Moritomo, A. Machidas, S. Mori, *et al.*, Phys. Rev. B **60**, 9220 (1999).
27. L. Balents and C. M. Varma, Phys. Rev. Lett. **84**, 1264 (2000).
28. V. Barzykin and L. P. Gor'kov, Phys. Rev. Lett. **84**, 2207 (2000).
29. R. Kayumoto, H. Yoshizawa, H. Kawano, *et al.*, Phys. Rev. B **60**, 9506 (1999).

*Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 424–428.*
*Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 480–485.*
*Original Russian Text Copyright © 2001 by Agrinskaya, Kozub, Ivanov, Ustinov, Chernyaev, Shamshur.*

# Low-Temperature Hopping Conduction over the Upper Hubbard Band in *p*-GaAs/AlGaAs Multilayered Structures

**N. V. Agrinskaya*, V. I. Kozub, Yu. L. Ivanov, V. M. Ustinov,**
**A. V. Chernyaev, and D. V. Shamshur**

*Ioffe Physicotechnical Institute, Russian Academy of Science, ul. Politekhnicheskaya 26, St. Petersburg, 194031 Russia*
**e-mail: nina.agrins@pop.ioffe.rssi.ru*

Received February 23, 2001

**Abstract**—The low-temperature 2D variable range hopping conduction over the states of the upper Hubbard band is investigated in detail for the first time in multilayered Be-doped *p*-type GaAs/AlGaAs structures with quantum wells of 15-nm width. This situation was realized by doping the layer in the well and a barrier layer close to the well for the upper Hubbard band ($A^+$ centers) in the equilibrium state filled with holes. The conduction was of the Mott hopping type in the entire temperature range (4–0.4 K). The positive and negative magnetoresistance branches as well as of non-Ohmic hopping conduction at low temperature are analyzed. The density of states and the localization radius, the scattering amplitude, and the number of scatterers in the upper Hubbard band are estimated. It is found that the interference pattern of phenomena associated with hopping conduction over the $A^+$ band is qualitatively similar to the corresponding pattern for an ordinary impurity band, but the tunnel scattering is relatively weak. *© 2001 MAIK “Nauka/Interperiodica”.*

## 1. INTRODUCTION

The conduction of 2D structures, including those on *p*-GaAs/AlGaAs, has become an object of intense investigations in recent years. This is due to a concentration-induced transition from the insulator-type temperature dependence to the metal-type dependence observed in purest structures at low temperatures.

It is important to note that the scaling theory of localization predicts an insulator-type behavior for 2D structures at low temperatures and rules out a metal–insulator transition. A few approaches were proposed for solving this problem [1–4]. For example, we proposed earlier a model explaining the observed behavior (including the suppression of the metal-type conduction by a parallel magnetic field). This model presumes a significant role of the conduction channel associated with the states in the upper hubbard band ($D^-$, $A^+$) [4].

In this connection, an analysis of conduction over the upper Hubbard band are of special importance. In the 3D case, in accordance with theoretical calculations, the binding energy of a small charged impurity ($A^+$ or $D^-$) relative to the bottom of the conduction band is quite small, $\varepsilon_- = 0.055\varepsilon_0$ ($\varepsilon_0$ is the binding energy of an isolated impurity). Thus, the conduction over the upper Hubbard band was regarded as noncompetitive as compared to the band conduction. However, conduction over the upper Hubbard band was observed in [5] for asymmetrically stressed Ge : Cu; it was noted that Hubbard bands are separated from each other and from the valence band. In this case, the Hubbard energy amounted to 3.7 meV in weakly doped samples and vanished as the impurity concentration approached the critical value. Since the number of doubly filled states in 3D semiconductors in equilibrium must be much smaller than the number of singly filled states (the only exception is $U^-$ centers), these results can be regarded as indirect.

A different situation may be realized in 2D systems with selective doping, where selective doping and bias voltage applied to the shutter can be used to controllably change the electron concentration in the well, thus changing the ratio of the concentrations of $D^-$ and $D^0$ centers [6]. In narrow quantum wells (for which the wave function scale becomes equal to the size of the well), the values of energy $\varepsilon_-$ and $\varepsilon_0$ increase. It is clear that this increase is stronger for the $D^-$ state than for $D^0$ in view of a considerably larger localization radius of an electron. In the case of an extremely narrow quantum well, the energies $\varepsilon_-$ and $\varepsilon_0$ increase by a factor of 10 and 4, respectively, as compared to the 3D case. In wells of finite width, a situation is possible when the energy of the $D^-$ state is lowered, while the energy of the $D^0$ state preserves its value, which may also reduce the Hubbard energy. This may facilitate the observation of the contribution from the upper Hubbard band and considerably affect the physics of the given systems.

In our previous publication [7], we analyzed a multilayered GaAs/AlGaAs system with a quantum well having a width of the order of 15 nm, which was doped with a Be acceptor impurity for which the localization radius (2 nm) was much smaller than the width of the well. Using selective doping of the barrier layer close to the well, a situation when the upper Hubbard band in equilibrium is filled with holes and conduction takes place over its states was realized in the well also. The Hall effect and the impurity and hopping conduction were investigated in the temperature range 300–1.7 K. The experiments proved that the binding energy of the ground state $A^0$ of acceptors did not change significantly in wells of width 15 nm as compared to the situation in the bulk of the sample, while the binding energy of the $A^+$ centers increases noticeably as compared to the bulk case and amounts to 7 meV (which is five times as large as in the bulk material). This can be explained by the closeness of the size of the well and the hole radius at an $A^+$ center. It should be noted for comparison that the theoretical calculations made on the basis of the Monte Carlo method in the effective-mass approximation, as well as magnetooptical experiments, give the binding energy of a $D^-$ center in 100-nm wells, which is seven times as high a in the bulk case [8].

In the present work, we investigate in detail for the first time the low-temperature 2D variable range hopping conduction realized over the states of the upper Hubbard band in multilayered *p*-GaAs/AlGaAs with a quantum well having a width of the order of 15 nm. The positive and negative branches of magnetoresistance, as well as non-Ohmic hopping conduction at low temperatures, are analyzed. Densities of state and localization radius are estimated as well as the scattering amplitude and the concentration of scatterers in the upper Hubbard band.

## 2. EXPERIMENT

The structures under investigation were grown on semi-insulating GaAs(100) substrates by molecular-beam epitaxy on a Riber 32P device equipped with solid Ga, Al, As, and Be sources. The growth was carried out under As-enrichment conditions at the substrate temperature 580°C. The growth rate was approximately 10 nm/min. The structures contain 10 quantum GaAs wells of width 15 nm, separated by $Al_{0.3}Ga_{0.7}As$ barriers of width 15 nm. Confining $Al_{0.3}Ga_{0.7}As$ layers of thickness 100 nm were deposited in front of the first and behind the last quantum well. Epitaxial growth was terminated by the deposition of coating GaAs layer of thickness 20 nm. In the samples under investigation, the middle region of quantum wells of width 5 nm was doped as well as the $Al_{0.3}Ga_{0.7}As$ barriers, the width of the middle doped region of the barrier being 5 nm. Consequently, the thickness of the undoped spacer layers from both sides of the barrier was 5 nm. We used beryllium as a *p*-type doping impurity introduced at the concentration $1 \times 10^{17}$ atoms/cm$^3$. The contacts with the samples were made by 2-min burning in deposited gold containing 3% Zn at 450°C.

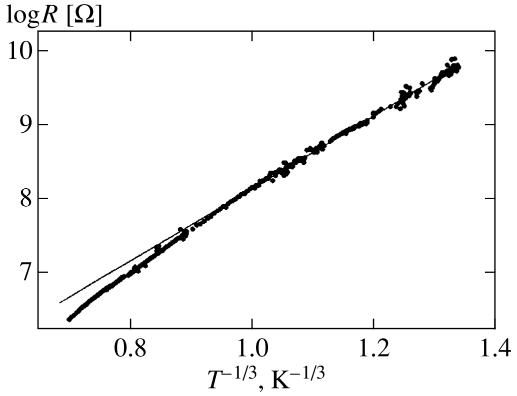

**Fig. 1.** Temperature dependence of the resistance of a multilayered *p*-GaAs/AlGaAs structure in the temperature range 0.4–4 K.

Figure 1 shows the temperature dependences of conductivity. It can be seen that for low voltages (0.1–0.01 V/cm) and currents (10–100 nA), the conduction is of the Ohmic type and its temperature dependence corresponds to hopping conduction with a variable distance of the jump. Since the jump distance at low temperatures is known to exceed the size of the structure, the hopping transport is two-dimensional and is described by the formula

$$\sigma = \sigma_0 \exp(-T_0/T)^{1/3}, \qquad (1)$$

where $T_0$ is the parameter associated with the density of states $N_{E_F}$ at the Fermi level and localization radius $a$:

$$T_0 = C(N_{E_F}a^2)^{-1}, \qquad (2)$$

$C = 13.8$ being a numerical coefficient. Parameter $T_0$ for our sample was estimated at 1500 K.

At the lowest temperature (0.4 K), we recorded the current–voltage characteristics. The magnetoresistance was measured in a constant magnetic field up to 5 T, perpendicular to the plane of the structure (it should be noted that magnetoresistance was virtually equal to zero for our fields parallel to the plane of the structure). The characteristic curves in the region of low fields for several temperatures are presented in Fig. 2, showing a region of negative magnetoresistance. A further increase in the field leads to the emergence of a region of positive magnetoresistance. It can be seen from Fig. 3 that the dependence $\log R(H) \propto H^{2/3}$ is observed in the region of strong fields.

## 3. DISCUSSION OF RESULTS

We assume that low-temperature transport is determined by 2D variable range hopping conduction. The 2D nature of hopping transport follows from the characteristic temperature dependence of conductivity (see Fig. 1) and a considerable anisotropy of magnetoresistance.

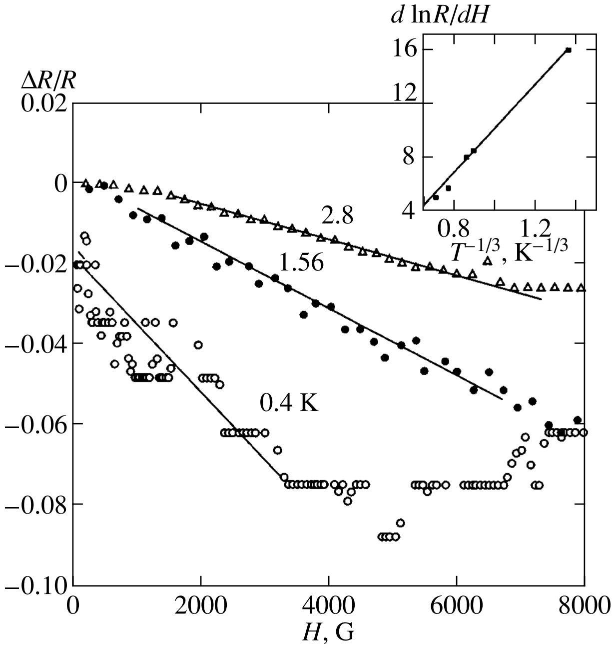

**Fig. 2.** Negative magnetoresistance in weak magnetic fields for various temperatures. The inset shows the temperature dependence of the linear segment of the slope of the negative magnetoresistance ln$R$ on $H$.

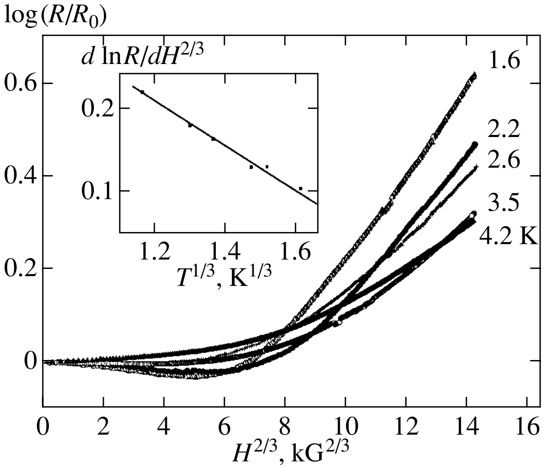

**Fig. 3.** Positive magnetoresistance in strong fields for various temperatures. The inset shows the temperature dependence of the linear segment of the slope of the positive magnetoresistance ln$R$ on $H^{2/3}$.

An analysis of magnetoresistance in the hopping mode makes it possible to estimate the localization radius. For example, such an estimate was obtained in [9] as applied to a 2D hopping transport on the basis of an analysis of the negative magnetoresistance whose detailed theory in the limit of strong scattering was constructed by Raikh and Wessels [10]. It should be noted that it is exactly in the limit of strong scattering that the behavior of negative magnetoresistance is universal and insensitive to the details of the scattering proper [10]. A characteristic feature of such a mode is the universal value of magnetoresistance at a minimum whose presence is determined by the competition between the negative interference magnetoresistance and the positive magnetoresistance determined by the compression of the wave functions of centers in the magnetic field. This value amounts to approximately 20%, which is noticeably higher than in our situation (see Fig. 2).

It should be noted that a certain decrease in the negative magnetoresistance (by half) may be due to the spin effects analyzed in [11] and disregarded in [10]. As a matter of fact, in the case of singly filled centers, the final state for a system for a jump with tunnel scattering depends on the spin state of the intermediate center and interference takes place only for a certain spin orientation of this center. However, for jumps in the upper Hubbard band, when doubly filled centers play the roles of the initial and intermediate states, the spin correlation of electrons at these centers rules out arbitrariness in the orientation of the hopping electron spin and interference always takes place. For this reason, spin effects cannot explain the relatively small value of the negative magnetoresistance observed by us.

Besides, the value of the resistance at the minimum depends on temperature in our experiments, which also contradicts the predictions for the case of strong scattering [10]. These circumstances indicate that the limit of strong scattering is not realized in our system. It should be noted that in accordance with the conclusions drawn in [10], the strong-scattering limit is realized in almost all cases, but the scattering potential was regarded in [10] as a purely Coulomb potential. In our opinion, the weak scattering observed in this situation is in accord with the pattern of hopping transport over filled centers (a hole jumps over $A^0$ centers) for which the potential relative to an "extra" electron is not a purely Coulomb potential and decreases faster with increasing distance from the center. An order-of-magnitude analysis of the situation with weak scattering was carried out by Spivak and Shklovskii [12]. In accordance with their estimates, the 2D negative interference magnetoresistance is defined as

$$\ln\left(\frac{\sigma(H)}{\sigma(0)}\right) = \frac{1}{2\Phi_0}N\mu^2a^2\left(\frac{T_0}{T}\right)^{1/3}H, \qquad (3)$$

where $N$ is the concentration of scatterers in a 2D layer, $\mu$ is the scattering amplitude, and $\Phi_0$ is the magnetic flux quantum. In accordance with this expression, the linear magnetoresistance coefficient depends on temperature as $(T_0/T)^{1/3}$, while Raikh and Wessels [10] predicted the $T^{-1/2}$ dependence for the strong-scattering limit. The inset in Fig. 2 shows the temperature dependence of the corresponding slope determined from experimental data. It can be seen that this dependence indeed corresponds to the weak scattering limit. By comparing this estimate with experimental results, we determined the value of $\mu$ (assuming that the values of

$a = 110$ Å and $T_0$ are known). The corresponding estimate resulted in $\mu \sim 2 \times 10^{-6}$ cm, i.e., $\mu \sim a$. In other words, the scattering amplitude is of the order of the localization radius, which corresponds to the weak scattering limit. Shklovskii and Spivak give the estimate $\mu \sim 1/Na^2$ for scattering at a filled center. In our situation, this gives $\mu \sim 4 \times 10^{-6}$ cm. Taking into account the approximate nature of the estimates, we can say that the theoretical value is in satisfactory agreement with the experimental result.

Another possibility of estimating the localization radius is based on an analysis of the positive contribution to magnetoresistance, which is associated with the compression of wave functions since this contribution is also universal in the region of weak fields.

In accordance with [10], this contribution in the case of hopping conduction with a varying distance of the jump for a constant density of states at the Fermi level is given by

$$\ln\left(\frac{R(H)}{R(0)}\right) = K\left(\frac{T_0}{T}\right)\left(\frac{eHa^2}{ch}\right)^2, \tag{4}$$

where $K = 0.0028$ is a numerical coefficient.

Figure 4 shows the low-temperature ($T = 0.4$ and 4.2 K) positive magnetoresistance, which is a quadratic function of magnetic field. It can be seen that in accordance with formula (4), the derivative of the logarithm of resistance with respect to the square of the magnetic field varies in inverse proportion to temperature (the slope decreases by a factor of 10 as the temperature increases from 0.4 to 4.2 K). The radius of states through which hopping conduction takes place is estimated from this dependence as 110 Å.

In the region of strong fields, an exponential positive magnetoresistance is observed, which is satisfactorily described by the law $\log R(H) \propto H^{2/3}$ (Fig. 3). The derivative of the logarithm of resistance with respect to $H^{2/3}$ decreases upon heating, and the temperature dependence of this quantity is shown in the inset in Fig. 3. The observed behavior is in accord with the predictions made by Shklovskii and Spivak [12], who estimated the contribution of the under-the-barrier scattering of a tunneling electron at localization centers to the positive magnetoresistance. In accordance with their results, the quadratic dependence of magnetoresistance upon an increase of the magnetic field is replaced by the law

$$\ln\frac{R(H)}{R(0)} = \left(\frac{14}{N^{2/3}N(E_F)s^2}\frac{e^2H^2}{c^2h^2}\frac{1}{T}\right)^{1/3}, \tag{5}$$

where $s$ is a certain number and $N$ is the concentration of scattering centers. By comparing the coefficient in this formula with the experimentally determined value (see inset in Fig. 3) and taking into account the estimates of $a$ and $T_0$ obtained above, we determined the concentration of scattering centers $N = 2 \times 10^{16}$ cm$^{-3}$, which is in accord (by the order of magnitude) with the concentration of localized states ($10^{17}$ cm$^{-3}$). Taking into account the dependence of this quantitative estimate on the model-sensitive factor $s$, we can state that the accuracy of our estimate is quite satisfactory, which additionally supports the earlier estimate of the localization radius.

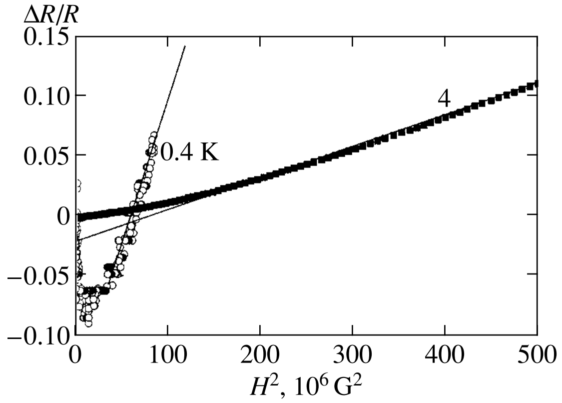

**Fig. 4.** Quadratic positive magnetoresistance for temperatures 0.4 and 4.2 K.

On the other hand, it should be noted that, in accordance with the data on the negative magnetoresistance and the positive magnetoresistance in strong fields, the pattern of interference phenomena in the hopping conduction in the $A^+$ band is at least qualitatively similar to the corresponding pattern for an ordinary impurity band. However, an important difference is the weak under-the-barrier scattering and the absence of spin effects (which are suppressed by the strong spin correlation at doubly filled sites).

Another independent method of estimating the radius involves an analysis of the non-Ohmic component of the low-temperature hopping conductivity. In accordance with the Shklovskii theory of nonlinear hopping conduction [13], nonlinearity starts manifesting itself in electric fields such that $eEL \gg kT$, where $E$ is the electric field, $L$ is the correlation length of the percolation cluster,

$$L \sim \xi^{\nu} r(T) \sim \xi^{\nu+1} a, \tag{6}$$

$r(T)$ is the jump distance, $a$ is the localization radius, $\xi = (T_0/T)^{1/3}$, and $\nu$ is the critical index equal to 1.34 for 2D conduction. The theoretical dependence of conductivity on the electric field has the form

$$\sigma \sim \sigma_0 \exp\left(\frac{eEL}{kT}\right)^{1/(\nu+1)} \sim \sigma_0 \exp\left(\frac{eEL}{kT}\right)^{0.43}. \tag{7}$$

In accordance with this relation, the derivative of the dependence of the logarithm of conductivity with respect to $E^{0.43}$ is equal to $(eL/kT)^{0.43}$. The field dependences of the resistance plotted on the appropriate scale are presented in Fig. 5 ($T = 0.4$ K). The experimentally obtained

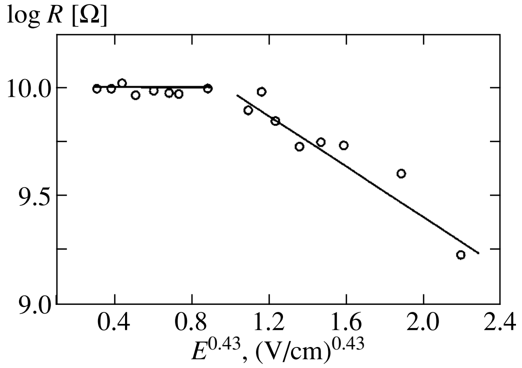

**Fig. 5.** Dependence of resistance on the electric field strength at 0.4 K.

value of $L$ is $3 \times 10^{-4}$ cm. Using relations (6), we obtain $a$ = 110 Å, which is in good agreement with the value of the radius obtained from the expression for the low-temperature magnetoresistance.

Further, we can try to estimate the impurity band width $W$ associated with the $A^+$ states. Assuming that the total number of states in the impurity band is equal to the impurity concentration $N_A = 1.5 \times 10^{11}$ cm$^{-2}$ and determining the density of states at the Fermi level as $N_{E_F} = N_A W = 13.8/(T_0 a)$, we estimate the width of the impurity band as $W$ = 3.5 meV.

## 4. CONCLUSIONS

The analysis of the experimental results obtained unambiguously indicates that the low-temperature transport in the $p$-GaAs/AlGaAs structures under investigation (with doped wells and barriers) is associated with a 2D hopping conduction over the upper Hubbard band. This follows primarily from the experimentally obtained value of binding energy as well as from the value of localization radius estimated from a number of independent experimental values, which is higher than the corresponding value for the ground state of an acceptor in the 3D case approximately by a factor of 4–5.[1] A number of other parameters such as the scattering cross section in the impurity band, the concentration of scattering centers, and the impurity band width, which were estimated on the basis of existing theoretical concepts under the assumption that transport occurs in the upper Hubbard band are also in accord with the model used here. An important conclusion is that the pattern of interference phenomena is qualitatively the same as for jumps in an ordinary impurity band, although quantitative differences are possible. In particular, an analysis of the negative magnetoresistance leads to the conclusion that tunnel scattering is weak. In our opinion, this indicates a non-Coulomb nature of the potential of the corresponding centers, which is expected for jumps over the upper Hubbard band (when intermediate centers are neutral).

Summing up, we emphasize that a situation in which hopping transport in the upper Hubbard band is controllable was realized by us for the first time.

## ACKNOWLEDGMENTS

This work was supported financially by the Russian Foundation for Basic Research (project nos. 00-02-16992 and 00-15-96750).

## REFERENCES

1. S. V. Kravchenko, G. V. Kravchenko, J. E. Furneaux, *et al.*, Phys. Rev. B **50**, 8039 (1994).
2. D. Simmonian *et al.*, Phys. Rev. Lett. **79**, 2304 (1997).
3. B. L. Altshuler and D. L. Maslov, Phys. Rev. Lett. **82**, 145 (1999).
4. V. I. Kozub, N. V. Agrinskaya, S. I. Khondaker, and I. Shlimak, cond-mat/9911450.
5. O. D. Dubon, W. Walukiewicz, J. W. Beeman, and E. E. Haller, Phys. Rev. Lett. **78**, 3519 (1997).
6. S. Huant, S. P. Najda, and B. Etienne, Phys. Rev. Lett. **65**, 1486 (1990).
7. N. V. Agrinskaya, Yu. L. Ivanov, V. M. Ustinov, and D. A. Poloskin, Fiz. Tekh. Poluprovodn. (St. Petersburg) **35**, 571 (2001) [Semiconductors **35**, 550 (2001)].
8. T. Pang and S. G. Louie, Phys. Rev. B **65**, 1635 (1990).
9. D. L. Watson, A. K. Savchenko, E. H. Linfield, *et al.*, Phys. Status Solidi B **205**, 187 (1998).
10. M. E. Raikh and G. F. Wessels, Phys. Rev. B **47**, 15609 (1993).
11. H. L. Zhao, B. Z. Spivak, M. P. Gelfand, and S. Feng, Phys. Rev. B **44**, 10760 (1991).
12. B. I. Shklovskii and B. Z. Spivak, in *Hopping Transport in Solids*, Ed. by M. Pollak and B. Shklovskii (Elsevier, Amsterdam, 1991), p. 271.
13. B. I. Shklovskii, Fiz. Tekh. Poluprovodn. (Leningrad) **13**, 93 (1979) [Sov. Phys. Semicond. **13**, 53 (1979)].

*Translated by N. Wadhwa*

---

[1] It should be emphasized that we obtained quantitative estimates using the numerical coefficients determined earlier for the standard 2D impurity band; strictly speaking, these numerical coefficients for states from the upper Hubbard band may be slightly different.

Journal of Experimental and Theoretical Physics, Vol. 93, No. 2, 2001, pp. 429–434.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 120, No. 2, 2001, pp. 486–492.
Original English Text Copyright © 2001 by Darmanyan, Kobyakov, Lederer.

## MISCELLANEOUS

# Asymmetric Dark Solitons in Nonlinear Lattices[¶]

**S. Darmanyan[a, *], A. Kobyakov[b], and F. Lederer[c]**
[a] *Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, 142190 Russia*
**e-mail: sdarmanyan@yahoo.com*
[b] *School of Optics/CREOL, University of Central Florida, 4000 Central Florida Blvd., Orlando, FL 32816-2700, USA*
[c] *Institute of Solid State Physics and Theoretical Optics, Friedrich-Schiller-Universität Jena, 07743, Jena, Germany*

Received February 16, 2001

**Abstract**—New types of stable discrete solitons are discovered. They represent the first example of asymmetric dark solitons and shock waves with a nonzero background. Both types of solutions exhibit a strong intrinsic phase dynamics. Their domains of existence and criteria of stability are identified. Numerical experiments support the analytical findings. *© 2001 MAIK "Nauka/Interperiodica".*

Numerous recent studies have evidenced that the inherent discreteness of nonlinear systems can qualitatively alter their dynamical behavior compared to their continuous counterparts. Because many physical systems are discrete by definition, these effects attract a steadily increasing interest in various branches of physics (for a detailed overview, see the review papers [1] and references therein). In these studies, particular emphasis was given to stationary, localized structures that are frequently termed as discrete solitons. One option to categorize them is by their degree of localization. Strongly localized solitons (SLSs), where the excitation is resting and involves only a few lattice sites, exhibit properties that originate from the very discreteness of the system [2]. Thus, their behavior differs in many aspects from solutions of related continuous models. It was shown that SLSs can significantly contribute to the heat transfer and other thermodynamic and magnetic effects in solids. Moreover, certain destabilization scenarios can be used for signal processing and switching applications in discrete optical systems such as the coupled waveguide arrays [1–4]. Up to now, various types of SLSs have been reported to exist. Bright [3–7] and dark [8–11] stable SLSs exhibiting interesting new topologies and shapes were identified in various nonlinear evolution equations.

However, similarly to continuum models, all these solutions are (anti-) symmetric and do not exhibit an intrinsic phase dynamics. The existence and stability of asymmetric bright SLSs that are quasi-periodic in time were studied in [12]. In this paper, we reveal that a discreteness may induce new soliton formation mechanisms resulting in the existence of shock waves with two finite backgrounds as well as asymmetric dark solitons. They exhibit a nontrivial intrinsic phase dynamics; i.e., the backgrounds oscillate at two different frequencies, and the transition region is characterized by combinations of these frequencies.

Our model is based on the discrete nonlinear Schrödinger equation (DNLSE), which is among the most prominent model equations in nonlinear physics. Vibron modes in biomolecules, the Heisenberg ferromagnet or Frenkel excitons in a chain with two-level atoms can be mentioned among numerous phenomena described by this equation [1]. Moreover, the DNLSE also describes light propagation in arrays of weakly coupled nonlinear optical waveguides exhibiting Kerr nonlinearity [13]. Recently, the existence and dynamics of discrete solitons in the latter environment were experimentally verified [14].

We consider the DNLSE in the generic form

$$i\dot{\psi}_n + c(\psi_{n-1} + \psi_{n+1}) + \lambda|\psi_n|^2\psi_n = 0, \qquad (1)$$

where $\psi_n$ denotes the amplitude excitation at the $n$th site, $c$ and $\lambda = \pm 1$ are the linear and nonlinear coupling coefficients, respectively, and the upper dot denotes the derivative with respect to the evolution variable $t$.

Traditionally, one seeks a resting solution to (1) with the common frequency $\omega$ in the form $\psi_n = f_n \exp(i\omega t)$, where the localization involves only several lattice sites $n$. In contrast to this conventional ansatz, we search for solutions that are characterized by a combinational frequency. This combinational frequency is determined by interaction of the lattice sites of the localization region with both backgrounds. We show that typical stable SLSs of this type are, e.g., shock waves with two finite backgrounds (Fig. 1a) or asymmetric dark solitons (Fig. 1b). A new family of symmetric dark solitons without the phase jump $\pi$ in the soliton center is also identified in what follows.

The asymmetric SLSs displayed in Fig. 1 have the form

$$\psi = \{\psi_n\} = \{(\ldots, 1, 1, 1)\exp(i\omega_1 t), \psi_{-N}$$
$$\ldots\psi_N, (A, A, A, \ldots)\exp(i\omega_2 t)\},$$

---

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

1063-7761/01/9302-0429$21.00 © 2001 MAIK "Nauka/Interperiodica"

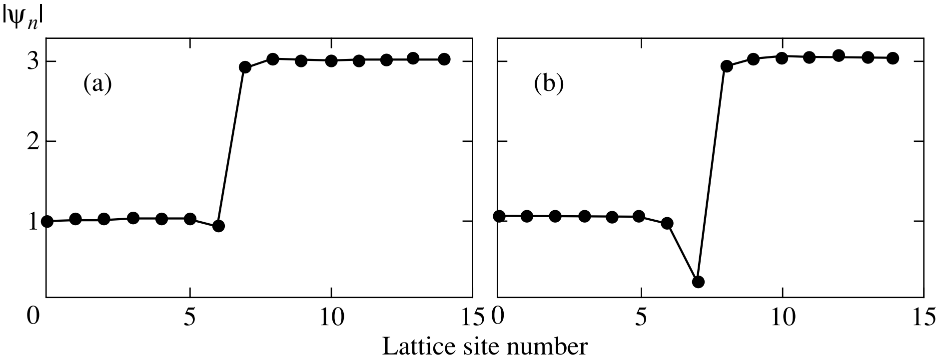

**Fig. 1.** Discrete shock wave with a finite background (a) and asymmetric discrete dark soliton (b).

where the amplitude of the left background is scaled to unity. Strong localization implies $c \ll 1$ [4, 6–11] and a small number $N$ of constituents of the transition region. It is evident from Eq. (1) that the two background frequencies, $\omega_1 = 2c + \lambda$ for $n > -N$ and $\omega_2 = 2c + \lambda A^2$ for $n > N$, do not coincide. For these solutions to exist, the localization between both backgrounds ($n = -N, \ldots, N$) must have the form

$$\psi_n(t) = f_n \exp(i\omega_1 t) + g_n \exp(i\omega_2 t) + m_n(t), \quad (2)$$

where $m_n(t)$ contains an infinite sum of terms with various combinational frequencies of both backgrounds. We follow the conventional terminology [3], assuming that odd (even) SLSs have an odd (even) number of transition sites, and we omit the site $n = 0$ for even modes. As can be seen in the ansatz, we assume unstaggered backgrounds, which requires $\lambda = -1$ for modulationally stable solutions [15]. Because (1) is invariant under the transformation $\lambda \longrightarrow -\lambda$, $t \longrightarrow -t$, $\psi_n \longrightarrow (-1)^n\psi_n$, the results also hold for staggered backgrounds with $\lambda = 1$.

In what follows, we assume that $A$ is real-valued, thus dealing with either the in-phase ($A > 0$) or out-of-phase ($A < 0$) background at $t = 0$. Substitution of (2) into (1) results in a system of equations where in the strong localization limit [2–7], we only keep the rms in the lowest order in the small parameter $c$.

## SHOCK WAVES WITH A FINITE BACKGROUND

We begin with SLSs of the narrowest possible width, namely, with finite background shock waves (Fig. 1a). It is an even SLS with $N = 1$, and therefore, only two sites $n = -1, 1$ constitute the transition region. Within the first-order approximation in $c$, the solution to Eq. (1) is given by

$$\begin{aligned} \psi_{-1} &\approx \left(1 - \frac{c}{2}\right)e^{i\omega_1 t} - \alpha A^3 e^{i\omega_2 t} + \alpha A e^{i(2\omega_1 - \omega_2)t}, \\ \psi_1 &\approx \left(A - \frac{c}{2A}\right)e^{i\omega_2 t} - \alpha e^{i\omega_1 t} + \alpha A^2 e^{i(2\omega_2 - \omega_1)t}, \end{aligned} \quad (3)$$

where

$$\alpha = \frac{c}{(A^2 - 1)^2 - 2c(1 + A^2)}$$

and the oscillation of each site in the transition region is determined by a combination of three frequencies. Other combinations of the background frequencies $\omega_1$ and $\omega_2$ appear only as higher order terms in $c$ and do not significantly contribute to the dynamics of this SLS. Two constraints must be satisfied for solution (3) to be valid, namely, $|\alpha A^3| \ll 1$ for $|A| > 1$ and $|\alpha| \ll |A|$ for $|A| < 1$.

We also mention that the limit $A \longrightarrow 1$ requires taking the second-order terms in $c$ into account. The transformation $\alpha \longrightarrow \alpha(1 \pm c)$, where the respective signs "+" and "–" correspond to the second and third terms in Eq. (3), provides a more accurate solution in this case. Without loss of generality, we consider the case where $|A| \geq 1$, thereby normalizing with respect to the lower background amplitude.

We performed numerical experiments to prove the existence and to probe the robustness of this new SLS. We directly integrated Eq. (1) using solution (3) as the initial condition. The results have shown that the soliton can be easily excited. Moreover, the solution is very robust against rather strong perturbations of the initial conditions. We used a steplike profile

$$f_n = (\ldots, 1, 1, 1, A, A, A, \ldots)$$

for excitation and obtained the robust propagation displayed in Fig. 2. A zoomed picture of the amplitude and phase evolution of the two sites in the transition region is shown in Fig. 3, where an excellent agreement between analytical (Eq. (3)) and numerical results can be recognized. However, this SLS exists only in a restricted domain in the parameter space because for $A$ approaching $A_\pm \longrightarrow 1 \pm \sqrt{c} + c^{3/2}/8$, the approximate solution diverges, see Eq. (3). For example, if $c = 0.08$, then $A_+ \approx 1.29$. Indeed, the numerical integration of Eq. (1) with the steplike initial condition reveals a rapid decay of the initial excitation even for $A = 1.45$ (Fig. 4).

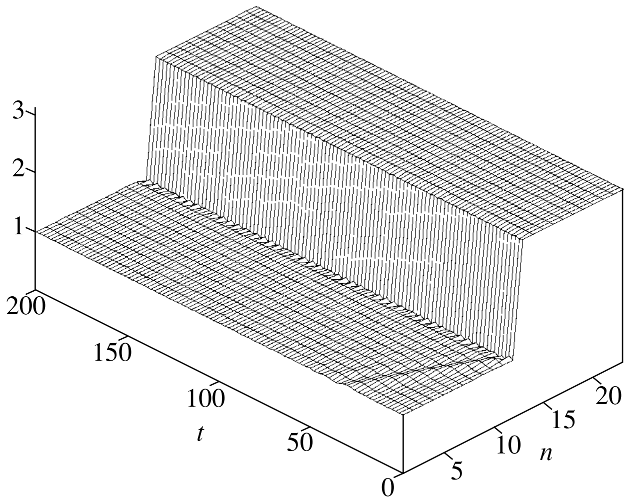

**Fig. 2.** Temporal evolution of the amplitude of a stable shock wave; $A = 3$, $c = 0.08$; step-like excitation.

This behavior can be easily explained by realizing that, e.g., for $n = -1$, the ratio of the amplitudes oscillating at $\omega_2$ and $\omega_1$ amounts to approximately 0.35. Thus, higher-order terms become essential and evoke the SLS decay. If we require that this ratio should be of the order of $c$, we can estimate the SLS robustness domain. The condition $\alpha A^3 \sim c \ll 1$ gives the approximate threshold value of the amplitude $A$ as $A_{th} \approx 1.9 + c$. For $A \gtrsim 1.9$, one can therefore expect a robust SLS behavior that has been confirmed by our numerical simulations.

## ASYMMETRIC DARK SOLITONS

Following the same approach, we can find an odd solution that takes form

$$\begin{aligned}
\psi_{-1} &\approx \left(1-\frac{c}{2}\right)e^{i\omega_1 t},\\
\psi_0 &\approx -ce^{i\omega_1 t} - \frac{c}{A}e^{i\omega_2 t},\\
\psi_1 &\approx \left(A-\frac{c}{2A}\right)e^{i\omega_2 t}.
\end{aligned} \tag{4}$$

To the best of our knowledge, the solution represents the first example of an asymmetric dark soliton (Fig. 1b) exhibiting a strong intrinsic phase dynamics. Numerical solution of Eq. (1) with initial condition (4) proves the robustness of the solution. Although both amplitudes with $n = 0$ are small, the presence of two frequency components is essential, because the backgrounds interact via the excitation at $n = 0$. Precisely

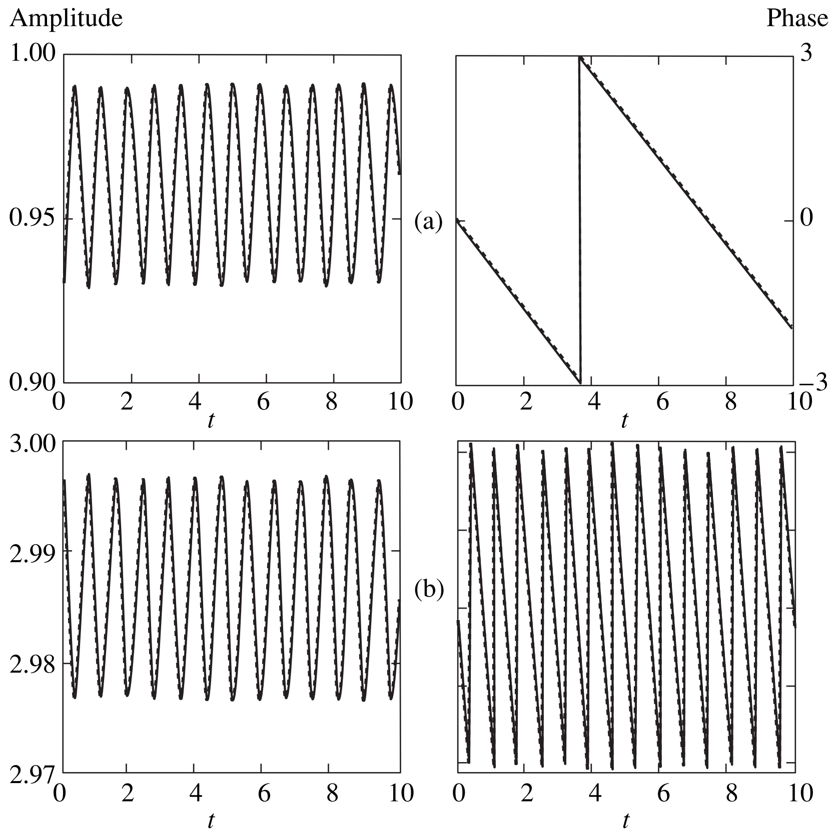

**Fig. 3.** Amplitude oscillations and phase evolution of the excitations in the shock wave transition region: (a) $n = -1$; (b) $n = 1$; the parameters are as in Fig. 2. The solid lines show analytical results (3) and the dashed lines correspond to the numerical integration of Eq. (1).

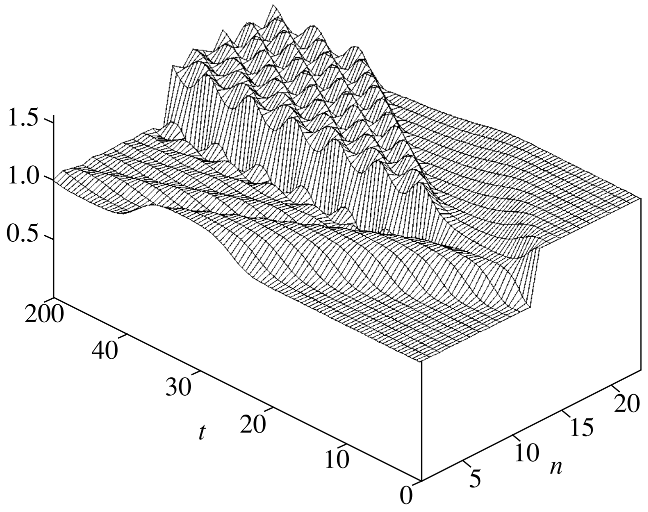

**Fig. 4.** Temporal evolution of the amplitude of an unstable shock wave; $A = 1.45$, $c = 0.08$; step-like excitation.

this interaction affects the stability of the dark soliton. The existence domain of this mode depends on the coupling constant $c$ and the ratio of the background amplitudes $A$. If the backgrounds are separated by more than two lattice constants, wide solitons form. In fact, such solitons can be viewed as two noninteracting discrete front waves, reported recently [7].

Whereas the canonical case where $A = -1$ has been investigated previously and both even and odd dark solutions have been found [8, 9], the case where $A = 1$ provides a new type of solutions, namely, symmetric dark solitons without a phase jump in the center representing a genuine dark soliton with regard to the amplitude. This particular solution has no intrinsic phase dynamics; i.e., all excitations oscillate with frequency $\omega = 2c + \lambda$. There are odd,

$$\psi_{-1} = \psi_1 \approx \left(1 - \frac{c}{2}\right)e^{i\omega t}, \quad \psi_0 \approx -2ce^{i\omega t}, \tag{5}$$

and even,

$$\psi_{-2} = \psi_2 \approx \left(1 - \frac{c}{2}\right)e^{i\omega t}, \quad \psi_{-1} = \psi_1 \approx -ce^{i\omega t}, \tag{6}$$

solutions.

Because asymmetric dark soliton (4) is a fairly exotic object, it is worthwhile to probe its stability by linear stability analysis. Introducing a complex perturbation at each site via $\psi_n \longrightarrow \psi_n + \epsilon_n$ and linearizing Eq. (1) with respect to perturbations $\epsilon_n$, we obtain the set of equations

$$\begin{gathered} i\dot{\epsilon}_{-2} - 2\epsilon_{-2} + c\epsilon_{-1} - \epsilon_{-2}^* e^{2i\omega_1 t} = 0, \\ i\dot{\epsilon}_{-1} - 2(1-c)\epsilon_{-1} + c(\epsilon_0 + \epsilon_{-2}) \\ - (1-c)\epsilon_{-1}^* e^{2i\omega_1 t} = 0, \\ i\dot{\epsilon}_0 + c(\epsilon_{-1} + \epsilon_1) = 0, \\ i\dot{\epsilon}_1 - 2(A^2 - c)\epsilon_1 + c(\epsilon_0 + \epsilon_2) \\ - (A^2 - c)\epsilon_1^* e^{2i\omega_2 t} = 0, \\ i\dot{\epsilon}_2 - 2A^2\epsilon_2 + c\epsilon_1 - A^2\epsilon_2^* e^{2i\omega_2 t} = 0, \end{gathered} \tag{7}$$

where only the sites that belong to the transition region and one site from each background were taken into account. Nevertheless, this set of equations can be easily extended to any number of background sites.

The approach successfully used in studying the stability of bright SLSs [4] cannot be applied here, because the coefficients in Eqs. (7) depend explicitly on the evolution variable. We therefore follow a different procedure to tackle the stability issue of multifrequency localized structures. In doing this, we introduce the Fourier transform of the perturbations,

$$\epsilon_n = \int_{-\infty}^{\infty} F_n(\Omega)e^{i\Omega t}d\Omega, \quad \epsilon_n^* = \int_{-\infty}^{\infty} \Phi_n(\Omega)e^{i\Omega t}d\Omega,$$

where $\Phi_n(\Omega) = F_n^*(-\Omega)$, and rewrite Eqs. (7) in the frequency domain. We eliminate functions $\Phi_0$ and $F_0$ and reduce the total number of equations to eight,

$$(2 + \Omega)F_{-2}(\Omega) + \Phi_{-2}(\Omega - 2\omega_1) - cF_{-1}(\Omega) = 0, \tag{8}$$

$$\begin{gathered} F_{-2}(\Omega) - (\Omega - 4c)\Phi_{-2}(\Omega - 2\omega_1) \\ - c\Phi_{-1}(\Omega - 2\omega_1) = 0, \end{gathered} \tag{9}$$

$$\begin{gathered} -cF_{-2}(\Omega) + \left(2 - 2c - \frac{c^2}{\Omega} + \Omega\right)F_{-1}(\Omega) \\ + (1-c)\Phi_{-1}(\Omega - 2\omega_1) - \frac{c^2}{\Omega}F_1(\Omega) = 0, \end{gathered} \tag{10}$$

$$\begin{gathered} -c\Phi_{-2}(\Omega - 2\omega_1) + (1-c)F_{-1}(\Omega) \\ + \left(2c - \Omega + \frac{c^2}{\Omega + 2 - 4c}\right)\Phi_{-1}(\Omega - 2\omega_1) \\ + \frac{c^2}{\Omega + 2 - 4c}\Phi_1(\Omega - 2\omega_1) = 0, \end{gathered} \tag{11}$$

$$\begin{gathered} -\frac{c^2}{\Omega}F_{-1}(\Omega) + \left(\Omega + 2A^2 - 2c - \frac{c^2}{\Omega}\right)F_1(\Omega) \\ + (A^2 - c)\Phi_1(\Omega - 2\omega_2) - cF_2(\Omega) = 0, \end{gathered} \tag{12}$$

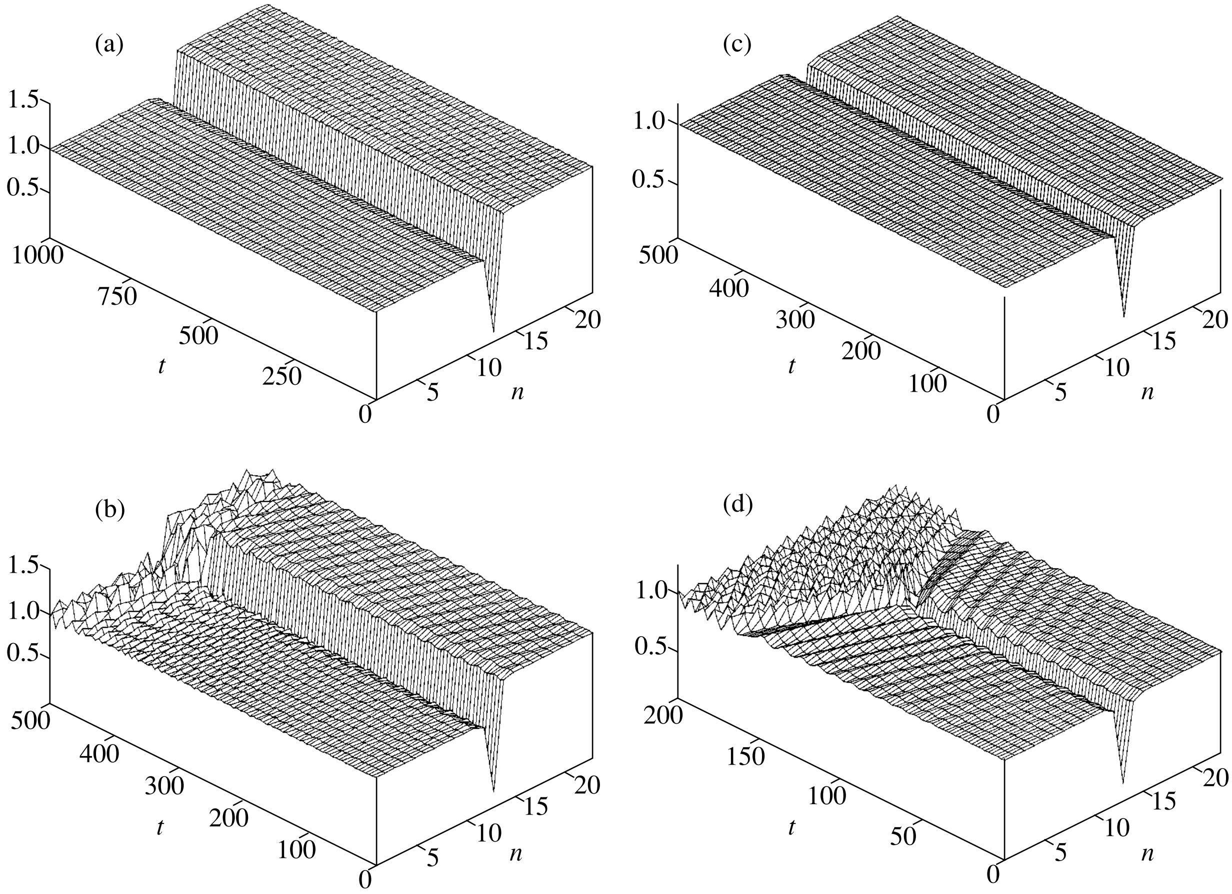

**Fig. 5.** Amplitude evolution of discrete dark solitons: (a) stable asymmetric dark soliton, $A = 1.4$, $c = 0.065$; (b) unstable asymmetric dark soliton, $A = 1.4$, $c = 0.1$; (c) stable symmetric dark soliton, $A = 1$, $c = 0.07$; (d) unstable symmetric dark soliton, $A = 1$, $c = 0.1$.

$$\begin{gathered}\frac{c^2}{\Omega + 2A^2 - 4c}\Phi_{-1}(\Omega - 2\omega_2) + (A^2 - c)F_1(\Omega) \\ -\left(\Omega - 2c - \frac{c^2}{\Omega + 2A^2 - 4c}\right)\Phi_1(\Omega - 2\omega_2) \\ -c\Phi_2(\Omega - 2\omega_2) = 0,\end{gathered} \tag{13}$$

$$\begin{gathered}-cF_1(\Omega) + (\Omega + 2A^2)F_2(\Omega) \\ + A^2\Phi_2(\Omega - 2\omega_2) = 0,\end{gathered} \tag{14}$$

$$\begin{gathered}-c\Phi_1(\Omega - 2\omega_2) + A^2F_2(\Omega) \\ -(\Omega - 4c)\Phi_2(\Omega - 2\omega_2) = 0,\end{gathered} \tag{15}$$

where all functions with shifted arguments must be considered independent. A complete set of these equations contains an infinite number of equations for the functions $F_n(\Omega - 2l\omega_1)$, $F_n(\Omega - 2l\omega_2)$, $\Phi_n(\Omega - 2l\omega_1)$ and $\Phi_n(\Omega - 2l\omega_2)$ with $n = \pm1, \pm2$ and $l = 0, 1, 2, 3, \ldots$. This fact is not surprising because Eqs. (7) explicitly depend on time and, therefore, their solutions contain all harmonics of the background frequencies $\omega_1$ and $\omega_2$.

The terms with denominators in Eqs. (8)–(15) are responsible for higher harmonics. They are of the second order in c and could therefore be omitted. A reason for keeping them is to account for possible resonances that appear as any denominators approaches zero, i.e., as $\Omega \longrightarrow 0$, $\Omega - 2\omega_1 \longrightarrow 0$, $\Omega - 2\omega_2 \longrightarrow 0$. Outside the resonance regions, these terms can be omitted and Eqs. (8)–(15) reduce to two sets of four closed equations allowing the solution of the respective eigenvalue problem. The solution reveals that all eigenvalues are real, i.e., the SLS is stable. Thus, only the resonance regions are potentially responsible for the onset of instability. To treat the set of equations (8)–(15), one needs to close it by truncating to a finite number of equations. To proceed in this way, we note that only the terms $\Phi_1(\Omega - 2\omega_1)$ in (11) and $\Phi_{-1}(\Omega - 2\omega_2)$ in (13) introduce new frequencies into the system. A more thorough analysis of Eqs. (8)–(15) shows that it is not necessary to consider these harmonics in the first-order approximation in $c$ because the amplitudes of these oscillations are of a higher order in $c$. In seeking the instability gain $\mathrm{Im}\Omega_j \sim c$, we can therefore drop these terms. We then obtain eight closed equations with the coefficients that depend nonlinearly on the eigenvalue $\Omega$.

The corresponding eigenvalue problem represents a polynomial of the 11th order possessing complex solutions in some domains of the parameter space $(c, A)$. We found that complex eigenvalues appear for $c > c_{cr1}(A)$. Our analysis also revealed the existence of stability windows for $c_{\mathrm{cr}2n}(A) < c < c_{\mathrm{cr}2n+1}(A)$, where $n = 1, 2, \ldots$ As was shown recently [10], the existence of such windows is due to the finite size of the system used for modeling. The windows tend to disappear with an increasing number of lattice sites. With additional sites taken into account, we indeed observed this phenomenon. We note that the results obtained also hold for symmetric dark soliton (5).

Thus, we conclude that both asymmetric and symmetric dark solitons destabilize provided the linear coupling exceeds the threshold $c = c_{\mathrm{cr}1}(A)$. It is important to note that the value $c_{\mathrm{cr}1}(A)$ slightly depends on both the number $N_s$ of sites regarded for the stability analysis, provided $N_s \geq 5$, and the ratio of the background amplitudes $A$. This value can be calculated with a good accuracy by taking five sites into account. To improve the accuracy, we also considered the case of seven sites involved. The result obtained was $c_{\mathrm{cr}1} \approx 0.085$. A direct numerical integration of Eq. (1) confirms this prediction. Representative examples are displayed in Fig. 5 for $A = 1.4$ (asymmetric dark soliton) and $A = 1$ (symmetric dark soliton without a phase jump). Figures 5a and 5c exhibit stable propagation below the critical coupling ($c < c_{\mathrm{cr}1}$), whereas the solitons decay beyond that threshold ($c = 0.1 > c_{\mathrm{cr}1}$), which is in agreement with the linear stability analysis (Figs. 5b and 5d).

In conclusion, we have shown that new types of solitons, not reported before in the literature, may exist in nonlinear lattices described by the discrete nonlinear Schrödinger equation. These solitons are shock waves with a finite background and asymmetric dark solitons. They are peculiar in that they exhibit a nontrivial intrinsic phase dynamics. Additionally, we found a symmetric dark soliton with the conventional phase dynamics but without a phase jump in the center. A linear stability analysis and numerical experiments revealed the domains of their robust behavior.

## ACKNOWLEDGMENTS

The authors (S.D. and F.L.) gratefully acknowledge the grants from the Deutsche Forschungsgemeinschaft (Bonn) in the framework of Sonderforschungsbereich 196 and INTAS no. 96-0336.

## REFERENCES

1. A. C. Scott, Phys. Rep. **217**, 1 (1992); S. Flach and C. R. Willis, Phys. Rep. **295**, 181 (1998); D. Hennig and G. P. Tsironis, Phys. Rep. **307**, 333 (1999); R. Lai and A. J. Sievers, Phys. Rep. **314**, 147 (1999); F. Lederer, S. Darmanyan, and A. Kobyakov, in *Spatial Solitons*, Ed. by W. Torruellas and S. Trillo (Springer-Verlag, Berlin, 2001) (in press).
2. A. S. Dolgov, Fiz. Tverd. Tela (Leningrad) **28**, 1641 (1986) [Sov. Phys. Solid State **28**, 907 (1986)]; A J. Sievers and S. Takeno, Phys. Rev. Lett. **61**, 970 (1988); J. B. Page, Phys. Rev. B **41**, 7835 (1990).
3. D. Cai, A. R. Bishop, and N. Grönbech-Jensen, Phys. Rev. Lett. **72**, 591 (1994).
4. S. Darmanyan, A. Kobyakov, and F. Lederer, Zh. Éksp. Teor. Fiz. **113**, 1253 (1998) [JETP **86**, 682 (1998)].
5. E. W. Laedke, O. Kluth, and K. H. Spatschek, Phys. Rev. E **54**, 4299 (1996).
6. D. Hennig and H. Gabriel, Phys. Rev. E **57**, 2371 (1998).
7. S. Darmanyan, A. Kobyakov, F. Lederer, and L. Vázquez, Phys. Rev. B **59**, 5994 (1999); A. Kobyakov, S. Darmanyan, T. Pertsch, and F. Lederer, J. Opt. Soc. Am. B **16**, 1737 (1999).
8. Yu. S. Kivshar, W. Krolikowski, and O. Chubykalo, Phys. Rev. E **50**, 5020 (1994).
9. V. V. Konotop and S. Takeno, Phys. Rev. E **60**, 1001 (1999); V. V. Konotop and M. Salerno, Phys. Rev. E **56**, 3611 (1997).
10. M. Johansson and Yu. S. Kivshar, Phys. Rev. Lett. **82**, 85 (1999).
11. A. Kobyakov, S. Darmanyan, F. Lederer, and E. Schmidt, Opt. Quantum Electron. **30**, 795 (1998).
12. M. Johansson and S. Aubry, Nonlinearity **10**, 1151 (1997).
13. D. N. Cristodoulides and R. I. Joseph, Opt. Lett. **13**, 794 (1988).
14. H. S. Eisenberg, Y. Silberger, R. Morandotti, *et al.*, Phys. Rev. Lett. **81**, 3383 (1998); R. Morandotti, U. Peschel, J. S. Aitchison, *et al.*, Phys. Rev. Lett. **83**, 2726 (1999); U. Peschel, R. Morandotti, J. S. Aitchison, *et al.*, Appl. Phys. Lett. **75**, 1348 (1999).
15. Yu. Kivshar and M. Peyrard, Phys. Rev. A **46**, 3198 (1992).