Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 649–657.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 723–732.
Original Russian Text Copyright © 2003 by Podolsky.

GRAVITATION, ASTROPHYSICS

Singularity Structure in Veneziano's Model

D. I. Podolsky*
Landau Institute for Theoretical Physics, Russian Academy of Sciences, ul. Kosygina 2, Moscow, 119334 Russia
**e-mail: podolsky@itp.ac.ru*

Received April 24, 2003

Abstract—We consider the structure of the cosmological singularity in Veneziano's inflationary model. The problem of choosing initial data in the model is shown to be unsolved—the spacetime in the asymptotically flat limit can be filled with an arbitrary number of gravitational and scalar field quanta. As a result, the universe acquires a domain structure near the singularity, with an anisotropic expansion of its own realized in each domain. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

An effective inflationary model is known to be very difficult to construct in the low-energy approximation of the string theory [1]. The effective potentials of the scalar field responsible for the inflationary pattern of the cosmological dynamics that arise in attempting to solve the problem do not ensure the satisfaction of the slow-roll condition $|\dot{H}| \ll H^2$. Thus, the Friedmann decelerating expansion is typical of this effective gravitational theory even in the presence of a scalar field.

Actually, this implies that the problem cannot be solved by a brute-force method: new ideas based on the nontrivial low-energy spectrum of the string theory or the nonperturbative effects arising in this theory should be invoked to obtain an inflationary scenario in terms of the string theory. One might probably expect the field dynamics at the inflationary phase to be also nontrivial.

The appearance of Veneziano's paper [2] may be considered to be the birthing of one of these nontrivial scenarios. The basic idea of this study is as follows.

Let us consider the low-energy effective action of an arbitrary theory of closed superstrings [3], with our analysis being restricted to the field problem with zero vacuum averages of the fermion fields, the *R*–*R*-sector fields, and the antisymmetric field $B_{\mu\nu}$. The sector left after this projection has the same structure for any superstring theory,[1] and the corresponding effective action (in dimensionless units) is

$$S = -\int \sqrt{-g} d^{d+1} x e^{-\phi}(R + g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi). \tag{1}$$

Here, *d* is the dimension of space. We will seek a spatially homogeneous classical solution to the equations of motion derived by varying this action. To this end, as usual, we should set

$$g_{\mu\nu} = (1, -a^2(t)\delta_{ij}), \tag{2}$$

$$\phi = \phi(t). \tag{3}$$

Substituting (2) and (3) into (1), we easily find that

$$S = -\int d^{d+1} x a^d e^{-\phi}(\dot{\phi}^2 - 2dH\dot{\phi} + d(d-1)H^2). \tag{4}$$

As can easily be seen, this action is invariant with respect to the field transformations

$$a(t) \longrightarrow \frac{1}{a(t)}, \tag{5}$$

$$\phi \longrightarrow \phi - 2d\ln a, \tag{6}$$

$$t \longrightarrow -t. \tag{7}$$

Therefore, the equations of motion following from (4) have a classical solution that describes the accelerating expansion corresponding to inflation at $-\infty < t < -0$ and the Friedmann decelerating expansion at $0 < t < \infty$:

$$\begin{gathered} a_+(t) = \left(\frac{t}{t_0}\right)^{1/\sqrt{d}}, \\ \phi = (\sqrt{d} - 1)\ln\frac{t}{t_0}, \quad t > 0, \end{gathered} \tag{8}$$

$$\begin{gathered} a_-(-t) = \left(-\frac{t}{t_0}\right)^{-1/\sqrt{d}}, \\ \phi = -(\sqrt{d} + 1)\ln\left(-\frac{t}{t_0}\right), \quad t < 0. \end{gathered} \tag{9}$$

The physical meaning of this solution is as follows. The

---

[1] Recall that there are four distinct theories of closed superstrings in ten-dimensionality: *IIA*, *IIB*, heterotic *SO*(32), and heterotic E_8.

universe is initially an asymptotically flat world in the sense that the Riemann tensor components tend to zero as $t \longrightarrow -\infty$. In addition, this world is absolutely cold—it contains no clustered matter. Starting from this maximally symmetric state, the universe undergoes superinflation (the prefix “super” implies that $\dot{H} > 0$) at $-\infty < t < 0$ and gives way to decelerating Friedmann expansion at $t > 0$, which corresponds to the transition from the superinflation branch to the dual branch.

Significantly, as $t \longrightarrow -0$, the spacetime curvature tends to infinity. Therefore, sooner or later, we will go outside the validity range for the low-energy approximation of the string theory and action (1) in our solution.

The invariance of action (4) with respect to transformations (5)–(7) was called SF duality.[2] If it holds not only for the low-energy approximation of the string theory but also for the total nonperturbative Green–Schwartz action, then a solution of type (8) and (9) actually corresponds to the saddle point in the complete field problem; i.e., it describes the cosmological dynamics in this theory. However, since the SF duality itself is an essentially nonperturbative effect, we cannot ascertain whether it exists in the nonperturbative string theory in the low-energy approximation. Gasperini and Veneziano [4] argued that the nonperturbative SF duality does hold.

The most interesting and critical (from the viewpoint of the scenario) point on the time axis is $t = 0$ at which the cosmological singularity is reached.[3]

Our objective is to elucidate the pattern of field dynamics near the singularity and to check within the framework of the low-energy approximation whether allowance for the fluctuations of the classical trajectory (8) and (9) leads to a general softening of the singularity.

2. AN ANISOTROPIC SOLUTION IN VENEZIANO’S MODEL

The equations of motion that follow from the four-dimensional theory with the action

$$S = -\frac{1}{\lambda_s^2}\int\sqrt{-g}d^4xe^{-\phi}(R + (\partial\phi)^2) \qquad (10)$$

(here, we introduced the constant λ_s to reduce the action to dimensionless form) are

$$-D_\mu D_\nu\phi + \frac{1}{2}g_{\mu\nu}(D\phi)^2 = R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R, \qquad (11)$$

$$D_\mu D^\mu\phi = (\partial\phi)^2. \qquad (12)$$

Analysis of these equations can be greatly simplified if we note that the theory described by action (10) is conformally equivalent to the general theory of relativity with the scalar field

$$S = -\frac{1}{\lambda_s^2}\int\sqrt{-g}d^4x\left(R - \frac{1}{2}(\partial\phi)^2\right). \qquad (13)$$

Indeed, if the metric $g_{\mu\nu(E)}$ is the solution of the Einstein equations, then $g_{\mu\nu(S)} = g_{\mu\nu(E)}\exp(\phi)$ is the solution of Eqs. (11) and (12). The reverse is also true (below, we call $g_{\mu\nu(E)}$ and $g_{\mu\nu(S)}$ the metrics in the Einstein and string frames, respectively).

Let us try to go beyond the scope of the homogeneous problem and to construct an exact solution that corresponds to a classical strong gravitational wave propagating against a homogeneous background and strong inhomogeneous scalar field perturbations in the Einstein frame. Since, in general, this solution, clearly, cannot be found, we restrict our analysis to the quasi-two-dimensional problem. Let all metric components and the scalar field depend on the coordinates t and x alone. This problem can be solved completely (see, e.g., [6]), and the answer is the axisymmetric Einstein–Rozen metric.

We seek the corresponding solution in the form

$$ds_{(E)}^2 = e^{2A}dt^2 - e^{2C}dx^2 - e^{2B}(e^\gamma dy^2 + e^{-\gamma}dz^2), \qquad (14)$$

where A, B, C, γ, and ϕ are functions of t and x alone. Since the Einstein equations are invariant with respect to the gauge transformations $\tilde{t} = \tilde{t}(t, x)$ and $\tilde{x} = \tilde{x}(t, x)$, we may set $g_{00} = -g_{11}$ and $g_{01} = 0$, i.e., $A(t, x) = C(t, x)$.

The Einstein equations impose the following constraints on the functions A, B, γ, and ϕ:

$$A'' - \ddot{A} - 2\ddot{B} + 2\dot{A}\dot{B} + 2A'B' - 2(\dot{B})^2 - \frac{(\dot{\gamma})^2}{2} = \frac{1}{2}(\dot{\phi})^2, \qquad (15)$$

$$-2\dot{B}' + 2A'\dot{B} + 2B'\dot{A} - 2B'\dot{B} - \frac{\dot{\gamma}\gamma'}{2} = \frac{1}{2}\phi'\dot{\phi}, \qquad (16)$$

$$\begin{aligned}&\ddot{A} - A'' - 2B'' + 2A'B' + 2\dot{A}\dot{B}\\ &- 2(B')^2 - \frac{(\gamma')^2}{2} = \frac{1}{2}(\phi')^2,\end{aligned} \qquad (17)$$

$$\ddot{B} + 2(\dot{B})^2 = B'' + 2(B')^2, \qquad (18)$$

$$\ddot{\gamma} + 2\dot{B}\dot{\gamma} = \gamma'' + 2B'\gamma'. \qquad (19)$$

Finally, the equation of motion for the field ϕ appears as

$$\ddot{\phi} + 2\dot{B}\dot{\phi} = \phi'' + 2B'\phi'. \qquad (20)$$

---

[2] SF stands for the scale factor.

[3] The singularity may be smoothed out in the nonperturbative string theory. The validity of this assumption is closely related to the possibility of solving the problem of passage through the singularity from superinflation to decelerating expansion [5] (see also [4]).

We can easily find from (18) that

$$B = \frac{1}{2}\ln(f_1(\xi) + f_2(\eta)), \tag{21}$$

where $\xi = t - x$, $\eta = t + x$, and $f_{1,2}$ are arbitrary functions of their argument. Let us now use the remaining gauge invariance with respect to the transformations $\tilde{\xi} = H_1(\xi)$ and $\tilde{\eta} = H_2(\eta)$ and assume that

$$B = \frac{1}{2}\ln\left(-\frac{t}{t_i}\right).$$

Equations (19) and (20) can then be easily solved:

$$\begin{aligned}\phi &= \psi\ln\left(-\frac{t}{t_i}\right)\\ &+ \sum_k [c_{1k}J_0(kt) + c_{2k}N_0(kt)]e^{ikx} + \text{c.c.},\end{aligned} \tag{22}$$

$$\begin{aligned}\gamma &= \beta\ln\left(-\frac{t}{t_i}\right)\\ &+ \sum_k [c_{3k}J_0(kt) + c_{4k}N_0(kt)]e^{ikx} + \text{c.c.}\end{aligned} \tag{23}$$

An expression for $A(t, x)$ can be derived by using the equations

$$\dot{A} = \frac{t}{4}\left((\phi')^2 + (\dot{\phi})^2 + (\gamma')^2 + (\dot{\gamma})^2 - \frac{1}{t^2}\right), \tag{24}$$

$$A' = \frac{t}{2}(\phi'\dot{\phi} + \gamma'\dot{\gamma}). \tag{25}$$

It is convenient to separate the function $A(t, x)$ into homogeneous and inhomogeneous parts. The former can be easily derived from Eq. (24):

$$\begin{aligned}A_{\text{hom}} &= \frac{1}{4}(\psi^2 + \beta^2 - 1)\ln(-t/t_i)\\ &+ \frac{1}{4}\sum_k \Bigg\{ c_{1k}c_{1k}^{+}\frac{(kt)^2}{2}[J_1^2(-kt) - J_0(-kt)J_2(-kt)]\\ &+ (c_{2k}c_{1k}^{+} + c_{1k}c_{2k}^{+})\frac{(kt)^2}{4}[2J_1(-kt)N_1(-kt)\\ &- J_2(-kt)N_0(-kt) - J_0(-kt)N_2(-kt)]\\ &+ c_{2k}c_{2k}^{+}\frac{(kt)^2}{2}[N_1^2(-kt) - N_0(-kt)N_2(-kt)] \Bigg\}\\ &+ \frac{1}{4}\sum_k \Bigg\{ c_{1k}c_{1k}^{+}\frac{(kt)^2}{2}[J_0^2(-kt) + J_1^2(-kt)]\\ &+ c_{2k}c_{2k}^{+}\frac{(kt)^2}{2}[J_0^2(-kt) - J_1^2(-kt)]\\ &+ (c_{1k}c_{2k}^{+} + c_{2k}c_{1k}^{+})\frac{(kt)^2}{4}[2J_0(-kt)N_0(-kt)\\ &- J_1(-kt)N_{-1}(-kt) - J_{-1}(-kt)N_1(-kt)] \Bigg\}\\ &+ (1, 2) \longrightarrow (3, 4).\end{aligned} \tag{26}$$

The inhomogeneous contribution to the function $A(t, x)$ is easier to determine from Eq. (25):

$$\begin{aligned}A_{inh} &= \psi\sum_k [c_{1k}J_0(-kt) + c_{2k}N_0(-kt)]e^{ikx}\\ &+ \sum_k \frac{klt}{k+l}e^{i(k+l)x}[c_{1k}J_1(-kt) + c_{2k}N_1(-kt)]\\ &\times (c_{1l}J_0(-lt) + c_{2l}N_0(lt))\\ &+ \sum_{k, l, k \neq l} \frac{klt}{k-l}[c_{1k}J_1(-kt) + c_{2k}N_1(-kt)]\\ &\times [c_{1l}J_0(-lt) + c_{2l}N_0(lt)] + \text{c.c.}\\ &+ (1, 2, \psi) \longrightarrow (3, 4, \beta).\end{aligned} \tag{27}$$

Let us first consider the homogeneous limit ($c_{\alpha k} = 0$, $\forall\alpha, k$). In the string frame, the spacetime metric is

$$\begin{aligned}ds^2 &= \left(-\frac{t}{t_i}\right)^{\frac{1}{2}(\psi^2 + \beta^2 - 1) + \psi_1}(dt^2 - dx^2)\\ &- \left(-\frac{t}{t_i}\right)^{1 + \psi_1 + \gamma_1}dy^2 - \left(-\frac{t}{t_i}\right)^{1 + \psi_1 - \gamma_1}dz^2.\end{aligned} \tag{28}$$

The scalar curvature corresponding to this metric is

$$R = -\frac{\psi^2}{t^2}\left(-\frac{t}{t_i}\right)^{(1 - \beta^2 - 2\psi - \psi^2)/2}. \tag{29}$$

It tends to zero as $t \longrightarrow -\infty$ for any point in parametric space (ψ, β). If $\beta = 0$ and $\psi^2 = 3$, then the spacetime (the background spacetime in the inhomogeneous problem) is isotropic. In this case, metric (14) is identical to Veneziano's solution, which is asymptotically equivalent to Minkowski's flat universe for $t \longrightarrow -\infty$. Below, precisely this choice of parameters will be of particular interest to us.

3. THE LIMIT OF ASYMPTOTICALLY FLAT SPACETIME: QUANTIZATION

Below, we will see that the modes in expressions (22) and (23) can be interpreted as dilatons and gravitons propagating against the background of curved spacetime. We choose the initial conditions as follows: we specify a sufficiently large time t_i [this time appears in (22), (23), and (26)] and discard all the modes that do not satisfy the condition $k|t_i| \gg 1$; i.e., we neglect the modes with a wavelength longer than the cosmological horizon in the initial state. However, the existence of these modes is in conflict with causality, unless the initial state itself defined in this way arose from inflation. Below, by the limit $t \longrightarrow -\infty$, we mean all t such that $|t| \gg |t_i|$. In this case, the following asymptotics is possible for all the modes without exception:

$$J_0(-kt) \approx \sqrt{\frac{2}{\pi(-kt)}}\cos\left(-kt - \frac{\pi}{4}\right),$$
$$N_0(-kt) \approx \sqrt{\frac{2}{\pi(-kt)}}\sin\left(-kt - \frac{\pi}{4}\right). \tag{30}$$

It thus follows that the "correct" modes[4] are

$$\frac{1}{2}e^{i\pi/4}H_0^{(1)}(-kt)e^{-ikx}$$

and

$$\frac{1}{2}e^{i\pi/4}H_0^{(1)+}(-kt)e^{ikx},$$

which correspond to the substitutions

$$b_{1k} = e^{i\pi/4}(c_{1k} - ic_{2k}), \quad b_{2k} = e^{i\pi/4}(c_{1k}^+ - ic_{2k}^+). \tag{31}$$

Let us now turn to quantization. We will construct the scalar product $\langle\phi, \phi\rangle$ (for the field γ corresponding to a graviton, the procedure is similar). To this end, we introduce an Abelian gauge field in such a way that the Lagrangian of the theory is

$$\sqrt{-g}L = \frac{1}{\lambda_s^2}\int dV\sqrt{-g}e^{-\phi}$$
$$\times[g^{\mu\nu}(\partial_\mu - iA_\mu)\phi(\partial_\nu - iA_\nu)\phi] \tag{32}$$

and then calculate $\sqrt{-g}\,\delta L/\delta A_0$. (Here, the integration of dV is performed over a spacelike hypersurface on which the result does not depend because of the validity of the Gauss theorem; we assume that only the background contributes to $\sqrt{-g}\,e^{-\phi}$, which corresponds to the assumption that the quanta propagate against a homogeneous background and that their inverse effect on the background metric may be disregarded.) Hence, we can easily find that

$$\langle\phi_1, \phi_2\rangle = -\frac{i}{\lambda_s^2}$$
$$\times\int(\phi_1\partial_\mu\phi_2^+ - \phi_2^+\partial_\mu\phi_1)n^\mu e^{-\phi}\sqrt{-g}dV. \tag{33}$$

It is convenient to take t = const as the hypersurface over which the integration is performed. In that case,

$$\langle\phi_1, \phi_2\rangle = -i\frac{L^2}{\lambda_s^2}\int(\phi_1\partial_\mu\phi_2^+ - \phi_2^+\partial_\mu\phi_1)$$
$$\times n^\mu\left(-\frac{t}{t_i}\right)^{1/2(\psi^2 + \beta^2 + 1) + \psi}dx, \tag{34}$$

where L is the linear size of the space. To correctly normalize the modes

$$u_k = \frac{1}{2}e^{i\pi/4}H_0^{(1)}(-kt)e^{ikx},$$

we calculate the commutator $[b_k, b_{k'}^+]$. For $t \longrightarrow -\infty$, we have

$$\langle u_k, u_{k'}\rangle = \frac{2}{t}e^{2A + 2B + \phi}\delta_{kk'} = F(k, t)\delta_{kk'}. \tag{35}$$

Hence, assuming that

$$[\phi(t, x), \phi(t, x')] = 0, \quad [\pi(t, x), \pi(t, x')] = 0, \tag{36}$$

$$[\phi(t, x), \pi(t, x')] = -\frac{i}{L^2}\delta(x - x'), \tag{37}$$

where

$$\pi(t, x) = \frac{\delta\sqrt{-g}L}{\delta\dot{\phi}} = -\frac{2}{\lambda_s^2}\sqrt{-g}e^{-\phi}g^{00}\dot{\phi},$$

we can easily find that

$$[b_k, b_{k'}^+] = \frac{\pi\lambda_s^2 t_i}{2L^3}\delta_{kk'} \tag{38}$$

and the correctly (half-quantum) normalized modes for $t \longrightarrow -\infty$ are

$$u_k = \sqrt{\frac{-\lambda_s^2 t_i}{L^3 kt}}e^{-ik(x - t)}.$$

---

[4] That is, modes corresponding to the positive frequency solutions.

To physically interpret these quantum modes, we calculate $\langle T_{00}\rangle$. In the Einstein equations, the (0, 0) component is

$$4\dot{A}\dot{B} - 2\ddot{B} - 2\dot{B}^2 = \frac{1}{2}[(\dot{\phi})^2 + (\phi')^2 + (\dot{\gamma})^2 + (\gamma')^2] = T_{00} \tag{39}$$

(we carried over γ to the right and now interpret this part of the metric as the contribution of gravitons). If the paired correlators are assumed (in this case, it does not matter whether we consider the amplitudes of the modes $b_{\alpha k}$ to be classical but randomly distributed Gaussian variables or operators) to be

$$\langle b_{\alpha k} b^{+}_{\beta l}\rangle = n_{\alpha}(k)\delta_{\alpha\beta}\delta_{kl}, \quad \langle b_{\alpha k} b_{\beta l}\rangle = 0, \tag{40}$$

$$n_1(k) = n_2(k), \quad n_3(k) = n_4(k) \tag{41}$$

(the physical meaning of this condition is that the flux of rightward-propagating quanta is equal to the flux of leftward-propagating quanta), then

$$\langle T_{00}\rangle = \langle\hat{H}\rangle = \sum_k \frac{\lambda_s^2 k t_i}{L^3 t}(n_1(k) + n_3(k)). \tag{42}$$

Thus, we see that for $t \longrightarrow -\infty$, the asymptotically flat world is filled with gravitational and scalar field quanta, which are naturally called gravitons and dilatons, respectively. Their energy density tends to zero as $t \longrightarrow -\infty$, which corresponds to a usual redshift.

Finally, let us find out what an observer moving inertially with respect to the cosmologically distinguished frame of reference (in which the background spacetime is homogeneous and isotropic) will see [7]. As the detector for photons of the field ϕ, we consider a type of matter for which the field ϕ enters into the effective gravitational constant. Let $|E_0\rangle$ and $|E\rangle$ be the quantum-mechanical states of the detector before and after the detector has interacted with the field ϕ. The perturbation of the Hamiltonian for the detector corresponding to this interaction is

$$V = \lambda d(\tau)\phi[x(\tau)]\left(-\frac{t}{t_i}\right)^{\psi}, \tag{43}$$

where $x(\tau)$ is the trajectory along which the detector moves, and $d(\tau)$ is the monopole moment operator for the detector. We can then easily see that the transition probability of the detector excited by quanta of the field ϕ from the initial state to all the possible states is

$$w = \lambda^2\left(-\frac{t}{t_i}\right)^{2\psi}\sum_E |\langle E|d(0)|E_0\rangle|^2 F(E - E_0). \tag{44}$$

The first factor under the summation sign describes the detector selectivity, and the second factor describes the response function, which is interest to us. We calculate it by assuming that the detector moves inertially with respect to the cosmologically distinguished frame of references (i.e., $x = vt/\sqrt{1 - v^2}$):

$$\begin{aligned} F(E) &= \int_a^b d\tau \int_a^b d\tau' \exp[iE(\tau - \tau')]G(x(\tau), x(\tau')) \\ &= \frac{\lambda_s^2}{L^3}\sqrt{1 - v^2}\,\tau_i^{1/\kappa}\int_{1/t_i}^{\infty}\frac{dk}{k}n_1(k)\int_a^b \frac{d\tau}{\tau^{1/2\kappa}} \\ &\times \exp\left\{-ik\sqrt{\frac{1 - v}{1 + v}}\,t_i\left(\frac{\tau}{\tau_i}\right)^{1/\kappa} + iE\tau\right\} \\ &\times \int_a^b \frac{d\tau'}{(\tau')^{1/2\kappa}}\exp\left\{ik\sqrt{\frac{1 - v}{1 + v}}\,t_i\left(\frac{\tau'}{\tau_i}\right)^{1/\kappa} - iE\tau'\right\} \\ &+ \left(\sqrt{\frac{1 - v}{1 + v}} \longrightarrow \sqrt{\frac{1 + v}{1 - v}}\right). \end{aligned} \tag{45}$$

Here, we use the following notation:

$$\kappa = \frac{1}{4}(\psi^2 + \beta^2 + 3) + \psi, \quad \tau_i = \frac{t_i}{\kappa},$$

$$t = \frac{t_i}{\sqrt{1 - v^2}}\left(\frac{\tau}{\tau_i}\right)^{1/\kappa},$$

t has the meaning of world time, and τ is the proper time of the detector.

Thus, calculating the response function $F(E)$ reduces to taking the integral

$$\begin{aligned} I(k, E) &= \int_a^b \frac{d\tau}{(\tau')^{1/2\kappa}} \\ &\times \exp\left\{-ik\sqrt{\frac{1 - v}{1 + v}}\,t_i\left(\frac{\tau}{\tau_i}\right)^{1/\kappa} + iE\tau\right\}. \end{aligned} \tag{46}$$

If the detector moves moderately fast and if the condition $a, b \gg \tau_i$ is satisfied, then integral (46) can be taken by the stationary-phase method:

$$I(k, E) \approx \frac{i\kappa E^{1/2\kappa - 1}}{kt_i}\sqrt{\frac{1 + v}{1 - v}}$$

$$\times\left\{\frac{\exp(iEb)}{(Eb)^{1/2\kappa}}\frac{\exp\left\{-ikt_i\sqrt{\frac{1-v}{1+v}}\left(\frac{b}{\tau_i}\right)^{1/\kappa}\right\}}{(b/\tau_i)^{1/\kappa-1}}\right.$$

$$\left.-\frac{\exp(iEa)}{(Ea)^{1/2\kappa}}\frac{\exp\left\{-ikt_i\sqrt{\frac{1-v}{1+v}}\left(\frac{a}{\tau_i}\right)^{1/\kappa}\right\}.}{(b/\tau_i)^{1/\kappa-1}}\right\}. \tag{47}$$

The final form of the detector response function $F(E)$ is

$$F(E) = \frac{\lambda_s^2}{L^3}\sqrt{1-v^2}\tau_i^{1/\kappa}\int\limits_{1/t_i}^{\infty}\frac{dkn_1(k)}{k^3E^2}$$

$$\times\left\{\left[\frac{\kappa\cos(Eb-\xi k(Eb)^{1/\kappa})}{\xi(Eb)^{1/\kappa-1}b^{1/2\kappa}}-\frac{\kappa\cos(Ea-\xi k(Ea)^{1/\kappa})}{\xi(Ea)^{1/\kappa-1}a^{1/2\kappa}}\right]^2\right.$$

$$\left.+\left[\frac{\kappa\sin(Eb-\xi k(Eb)^{1/\kappa})}{\xi(Eb)^{1/\kappa-1}b^{1/2\kappa}}-\frac{\kappa\sin(Ea-\xi k(Ea)^{1/\kappa})}{\xi(Ea)^{1/\kappa-1}a^{1/2\kappa}}\right]^2\right\} \tag{48}$$

$$+\left(\frac{1-v}{1+v}\longrightarrow\frac{1+v}{1-v}\right),$$

where

$$\xi = \sqrt{\frac{1-v}{1+v}}\frac{t_i}{(E\tau_i)^{1/\kappa}}.$$

We see that, if the detector is in a switched-on state at $\tau \in (a, b)$, then it is excited by quanta with momenta k from the interval

$$\left(E\left(\frac{a}{t_i}\right)^{1/\kappa}\sqrt{\frac{1-v}{1+v}}, E\left(\frac{b}{t_i}\right)^{1/\kappa}\sqrt{\frac{1-v}{1+v}}\right).$$

So, for $t \longrightarrow -\infty$, the asymptotically flat spacetime (the background-related energy decreases as $1/t^2$) is filled with particles whose energy density cannot be ignored (it decreases as $1/t$). The cosmological solutions with zero occupation numbers have a zero measure in the function space of the problem and, in this sense, are atypical. Thus, there is an infinite arbitrariness in choosing the initial state for Veneziano's model, and the problem of initial data remains to be solved (see [8]).

4. THE SPACETIME STRUCTURE NEAR THE SINGULARITY

In the limit $t \longrightarrow -0$, the spacetime metric in the string frame is asymptotically equivalent to the Kasner metric with the indices

$$p_1 = \frac{[(\beta+\beta_1)^2+(\psi+\psi_1)^2-1]/2+\psi+\psi_1}{[(\beta+\beta_1)^2+(\psi+\psi_1)^2+3]/2+\psi+\psi_1}, \tag{49}$$

$$p_2 = \frac{1+\psi+\psi_1+\beta+\beta_1}{[(\beta+\beta_1)^2+(\psi+\psi_1)^2+3]/2+\psi+\psi_1}, \tag{50}$$

$$p_3 = \frac{1+\psi+\psi_1-\beta-\beta_1}{[(\beta+\beta_1)^2+(\psi+\psi_1)^2+3]/2+\psi+\psi_1}, \tag{51}$$

where

$$\psi_1 = \sum_k\frac{2}{\pi}N[e^{i\pi/4}(b_{1k}e^{ikx}+b_{2k}e^{-ikx})+\text{c.c.}], \tag{52}$$

$$\beta_1 = \sum_k\frac{2}{\pi}N[e^{i\pi/4}(b_{3k}e^{ikx}+b_{4k}e^{-ikx})+\text{c.c.}], \tag{53}$$

and $N = \sqrt{\pi\lambda_s^2t_i/2L^3}$ is the normalization constant for the modes calculated in the preceding section.

The contribution to the action of dilaton ϕ from the part of the classical trajectory (22) in the vicinity of the cosmological singularity ΔS_{sing} logarithmically diverges—it is proportional to $\ln(\Lambda/t_0)$, where Λ is the beginning of the Kasner epoch and $t_0 \longrightarrow 0$ is some small time scale. At the same time, the remaining contribution to the action is finite. This can be interpreted as the destruction of quantum coherence [9] between the modes c_{1k} and c_{2k} (c_{3k} and c_{4k}) for which the condition $|kt| \ll 1$ is satisfied. Because of this destruction, the modes with a wavelength longer than the cosmological horizon "freeze"—their amplitudes may be considered to be classical, randomly distributed variables rather than operators.[5] These modes contribute to the Kasner indices. The physical cause of this effect is that the radius of the cosmological horizon H^{-1} specifies the causal connectivity scale in the theory; accordingly, the quantum quantities can correlate only on scales smaller than H^{-1}. Amplitude freezing implies that an observer living at the Kasner epoch will always record the same Kasner indices, irrespective of the number of experiments that he or she carries out. Nevertheless, before the time the Kasner asymptotics becomes valid, we

---

[5] Strictly speaking, this can be done if the effective occupation numbers $\langle n_k\rangle \gg 1$. In this case, the noncommutativity of the coordinates and momenta may be disregarded. For a more detailed discussion, see [10].

cannot predict with confidence precisely what Kasner indices the observer will record—in this sense, the Kasner indices are random variables. Thus, the spacetime structure becomes stochastic as $t \longrightarrow -0$, and it makes sense to discuss the behavior of the various correlation functions of the Kasner indices. Below, we disregard the time-independent contributions to the metric, because they bear no relation to the behavior of the metric singularity for $t \longrightarrow -0$.

Technically, it is more convenient not to pass to world time but work with a metric of the form

$$ds^2 = t^{q_1}(dt^2 - dx^2) - t^{q_2}dy^2 - t^{q_3}dz^2.$$

The quantities q_1, q_2, and q_3 are related to the Kasner indices by

$$q_1 = \frac{2p_1}{1-p_1}, \quad q_2 = \frac{2p_2}{1-p_1}, \quad q_3 = \frac{2p_3}{1-p_1}. \tag{54}$$

Since the identity $p_1^2 + p_2^2 + p_3^2 = 1$ holds in the string frame, the following relation is also valid:

$$q_2^2 + q_3^2 = 4(q_1 + 1). \tag{55}$$

The stochastic spacetime structure can be completely determined by calculating the distribution function

$$\begin{aligned} F(\lambda, \mu) &= \langle \delta(q_2 - \lambda)\delta(q_3 - \mu) \rangle \\ &= \int dxdy e^{-i(\lambda x + \mu y)} \langle e^{i(q_2 x + q_3 y)} \rangle. \end{aligned}$$

Since q_1 can be unambiguously determined from the known quantities q_2 and q_3, this distribution function allows expressions for any correlators of the indices q_i to be derived. After simple calculations, we obtain

$$\begin{aligned} F(\lambda, \mu) &= \frac{\pi}{\sqrt{N_1 N_3}} \\ &\times \exp\left\{ -\frac{\left(1 + \psi - \dfrac{\lambda + \mu}{2}\right)^2}{2N_1} - \frac{\left(\beta - \dfrac{\lambda - \mu}{2}\right)^2}{2N_3} \right\}, \end{aligned} \tag{56}$$

where

$$N_1 = \sum_k \frac{n_{1k} N^2}{\pi^2}, \quad N_3 = \sum_k \frac{n_{3k} N^2}{\pi^2}.$$

This function characterizes the spacetime "in the infinitely small." Since the problem is translationally invariant, all points in space are equal in rights: the locally measured Kasner indices are the random variables described by the distribution function (56). However, the probability that the Kasner indices are small at a given point in space and large at a point infinitely close to it approaches zero. The nonlocal correlation properties of the Kasner indices are specified in part by the two-point correlation functions[6]

$$\begin{aligned} \langle q_2(x) q_2(x') \rangle &= (1 + \psi + \beta)^2 \\ &+ \sum_k \frac{16 N^2 (n_{1k} + n_{3k})}{\pi^2} \cos[k(x - x')], \end{aligned} \tag{57}$$

$$\begin{aligned} \langle q_3(x) q_3(x') \rangle &= (1 + \psi - \beta)^2 \\ &+ \sum_k \frac{16 N^2 (n_{1k} + n_{3k})}{\pi^2} \cos[k(x - x')], \end{aligned} \tag{58}$$

$$\begin{aligned} \langle q_2(x) q_3(x') \rangle &= (1 + \psi)^2 - \beta^2 \\ &+ \sum_k \frac{16 N^2 (n_{1k} - n_{3k})}{\pi^2} \cos[k(x - x')]. \end{aligned} \tag{59}$$

We can see that the correlation is oscillatory in pattern, which is an artifact of the method of specifying the initial conditions for $t \longrightarrow -\infty$.

In conclusion, let us consider the transition possibility from the superinflationary expansion of the universe to its contraction as the cosmological singularity is approached.

If the spacetime metric is isotropic and uniform, then we say that the universe undergoes superinflationary expansion when the following condition is satisfied:

$$\dot{H} = \frac{\ddot{a}}{a} - \left(\frac{\dot{a}}{a}\right)^2 > 0,$$

where a is the scale factor. If the latter changes with time as $a = (-t)^p$, then the inequality is satisfied for $p < 0$. This criterion can be generalized to the anisotropic case in two ways (as we will see below, the results depend only slightly on the choice of a criterion):

(i) by introducing not one but three scale factors and requiring the satisfaction of the inequalities $p_1 < 0$, $p_2 < 0$, and $p_3 < 0$;

(ii) by requiring the satisfaction of the inequality $p_1 + p_2 + p_3 < 0$, which characterizes the behavior of the comoving 4-volume element.

Since the spacetime acquires a stochastic structure that is locally characterized by distribution (56) as the singularity is approached, there is a nonzero probability that superinflation will stop and will give way to con-

---

[6] The remaining paired correlation functions can be easily calculated by using Eq. (55), which relates the Kasner indices. We do not give the corresponding expressions here because they are cumbersome.

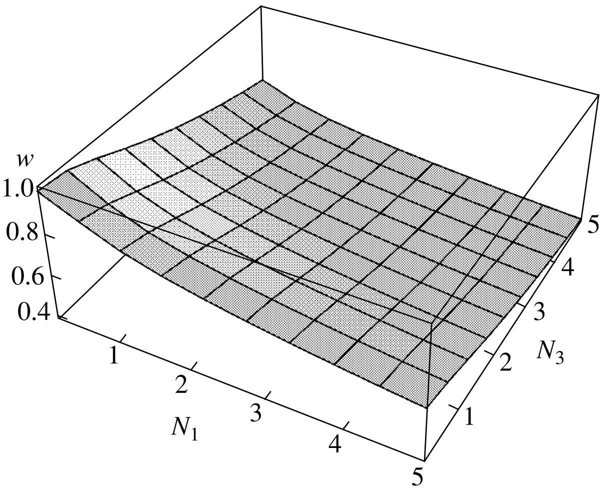

Fig. 1. Behavior of the probability of maximum superinflation duration, $w(\sum p_i < 0)$, as a function of N_1 and N_3 for $\psi = -\sqrt{3}$ and $\beta = 0$ at relatively low graviton and dilaton densities in the initial state.

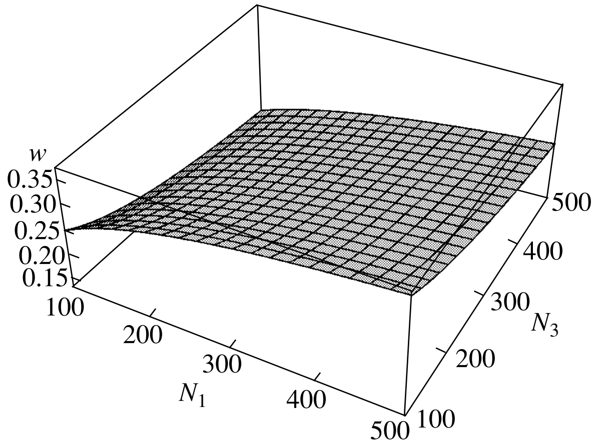

Fig. 2. Behavior of the transition probability from superinflationary expansion to contraction in all directions, $w(\forall p_i > 0)$, as a function of N_1 and N_3 for $\psi = -\sqrt{3}$ and $\beta = 0$ and high energy densities of gravitons N_1 and dilatons N_3 in the initial state.

traction. It is easy to see that, according to criterion (i), the inverse (in a sense) quantity—the probability that superinflation continues until the fall of the universe to a singularity of the second type—is

$$\begin{aligned} w(\forall p_i < 0) &= w(\forall q_i < 0) \\ &= \frac{1}{2\pi\sqrt{N_1 N_3}} \int_0^{\sqrt{2}} \rho d\rho \int_{3\pi/4}^{5\pi/4} d\phi \\ &\times \exp\left\{-\frac{(1+\psi-\rho\cos\phi)^2}{2N_1} - \frac{(\beta-\rho\cos\phi)^2}{2N_3}\right\}. \end{aligned} \tag{60}$$

In the physically interesting case with $N_1, N_3 \longrightarrow \infty$ and $\psi, \beta \sim 1$, we have

$$w(\forall p_i < 0) \approx \frac{1}{8\sqrt{N_1 N_3}}. \tag{61}$$

The probability that superinflationary expansion will give way to contraction in all directions in the same limit behaves as

$$w(\forall p_i > 0) \approx \frac{1}{\pi}\arctan\sqrt{\frac{N_1}{N_3}} - \frac{1}{8\sqrt{N_1 N_3}}. \tag{62}$$

If, however, we proceed from criterion (ii), then the probability of maximum superinflation duration is

$$\begin{aligned} & w(\sum p_i < 0) \\ &= \frac{1}{2\pi\sqrt{N_1 N_3}} \int_0^{\sqrt{6}} \rho d\rho \int_0^{2\pi} d\phi \exp\left\{-\frac{(3+\psi-\rho\cos\phi)^2}{2N_1}\right. \\ &\left. - \frac{(\beta-\rho\sin\phi)^2}{2N_3}\right\} \approx \frac{3}{\sqrt{N_1 N_3}}, \\ & N_1, N_3 \longrightarrow \infty, \quad \psi, \beta \sim 1. \end{aligned} \tag{63}$$

Since asymptotics (60) and (63) are virtually independent of the choice of a superinflation duration criterion, we can say that superinflationary expansion necessarily gives way to contraction as the cosmological singularity is approached.

Figure 1 shows the behavior of the probability of maximum superinflation duration $w(\sum p_i < 0)$ as a function of N_1 and N_3 for $\psi = -\sqrt{3}$ and $\beta = 0$, which corresponds to the absence of a seed anisotropy, i.e., Veneziano's universe. Interestingly, at moderately large N_1 and N_3, an increase in the graviton energy density with respect to the dilaton energy density generally causes this probability to decrease. Nevertheless, at large N_1 and N_3, the graviton and dilaton energy densities have the same weight from the viewpoint of their influence on the probability of maximum superinflation duration—an increase in the number of gravitons and dilatons in the initial state always causes this probability to decrease.

When considering the asymptotics of the probability of transition from superinflation to contraction in all directions, $w(\forall p_i > 0)$ at large N_1 and N_3, we arrive at a similar picture (see Fig. 2): an increase in the number of gravitons in the initial state generally causes this probability to decrease, while an increase in the number of dilatons causes it to increase. This implies that the statistical weight of the states describing anisotropic expansion (when there is expansion in two of the three directions in space and contraction in the third direction, etc.) becomes large; the higher the graviton energy density in the initial state, the larger this weight.

5. CONCLUSIONS

We have considered the structure of the cosmological singularity in Veneziano's model. As we showed, the problem of uniqueness in choosing the initial conditions in Veneziano's scenario has yet to be solved—the asymptotically flat world that corresponds to the initial state in the scenario can be filled with gravitational and scalar field quanta. In order to understand what influence any variations in initial data has on the spacetime structure near the singularity, we constructed an exact solution that described the gravitational and scalar field quanta propagating against the background of asymptotically flat spacetime in the limit $t \longrightarrow -\infty$. The spacetime near the singularity was shown to be described by the Kasner metric with indices that are classical random variables. Our calculations of the distribution function for the Kasner indices and the probabilities of maximum superinflation duration and a transition from superinflation to contraction in all directions (which corresponds to the transition from a singularity of the second type to an ordinary singularity corresponding to the invariants of the curvature tensor becoming infinite) indicate that an increase in the number of gravitons in the initial state causes the statistical weight of the evolutions corresponding to anisotropic expansion to increase.

Physically, this implies that the spacetime near the singularity acquires a domain structure, with an anisotropic expansion of its own being realized in each domain.

Thus, we can say that the quantization of Veneziano's scenario in the low-energy approximation of the string theory does not lead to a general softening of the singularity and ideas of the nonperturbative string theory must be invoked to solve the problem of the transition from inflation to decelerating Friedmann expansion.

ACKNOWLEDGMENTS

I am grateful to A.A. Starobinsky for helpful discussions. This study was supported in part by the Russian Foundation for Basic Research (project nos. 02-02-16817 and MAC 02-02-06914) and the Basic Research Program "Nonstationary Phenomena in Astronomy" (Russian Academy of Sciences).

REFERENCES

1. B. Campbell, A. Linde, and K. Olive, Nucl. Phys. B **355**, 146 (1991).
2. G. Veneziano, Phys. Lett. B **265**, 287 (1991).
3. M. B. Green, J. H. Schwarz, and E. Witten, *Superstring Theory* (Cambridge Univ. Press, Cambridge, 1987; Mir, Moscow, 1990).
4. M. Gasperini and G. Veneziano, Phys. Rep. **373**, 1 (2003).
5. M. Gasperini, J. Maharana, and G. Veneziano, Nucl. Phys. B **472**, 349 (1996).
6. J. Barrow and K. Kunze, Phys. Rev. D **56**, 741 (1997).
7. N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge Univ. Press, Cambridge, 1982; Mir, Moscow, 1984).
8. N. Kaloper, A. Linde, and R. Bousso, Phys. Rev. D **59**, 043508 (1999).
9. A. Starobinsky, in *Field Theory, Quantum Gravity, and Strings*, Ed. by H. J. de Vega and N. Sanchez (Springer, New York, 1986).
10. D. Polarski and A. Starobinsky, Class. Quantum Grav. **13**, 377 (1996).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 658–667.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 733–743.
Original Russian Text Copyright © 2003 by Elizarov, Tupitsyn.

ATOMS, SPECTRA, RADIATION

Analysis of the Angular Distribution of Auger Electrons in a Xe Atom

A. Yu. Elizarov[a],* and I. I. Tupitsyn[b],**
[a]*Ioffe Physicotechnical Institute, ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia*
**e-mail: a.elizarov@mail.ioffe.ru*
[b]*St. Petersburg State University, Universitetskaya nab. 7/9, St. Petersburg, 198904 Russia*
***e-mail: tup@tup.usr.pu.ru*

Received November 22, 2002

Abstract—The angular distribution of Auger electrons, and the numerically calculated anisotropy parameters of the angular distribution α_2 for ($M_3 \longrightarrow N_2N_3$), ($M_3 \longrightarrow N_3N_3$), ($M_4 \longrightarrow N_1N_3$), ($M_4 \longrightarrow N_4N_5$), ($M_4 \longrightarrow N_5N_5$), and ($M_{4,\,5} \longrightarrow O_{2,\,3}O_{2,\,3}$) transitions in a Xe atom are given. The matrix elements are calculated by using the nonrelativistic Hartree–Fock method in *LS* coupling and the relativistic Hartree–Fock–Dirac method in both *jj* coupling (the single-configuration approximation) and intermediate coupling (the multiconfiguration approximation). *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The existence of anisotropy in the angular distribution of Auger electrons in atoms was first predicted by Mehlhorn [1]. This effect has since been extensively studied experimentally and theoretically. Of particular interest in these studies are the Auger electrons produced when the deep vacancies formed by the interaction of atoms with synchrotron radiation are filled [2]. A general theory for the anisotropy in the angular distribution of Auger electrons was developed in [3–6], where the density-matrix formalism was used. Subsequently, the angular distribution coefficients were calculated by several theoretical groups [7–9]. However, as was noted in [6, 9, 10], there are often discrepancies between experimental data and calculations, indicating that the studies of the angular distribution of Auger electrons should be continued. The causes of these discrepancies have not yet been completely elucidated. They can result from imperfection of the theory for the angular distribution of Auger electrons itself (in particular, the necessity of allowing for the interference of the final states of an ion + Auger electron system [11]) and from the use of different approximations in calculating the wave functions of the initial and final ion states and the continuum wave functions.

In this paper, we use the theory for the angular distribution of Auger electrons developed in [6–8]. This theory is generalized to atoms with open valence shells for both *LS* and *jj* couplings. In addition, we derive expressions for the anisotropy coefficients of the angular distribution α_2 for intermediate coupling, i.e., for the multiconfiguration relativistic wave functions of the final and initial ion states. The wave functions of ions were calculated by the Hartree–Fock method and the multiconfiguration relativistic Hartree–Fock–Dirac method.

The parameters of the angular Auger electron distribution depend on the quality of the continuum wave function, which was calculated in many previous studies by using rough approximations. In particular, the exchange interaction between the continuum electron and the core electrons, as well as the nondiagonal Lagrange factors that ensure the orthogonality of the continuum wave function to single-electron occupied ion states, was disregarded. In this paper, we consider the influence of these approximations on the angular distribution parameter α_2. We determine the continuum wave function in our relativistic calculations by solving the relativistic Hartree–Fock–Dirac equations. In calculating the continuum wave functions, the influence of relativistic effects on α_2 may prove to be significant, because the behavior of the continuum wave function in the core region mainly contributes to α_2.

In the next section, we describe the basic theoretical principles used to calculate the anisotropy parameter for the angular distribution of Auger electrons. In Section 3, we give the angular anisotropy parameters α_2 for ($M_3 \longrightarrow N_2N_2$), ($M_3 \longrightarrow N_3N_3$), ($M_4 \longrightarrow N_1N_3$), ($M_4 \longrightarrow N_4N_4$), ($M_4 \longrightarrow N_5N_5$), and ($M_{4,\,5} \longrightarrow O_{2,\,3}O_{2,\,3}$) transitions in a Xe atom calculated by using various approximations.

2. THE GENERAL THEORY

The scattering of electrons or photons by an atom A(J_0) can give rise to a vacancy in the inner shell of $A^+(J_1)$. The formed vacancy can be filled with the emis-

1063-7761/03/9704-0658$24.00 © 2003 MAIK "Nauka/Interperiodica"

sion of a photon or an electron (the Auger process). In this paper, we consider the Auger process

$$\begin{aligned} &\mathrm{A}(J_0) + e \longrightarrow e_s + e_d + \mathrm{A}^{+}(J_1) \\ &\qquad\qquad\qquad\qquad \searrow \mathrm{A}^{+}(J_1) \longrightarrow \mathrm{A}^{2+}(J) + e_A. \end{aligned} \tag{1}$$

Here, e_s and e_d are the scattered and dislodged electrons, respectively. The Auger process results in the formation of a doubly charged ion $\mathrm{A}^{2+}(J)$ and a continuum Auger electron e_A in the final state. Since the lifetime of the excited state is much longer than the collision time, we can imagine that the scattering takes place in two independent stages: first, the formation of a vacancy and, second, Auger decay [1]. To eliminate the interference between the states of the two electrons, e_s and e_A, let us consider a process in which the energies of the scattered and Auger electrons are unequal. The quantum states of the electrons e_s and e_d and the doubly charged ion are not fixed in the experiment. In addition, we assume that neither the electron (e) nor the atom $\mathrm{A}(J_0)$ are polarized in the initial state ($A(\gamma_0 J_0)$).

A preferential direction is formed when electrons are scattered by an atom in a system of colliding particles, which, as was shown in [12], leads to anisotropy in the angular intensity distribution $I(\Theta)$ of Auger electrons. An expression for the angular distribution $I(\Theta)$ of Auger electrons was derived by several authors [1, 4–7]. In [6], the parameters $A(KkQ)$, which contain information about the dynamics and geometry of the Auger process, were introduced to describe the angular distribution $I(\Theta)$. General expressions for $A(KkQ)$ are given in the same paper:

$$\begin{aligned} A(KkQ) &= \sqrt{(2K+1)(2k+1)} \\ &\times \sum_{M, M_1, M_1'} \sum_{m_s, m_s'} (-1)^{J_1 - M_1 + 1/2 - m_s} \begin{pmatrix} J_1 & J_1 & K \\ M_1 & -M_1' & -Q \end{pmatrix} \\ &\times \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & k \\ m_1 & -m_s' & -Q \end{pmatrix} \langle JM, p^{(-)} m_s | V | J_1 M_1 \rangle \\ &\times \langle JM, p^{(-)} m_s' | V | J_1 M_1' \rangle^*, \end{aligned} \tag{2}$$

where V is the electron–electron Coulomb interaction operator:

$$V = \frac{1}{2} \sum_{i \neq j} v(\mathbf{r}_i, \mathbf{r}_j), \quad v(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|}. \tag{3}$$

2.1. *Intermediate and jj Couplings*

By expanding the wave function of the Auger electron continuum $\langle p^{(-)} m_s|$ in terms of spherical waves and taking into account the spin–orbit interaction [7], we can derive the following expression for $A(KkQ)$:

$$\begin{aligned} A(KkQ) &= \frac{1}{4\pi p} \sqrt{(2K+1)(2k+1)} \\ &\times \sum_{l, l'} i^{l'-l} \exp[i(\sigma_l^j - \sigma_{l'}^{j'})] \sum_{j, j'} (-1)^{J + J_1 + j + Q + l'} \\ &\times \sqrt{(2l+1)(2l'+1)(2j+1)(2j'+1)} \\ &\times \begin{Bmatrix} J & J_1 & j \\ K & j' & J_1 \end{Bmatrix} \sum_X (2X+1) \begin{pmatrix} X & l' & l \\ 0 & 0 & 0 \end{pmatrix} \\ &\times \begin{pmatrix} K & X & k \\ -Q & 0 & Q \end{pmatrix} \begin{Bmatrix} \frac{1}{2} & \frac{1}{2} & k \\ j' & j & K \\ l' & l & X \end{Bmatrix} \\ &\times \langle (J, \varepsilon j) J_1 \| V \| J_1 \rangle \langle (J, \varepsilon j') J_1 \| V \| J_1 \rangle, \end{aligned} \tag{4}$$

where $l + l'$ is even. Here, we use the notation of the Clebsch–Gordan coefficients, the $6j$ and $9j$ symbols from [13], and σ_l^j is the phase shift for the continuum electron in a state $\langle lj|$.

The angular distribution coefficients α_K are related to the parameters $A(KkQ)$ as follows [6]:

$$\alpha_K(J) = \frac{A(K00)}{A(000)}. \tag{5}$$

For α_K, we can easily obtain

$$\begin{aligned} \alpha_K &= \left(\sum_{lj} \langle J_1 \| V \| (Jj) J_1 \rangle^2 \right)^{-1} \sqrt{(2K+1)(2J_1+1)} \\ &\times \sum_{ll'} \sum_{jj'} (-1)^{J + J_1 - 1/2 + l'} i^{l+l'} C(K)_{jj'} \\ &\times \cos(\sigma_{l'} - \sigma_l) \langle J_1 \| V \| (Jj) J_1 \rangle \langle J_1 \| V \| (Jj') J_1 \rangle, \end{aligned} \tag{6}$$

where

$$\begin{aligned} C(K)_{jj'} &= -(-1)^{K+2j'} \sqrt{\frac{(2j+1)(2j'+1)}{2K+1}} \\ &\times C^{K0}_{j'-1/2,\, j1/2} \begin{Bmatrix} J & J_1 & j \\ K & j' & J_1 \end{Bmatrix}. \end{aligned} \tag{7}$$

The above matrix elements $\langle J_1 \| V \| (Jj)J_1 \rangle$ were determined for the initial and final multielectron states of an arbitrary atom. In general, these matrix elements can be calculated by using the Wigner–Eckart theorem [13, 14] if the multielectron wave function Ψ_{J_1, M_1} of the initial state of $A^+(J_1)$, the wave function $\Psi_{J, M}$ of the final state of $A^{2+}(J)$, and the single-electron wave function ψ_{jm} of the Auger electron are known:

$$\langle J_1 \| V \| (J, j)J_1 \rangle = \frac{\sqrt{2J_1 + 1}}{C^{J_1 M_1}_{JM, jm}} \langle J_1 M_1 | V | JM, jm \rangle. \tag{8}$$

The wave functions Ψ_{J_1, M_1} and $\Psi_{J, M}$ can be calculated by the relativistic single-configuration Hartree–Fock–Dirac method. In this case, the coefficients α_K are determined in *jj* coupling. For atoms with closed shells, this approach is equivalent to the approach developed in [6–8] and is justified only for heavy atoms with closed shells, where the final state of $A^{2+}(J)$ has vacancies at deep core levels. More accurate results can be obtained by the multiconfiguration Hartree–Fock method. In particular, we can take into account the superposition of all the relativistic configurations that correspond to one nonrelativistic configuration of the ion, i.e., all the configurations that have the same populations of the nonrelativistic shells (*nl*) and different populations of the relativistic shells (*nlj*). We will call this approximation intermediate coupling. Obviously, the intermediate coupling transforms into *LS* coupling in the nonrelativistic limit. This cannot be asserted for pure *jj* coupling for systems with open shells and, in particular, for ions with two vacancies at inner levels.

For atoms with closed valence shells, we can derive an expression for the above matrix elements $\langle J_1 \| V \| (Jj)J_1 \rangle$ in hole representation in a pure *jj* coupling:

$$\begin{aligned} \langle (J, \varepsilon j)J_1 \| V \| J_1 \rangle &= (-1)^{\Theta + j} \\ &\times \langle (l_f j_f, l'_f j'_f)J \| v \| (l_i j_i, \varepsilon l j)J \rangle, \end{aligned} \tag{9}$$

where $l_f j_f$ and $l'_f j'_f$ are the quantum numbers of two vacancies of $A^{2+}(J)$, and $l_i j_i$ are the quantum numbers of one vacancy of the initial state of $A^+(J_1)$. Θ is a half-integer and depends on the choice of phase factors for the wave functions of the initial and final ion states. When (9) is substituted into (4) and (6), the dependence on Θ vanishes and the additional phase factor is $(-1)^{j + j' + 1}$. The derived expressions for $A(KkQ)$ and α_K in hole representation are similar to (9) and (25) from [7].

The above matrix element in hole representation for *jj* coupling was obtained in [15] and is given in final form in [7]. In our notation, it is

$$\begin{aligned} &\langle (l_f j_f, l'_f j'_f)J \| v \| (l_i j_i, \varepsilon l j)J \rangle \\ &= \tau(-1)^{j'_f + j_i} \sqrt{(2j_f + 1)(2j'_f + 1)(2J + 1)} \\ &\times \Bigg[(-1)^J \sum_k C^{j_i 1/2}_{j_f 1/2, k0} C^{j 1/2}_{j'_f 1/2, k0} \begin{Bmatrix} j'_f & j & k \\ j_i & j_f & J \end{Bmatrix} \\ &\times R^k(n_f l_f j_f, n'_f l'_f j'_f, n_i l_i j_i, \varepsilon l j) \\ &+ \sum_k C^{j_i 1/2}_{j'_f 1/2, k0} C^{j 1/2}_{j_f 1/2, k0} \begin{Bmatrix} j_f & j & k \\ j_i & j'_f & J \end{Bmatrix} \\ &\times R^k(n_f l_f j_f, n'_f l'_f j'_f, \varepsilon l j, n_i l_i j_i) \Bigg], \end{aligned} \tag{10}$$

where the coefficient τ depends on whether the two vacancies of $A^{2+}(J)$ are equivalent or not:

$$\tau = \begin{cases} 1, & n_f l_f j_f \neq n'_f l'_f j'_f \\ \dfrac{1}{\sqrt{2}}, & n_f l_f j_f = n'_f l'_f j'_f. \end{cases} \tag{11}$$

The radial integrals R^k in (10) are identical to the standard radial integrals used in the Hartree–Fock–Dirac method [16]:

$$\begin{aligned} R^k(A, B, C, D) &= \int_0^\infty dr_1 \int_0^\infty dr_2, \\ &[P_A(r_1)P_C(r_1) + Q_A(r_1)Q_C(r_1)] \\ &\times \gamma_k(r_1, r_2)[P_B(r_1)P_D(r_1) + Q_B(r_1)Q_D(r_1)], \end{aligned} \tag{12}$$

where *A*, *B*, *C*, and *D* number the relativistic shells; *P* and *Q* are the large and small components of the radial wave function, respectively; and

$$\gamma_k(r_1, r_2) = \frac{r_<^k}{r_>^{k+1}} \tag{13}$$

2.2. LS Coupling

If the relativistic effects are negligible, then the approximation of *LS* coupling can be used to describe the ion states. The standard transformation from *jj* coupling to *LS* coupling for the final state should be used to derive an expression for the angular distribution parameters $A(K00)$ in LS coupling. For the above matrix element in LS coupling, we then obtain (see, e.g., [17])

$$\langle (J, \varepsilon l j) J_1 \| V \| J_1 \rangle = \sqrt{(2L_1+1)(2S_1+1)(2J+1)(2j+1)} \times \begin{Bmatrix} L & S & J \\ l & \frac{1}{2} & J \\ L_1 & S_1 & J_1 \end{Bmatrix} \times \left\langle \left(\left(LS, \varepsilon l \tfrac{1}{2} \right) L_1 S_1 \right) J_1 \| V \| (L_1 S_1) J_1 \right\rangle. \tag{14}$$

Using the Wigner–Eckart theorem, we can easily obtain

$$\left\langle \left(\left(LS, \varepsilon l \tfrac{1}{2} \right) L_1 S_1 \right) J_1 \| V \| (L_1 S_1) J_1 \right\rangle = \sqrt{\frac{(2J_1+1)}{(2L_1+1)(2S_1+1)}} \times \left\langle \left(LS, \varepsilon l \tfrac{1}{2} \right) L_1 S_1 \| V \| L_1 S_1 \right\rangle. \tag{15}$$

In this case, instead of (14), we have

$$\langle (J, \varepsilon l j) J_1 \| V \| J_1 \rangle = \sqrt{(2J_1+1)(2J+1)(2j+1)} \times \begin{Bmatrix} L & S & J \\ l & \frac{1}{2} & j \\ L_1 & S_1 & J_1 \end{Bmatrix} \left\langle \left(LS, \varepsilon l \tfrac{1}{2} \right) L_1 S_1 \| V \| L_1 S_1 \right\rangle. \tag{16}$$

For the parameters $A(K00)$ in LS coupling, we can then derive the expression

$$\begin{aligned} A(K00) = {} & \frac{1}{4\sqrt{2}\pi p} \sum_{l, l'} i^{l'-l} \exp[i(\sigma_l - \sigma_{l'})] \\ & \times \sqrt{(2l+1)(2l'+1)} C^{K0}_{l0, l'0} \sum_{j, j'} (-1)^{1/2 - J - J_1 + l + l'} \\ & \times (2J_1+1)(2J+1)(2j+1)(2j'+1) \begin{Bmatrix} \frac{1}{2} & j' & l' \\ K & l & j \end{Bmatrix} \\ & \times \begin{Bmatrix} J & J_1 & j \\ K & j' & J_1 \end{Bmatrix} \begin{Bmatrix} J_1 & L_1 & S_1 \\ J & L & S \\ j & l & \frac{1}{2} \end{Bmatrix} \begin{Bmatrix} J_1 & L_1 & S_1 \\ J & L & S \\ j' & l' & \frac{1}{2} \end{Bmatrix} \\ & \times \left\langle \left(LS, \varepsilon l \tfrac{1}{2} \right) L_1 S_1 \| V \| L_1 S_1 \right\rangle \left\langle \left(LS, \varepsilon l' \tfrac{1}{2} \right) L_1 S_1 \| V \| L_1 S_1 \right\rangle. \end{aligned} \tag{17}$$

Here, L_1, S_1, and J_1 are the quantum numbers that describe the state of the singly charged ion A^+; L, S, and J are the quantum numbers that describe the state of the doubly charged ion A^{++}; l and l' are the orbital quantum numbers of the continuum (Auger) electron. Since the nonrelativistic wave function of the Auger electron does not depend on the quantum number j, the above matrix elements in (17) do not depend on j and j'. Performing the summation over j and j', we can then transform (17) to

$$\begin{aligned} A(K00) = {} & \frac{1}{\sqrt{2}} \frac{1}{4\pi p} (2J+1)(2J_1+1) \\ & \times \sum_{l, l'} i^{l'-1} \exp[i(\sigma_l - \sigma_{l'})] \sqrt{(2l+1)(2l'+1)} C^{K0}_{l0, l'0} \\ & \times \sum_x (-1)^{x+J_1} (2x+1) \begin{Bmatrix} x & J_1 & l \\ K & l' & J_1 \end{Bmatrix} \\ & \times \begin{Bmatrix} L & S & J \\ \frac{1}{2} & x & S_1 \end{Bmatrix}^2 \begin{Bmatrix} L & L_1 & l \\ J_1 & x & S_1 \end{Bmatrix} \begin{Bmatrix} L & L_1 & l' \\ J_1 & x & S_1 \end{Bmatrix} \\ & \times \left\langle \left(LS, \varepsilon l \tfrac{1}{2} \right) L_1 S_1 \| V \| L_1 S_1 \right\rangle \left\langle \left(LS, \varepsilon l' \tfrac{1}{2} \right) L_1 S_1 \| V \| L_1 S_1 \right\rangle. \end{aligned} \tag{18}$$

The above matrix elements can be calculated by using the Wigner–Eckart theorem [13, 14] if the multielectron wave function $\Psi_{L_1, M_{L_1} S_1, M_{S_1}}$ of the initial state of A^+, the wave function $\Psi_{L, M_L S, M_S}$ of the final state of A^{2+}, and the single-electron wave function of the Auger electron $\psi_{lm, 1/2 m_s}$ are known:

$$\left\langle \left(LS, \varepsilon l \tfrac{1}{2} \right) L_1 S_1 \| V \| L_1 S_1 \right\rangle = \frac{\sqrt{(2L_1+1)(2S_1+1)}}{C^{L_1 M_{L_1}}_{L M_L, lm} C^{S_1 M_{S_1}}_{S M_S, lm}} \times \langle L_1 M_{L_1} S_1 M_{S_1} | V | L M_L, lm \rangle. \tag{19}$$

For atoms with closed valence shells, the matrix elements in (17) and (18) can be calculated in hole representation [15] by analogy with the case of *jj* coupling (9):

$$\left\langle\left(LS, \varepsilon l\frac{1}{2}\right)L_1S_1\|V\|L_1S_1\right\rangle$$

$$= \left\langle\left(l_f\frac{1}{2}, l'_f\frac{1}{2}\right)LS\|v\|\left(\varepsilon l\frac{1}{2}, l_i\frac{1}{2}\right)LS\right\rangle$$

$$= \tau(-1)^{l_f+l_i}\sqrt{(2l_f+1)(2l'_f+1)}$$

$$\times\left[(-1)^L\sum_k R^k(n_fl_f, n'_fl'_f, n_il_i, \varepsilon l)\right. \tag{20}$$

$$\times C^{l_i0}_{l_f0,\,k0}C^{l0}_{l'_f0,\,k0}\begin{Bmatrix} l'_f & l_f & L \\ l_i & l & k \end{Bmatrix}$$

$$+ (-1)^S\sum_k R^k(n_fl_f, n'_fl'_f, \varepsilon l, n_il_i)$$

$$\left.\times C^{l0}_{l_f0,\,k0}C^{l_i0}_{l'_f0,\,k0}\begin{Bmatrix} l'_f & l_f & L \\ l_i & l & k \end{Bmatrix}\right],$$

where $R^k(nl, n'l', n_1l_1, n'_1l'_1)$ is the radial integral:

$$R^k(A, B, C, D) = \int_0^\infty dr_1\int_0^\infty dr_2P_A(r_1)$$
$$\times P_B(r_2)\gamma_k(r_1, r_2)P_C(r_1)P_D(r_2). \tag{21}$$

2.3. Calculating the Continuum Wave Functions

We determined the nonrelativistic continuum wave functions in the Hartree–Fock approximation by solving the equation

$$-\frac{1}{2}\frac{d^2}{dr^2}P_{\varepsilon l}(r) + \left[\frac{l(l+1)}{2r^2} + V_c(r)\right]P_{\varepsilon l}(r) + W_{\text{ex}}(r)$$
$$= \varepsilon P_{\varepsilon l}(r) + \sum_{nl}\lambda_{\varepsilon l,\,nl}P_{nl}(r), \tag{22}$$

where n and l are the quantum numbers of the occupied atomic shells of A^{2+}, $\lambda_{\varepsilon l,\,nl}$ are the nondiagonal Lagrange factors that ensure the orthogonality of the continuum wave function $P_{\varepsilon l}$ to the atomic radial wave functions P_{nl} of the same symmetry, $V_c(r)$ is the Coulomb potential, and W_{ex} is the result of the action of the nonlocal exchange operator on the radial continuum wave function.

The continuum wave function $P_{\varepsilon l}$ used above was normalized to the δ function in energy, $\langle P_{\varepsilon l}|P_{\varepsilon' l}\rangle = \delta(\varepsilon - \varepsilon')$. In this case, the asymptotics of $P_{\varepsilon l}$ is

$$P_\varepsilon(r) \approx \sqrt{\frac{2}{\pi p}}\sin(\tau + \sigma_l),$$
$$\tau = pr + \frac{Z}{p}\ln(2pr) - l\frac{\pi}{2}, \tag{23}$$

where $p = \sqrt{2\varepsilon}$, Z is the ion charge, and σ_l is the scattering phase. By joining an arbitrary unnormalized regular (at zero) solution $\tilde{P}_\varepsilon(r) = NP_\varepsilon(r)$ of Eq. (22) and its derivative with asymptotics (23), we can determine the normalization N and the phase of the solution $\tilde{P}_\varepsilon(r)$. However, to achieve a high accuracy, the joining should be performed at distances on the order of 10^4–10^5 atomic units; i.e., we should find the regular (at zero) solution $\tilde{P}_\varepsilon(r)$ of Eq. (22) in a very large interval of the radial variable τ. The procedure of joining the solution and its derivative with a linear combination of the regular, F, and irregular, G, (at zero) Coulomb functions [18] is much more efficient:

$$\tilde{P}_\varepsilon(R_a) = AF(R_a) + BG(R_a). \tag{24}$$

Here, R_a is the joining point. The Coulomb functions can be calculated, for example, by using an efficient procedure described in [19], where the technique of continued fractions was used. In this case, the joining point R_a can be chosen in the region where all of the atomic radial wave functions are negligible and the atomic potential is a Coulomb one with a high accuracy, i.e., at $R_a \sim 20$–50 atomic units. The normalization N and the phase σ_l can be determined from the coefficients A and B by using the relations

$$\cos\sigma_l = \frac{1}{N}(A\cos\sigma_l^0 - B\sin\sigma_l^0),$$
$$\sin\sigma_l = \frac{1}{N}(A\sin\sigma_l^0 + B\cos\sigma_l^0), \tag{25}$$
$$N = \sqrt{A^2 + B^2},$$

where σ_l^0 is the phase of the Coulomb functions [18],

$$\sigma_l^0 = \arg\Gamma(l + 1 + i\eta), \quad \eta = -\frac{Z}{\sqrt{2\varepsilon}}. \tag{26}$$

The relativistic continuum wave functions were determined in the Hartree–Fock–Dirac approximation by solving the equations [16]

$$\begin{gathered}
c\left(-\frac{d}{dr}+\frac{k}{r}\right)Q_{\varepsilon k}+V_cP_{\varepsilon k}+W^P_{\mathrm{ex}} = \varepsilon P_{\varepsilon k}+\sum_{nl}\lambda_{\varepsilon l,\,nl}P_{nk},\\
c\left(\frac{d}{dr}+\frac{k}{r}\right)P_{\varepsilon k}+[-2c^2+V_c]Q_{\varepsilon k}+W^Q_{\mathrm{ex}}\\
= \varepsilon Q_{\varepsilon k}+\sum_{nl}\lambda_{\varepsilon l,\,nl}Q_{nk}.
\end{gathered}\tag{27}$$

Here, $P_{\varepsilon k}$ and $Q_{\varepsilon k}$ are the large and small components of the radial continuum wave function, P_{nk} and Q_{nk} are the large and small components of the radial wave functions for the occupied shells of A^{2+}, W^P_{ex} and W^Q_{ex} are the large and small components resulting from the action of the nonlocal exchange operator on the two-component radial continuum wave function, and k is the relativistic quantum number.

The normalization to energy for the relativistic wave functions is

$$\int_0^\infty dr[P_\varepsilon(r)P_{\varepsilon'}(r)+Q_\varepsilon(r)Q_{\varepsilon'}(r)] = \delta(\varepsilon-\varepsilon'). \tag{28}$$

The radial continuum wave function normalized to the δ function in energy at large r has the asymptotics [20]

$$\begin{gathered}
P(r)\approx\frac{1}{c}\left(\frac{\varepsilon+2c^2}{\pi p}\right)^{1/2}\sin(\tau+\sigma_k),\\
Q(r)\approx\frac{1}{c}\left(\frac{\varepsilon}{\pi p}\right)^{1/2}\cos(\tau+\sigma_k),\\
\tau = pr-\eta\ln(2pr)-\frac{\pi l^*}{2},
\end{gathered}\tag{29}$$

where

$$l^* = \begin{cases}\gamma, & k>0\\ \gamma-1, & k<0\end{cases}\tag{30}$$

and

$$\begin{gathered}
\gamma = \sqrt{k^2-\frac{Z^2}{c^2}},\quad p = \frac{1}{c}\sqrt{(\varepsilon+c^2)^2-c^4},\\
\eta = -\frac{Z(\varepsilon+c^2)}{c^2p}.
\end{gathered}\tag{31}$$

Just as in the nonrelativistic case, σ_k is the phase of the continuum wave function.

Just as in the nonrelativistic case, to determine the phase and the normalization of an arbitrary unnormalized regular (at zero) solution of the Dirac equation $\tilde{P}_\varepsilon$, $\tilde{Q}_\varepsilon$, we joined the solution found with a linear combination of the regular, P_r and Q_r, and irregular, P_i and Q_i, (at zero) relativistic Coulomb functions:

$$\begin{aligned}
\tilde{P}_\varepsilon(R_a) &= AP_r(R_a)+BP_i(R_a),\\
\tilde{Q}_\varepsilon(R_a) &= AQ_r(R_a)+BQ_i(R_a).
\end{aligned}\tag{32}$$

The asymptotics of the relativistic Coulomb functions can be chosen in the form

$$\begin{aligned}
P_r(r) &\approx \frac{1}{c}\left(\frac{\varepsilon+2c^2}{\pi p}\right)^{1/2}\sin(\tau+\sigma_k^0),\\
Q_r(r) &\approx \frac{1}{c}\left(\frac{\varepsilon}{\pi p}\right)^{1/2}\cos(\tau+\sigma_k^0),\\
P_i(r) &\approx \frac{1}{c}\left(\frac{\varepsilon+2c^2}{\pi p}\right)^{1/2}\cos(\tau+\sigma_k^0),\\
Q_i(r) &\approx -\frac{1}{c}\left(\frac{\varepsilon}{\pi p}\right)^{1/2}\sin(\tau+\sigma_k^0).
\end{aligned}\tag{33}$$

Here, σ_k^0 is the phase of the relativistic Coulomb functions. Using the standard expression [20], we can easily obtain the following expression for this phase:

$$\begin{aligned}
\sigma_k^0 &= \arg\Gamma(l^*+1+i\eta)\\
&+\frac{1}{2}\arg\left(\frac{k+i\eta c^2/(\varepsilon+c^2)}{\gamma k/|k|+i\eta}\right).
\end{aligned}\tag{34}$$

For the relativistic wave functions with asymptotics (33), the Wronskian is

$$W = P_iQ_r-P_rQ_i = \frac{1}{c\pi}. \tag{35}$$

To calculate the relativistic Coulomb functions, we used the transformation suggested in [21, 22]. This transformation reduces the Dirac radial Coulomb equation to two differential equations that are formally identical to the Schrödinger nonrelativistic equations. It can be written as

$$\begin{gathered}
\begin{pmatrix}P'\\ Q'\end{pmatrix} = U\begin{pmatrix}P\\ Q\end{pmatrix},\\
U = \begin{pmatrix}1 & X\\ X & 1\end{pmatrix},\quad X = -\frac{Z}{c}\frac{k}{|k|}\frac{1}{|k|+\gamma}.
\end{gathered}\tag{36}$$

Using the transformation U, we can derive the second-order differential equations

$$\left[-\frac{d^2}{dr^2}+\frac{l_1^*(l_1^*+1)}{r^2}-\frac{2Z^*}{r}\right]P' = 2\varepsilon^* P',$$
$$\left[-\frac{d^2}{dr^2}+\frac{l_2^*(l_2^*+1)}{r^2}-\frac{2Z^*}{r}\right]Q' = 2\varepsilon^* Q', \tag{37}$$

where

$$\varepsilon^* = \frac{(\varepsilon+c^2)^2-c^4}{2c^2} = \varepsilon\left(1+\frac{\varepsilon}{2c^2}\right),$$
$$Z^* = \frac{Z(\varepsilon+c^2)}{c^2} = Z\left(1+\frac{\varepsilon}{c^2}\right),$$
$$l_1^* = \begin{cases}\gamma, & k>0\\ \gamma-1, & k<0,\end{cases} \tag{38}$$
$$l_2^* = \begin{cases}\gamma-1, & k>0\\ \gamma, & k<0.\end{cases}$$

Differential equations (37) differ from the equations in [21] because they are written in a form more convenient for solving the problem formulated here. The regular, F_1, and irregular, G_1, solutions of the first equation in (37), as well as the regular, F_2, and irregular, G_2, solutions of the second equation, can be found by using the same procedure [19] as in the nonrelativistic case. For the relativistic Coulomb functions, we can then easily obtain

$$P_r = \frac{N_0}{1-X^2}(F_1-XF_2),$$
$$Q_r = \frac{N_0}{1-X^2}(F_2-XF_1),$$
$$P_i = \frac{N_0}{1-X^2}(G_1-XG_2), \tag{39}$$
$$Q_i = \frac{N_0}{1-X^2}(G_2-XG_1),$$

where the normalization N_0 is defined by

$$N_0^2 = \frac{(1-X^2)}{c\pi\gamma p}\left[c(\gamma+|k|)+\frac{\varepsilon|k|}{c}\right]. \tag{40}$$

3. RESULTS AND DISCUSSION

We calculated the anisotropy parameters α_2 of the angular Auger electron distribution for $(M_3 \longrightarrow N_2N_2)$,

Table 1. The coefficients α_2 for some of the Xe Auger transitions in the approximation of *LS* coupling

Term	α_2^{11}	α_2^{00}	α_2 [10]	α_2 [9]	α_2 [24]	α_2 [25]
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^1S_0$	−1.0000	−1.0000		−1.000		−1.000
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^3P_0$	−1.0000	−1.0000		−1.000		−1.000
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^3P_1$	−0.8000	−0.8000		−0.800		−0.800
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^3P_2$	0	0.7100		0		0
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^1D_2$	−0.2240	−0.1917		−0.167		−0.189
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^3F_2$	0.5817	0.5867		0.558	0.55	0.607
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^3F_3$	0.4597	0.4659		0.43	0.42	0.493
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^3F_4$	−0.7513	−0.7390		−0.806	−0.82	−0.608
$\mathrm{Xe}(M_4 \longrightarrow N_{4,5}N_{4,5})^1G_4$	−0.6203	−0.6144		−0.640		−0.499
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^1S_0$	−1.0690	−1.0690	−1.069	−1.069		−1.069
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^3P_0$	−1.0690	−1.0690	−1.069	−1.069		−1.069
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^3P_1$	−0.7483	−0.7483	−0.749	−0.748		−0.748
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^3P_2$	−0.3818	−0.3818	−0.371	−0.382		−0.382
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^1D_2$	−0.2394	−0.2050	−0.124	−0.178		−0.202
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^3F_2$	0.5157	−0.7134	0.738	0.0056	−0.02	0.115
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^3F_3$	0.3695	0.3338	0.336	0.322	0.32	0.412
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^3F_4$	0.4658*	0.3774	0.386	0.435	0.420	0.506
$\mathrm{Xe}(M_5 \longrightarrow N_{4,5}N_{4,5})^1G_4$	−0.6631	−0.6568	−0.710	−0.685		−0.533

* For this line, the experimental values of $\alpha_2 = 0.431 \pm 0.12$ were presented in [23].

Table 2. The coefficients α_2 for some of the Xe transitions calculated in the approximations of *LSJ* and intermediate (I) couplings

Term		$_{(00)}\alpha_2$ [7]	$_{(11)}\alpha_2^{\mathrm{HF}(LSJ)}$	$_{(11)}\alpha_2^{\mathrm{HFD}}$	$\alpha_2^{\mathrm{HFD(I)}}$
$\mathrm{Xe}(M_3 \longrightarrow N_2N_3)^3P_2$	α_2	−0.0905	0	−0.0836	0
$\mathrm{Xe}(M_3 \longrightarrow N_3N_3)^3P_2$	α_2	0.5431	0.8000	0.5332	0.6212
$\mathrm{Xe}(M_4 \longrightarrow N_4N_5)^3F_4$	α_2	−0.6805	−0.7948	−0.6706	−0.8266
$\mathrm{Xe}(M_4 \longrightarrow N_5N_5)^3F_4$	α_2	0.4161	0.4409	0.2703	0.3837
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)^3F_4$	α_2	−0.6041	−0.8004	−0.6034	−0.8314
$\mathrm{Xe}(M_5 \longrightarrow N_5N_5)^3F_4$	α_2	0.1544	0.4370	0.1688	0.3796
$\mathrm{Xe}(M_4 \longrightarrow N_1N_3)^3P_2$	α_2	0.4760	0.6818	0.5058	0.5034

Note: α_2^{HF} (α_2^{HFD}) is the anisotropy parameter for the angular distribution of photoelectrons when the wave functions are calculated by the Hartree–Fock (Hartree–Fock–Dirac) method.

Table 3. The coefficients α_2 for some of the Xe Auger transitions in the approximations of *LS* and intermediate couplings

Auger transition	Experiment [26]	MHFD [10]	HF(*LS*)	HFD(I)
$\mathrm{Xe}(N_4 \longrightarrow O_{2,3}O_{2,3})^1S_0$		−1.000	−1.000	−1.0000
$\mathrm{Xe}(N_4 \longrightarrow O_{2,3}O_{2,3})^3P_2$	0.72 ± 0.13	0.231	0.000	−0.1674
$\mathrm{Xe}(N_4 \longrightarrow O_{2,3}O_{2,3})^3P_1$	−0.73 ± 0.11	−0.837	−0.800	−0.8321
$\mathrm{Xe}(N_4 \longrightarrow O_{2,3}O_{2,3})^3P_0$		−1.000	−1.000	−1.0000
$\mathrm{Xe}(N_4 \longrightarrow O_{2,3}O_{2,3})^1D_2$	0.05 ± 0.06	−0.116	0.5160	0.3634
$\mathrm{Xe}(N_5 \longrightarrow O_{2,3}O_{2,3})^1S_0$		−1.069	−1.069	−1.0690
$\mathrm{Xe}(N_5 \longrightarrow O_{2,3}O_{2,3})^3P_2$	0.47 ± 0.13	−0.385	−0.382	−0.2017
$\mathrm{Xe}(N_5 \longrightarrow O_{2,3}O_{2,3})^3P_1$	−0.77 ± 0.17	−0.743	−0.748	−0.7309
$\mathrm{Xe}(N_5 \longrightarrow O_{2,3}O_{2,3})^3P_0$	−1.07 ± 0.10	−1.069	−1.069	−1.0690
$\mathrm{Xe}(N_5 \longrightarrow O_{2,3}O_{2,3})^1D_2$	0.24 ± 0.10	0.094	0.551	0.6167

($M_3 \longrightarrow N_3N_3$), ($M_4 \longrightarrow N_1N_3$), ($M_4 \longrightarrow N_4N_4$), ($M_4 \longrightarrow N_5N_5$), and ($M_{4,5} \longrightarrow O_{2,3}O_{2,3}$) transitions in a Xe atom. We obtained satisfactory agreement with the only known experimental value of α_2 for the Xe ($M_5 \longrightarrow N_{4,5}N_{4,5}$) transition, $\alpha_2 = 0.431 \pm 0.12$ [23] (see Table 1). The calculated value of α_2 in the relativistic approximation for intermediate coupling is $\alpha_2 = 0.3796$ (see Table 2). For ($M_{4,5} \longrightarrow O_{2,3}O_{2,3}$) Auger transitions, the agreement with the experimental value is much poorer. The parameters α_2 for some of the transitions differ even in sign (see Table 3), indicating that further studies should be carried out.

In our calculations, we took into account the exchange interaction and orthogonalized the wave function of the partial Auger electron wave to the core wave functions of A^{2+} by using the Lagrange factors. Our calculations for *jj* coupling are presented in Table 4, where the following notation is used: α_2^{00}—the calculations without orthogonalization and exchange, α_2^{01}—the calculations without orthogonalization, and α_2^{11}—the calculations with both orthogonalization and exchange. We see from this table that the values of α_2 greatly differ, depending on whether orthogonalization and exchange were taken into account for various transitions. Our calculations presented in Tables 1 and 2 are in good agreement with the calculations of other authors [7, 9, 10]. In these papers, the atomic wave functions were determined by solving the relativistic Hartree–Fock–Dirac equation using the procedure described in [27]. Our calculations of the parameter α_2 for ($M_{4,5} \longrightarrow N_{4,5}N_{4,5}$) Auger transitions in a Xe atom in the approximations of both *LS* and *jj* couplings show that an allowance for the exchange and orthogonalization of the Auger electron wave function to the core orbitals and an allowance for relativistic effects have similar effects on the parameter α_2. The changes in α are largest for intermediate coupling in the multiconfiguration case. As was noted in [7], for *LS* coupling, the orbital angular momentum of the partial Auger electron

Table 4. The coefficients α_2 calculated in the approximation of *jj* coupling for the Auger process in Xe ($M_{4,5} \longrightarrow N_{4,5}N_5$), where the orbital angular momentum of the Auger electron is $l'_1 = 2, 4, 6$

Term	*jj*	α_2^{00}	α_2^{01}	α_2^{11}
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{3}{2}\right)_0$	−1.0690	−1.0690	−1.0690
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{3}{2}\right)_2$	−0.3059	0.2914	−0.7881
$\mathrm{Xe}(M_4 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{5}{2}\right)_1$	−0.8000	−0.8000	−0.8000
$\mathrm{Xe}(M_4 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{5}{2}\right)_2$	0.0408	0.0258	0.0247
$\mathrm{Xe}(M_4 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{5}{2}\right)_3$	0.4616	0.4574	0.4553
$\mathrm{Xe}(M_4 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{5}{2}\right)_4$	−0.6398	−0.6644	−0.4553
$\mathrm{Xe}(M_4 \longrightarrow N_5N_5)$	$\left(\frac{5}{2}, \frac{5}{2}\right)_0$	−1.0000	−1.0000	−1.0000
$\mathrm{Xe}(M_4 \longrightarrow N_5N_5)$	$\left(\frac{5}{2}, \frac{5}{2}\right)_2$	−0.7806	−0.7605	−0.7591
$\mathrm{Xe}(M_4 \longrightarrow N_5N_5)$	$\left(\frac{5}{2}, \frac{5}{2}\right)_4$	−0.1992	0.1426	−0.1441
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{3}{2}\right)_0$	−1.0690	−1.0690	−1.0690
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{3}{2}\right)_2$	−0.7877	−0.7873	−0.7881
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{5}{2}\right)_1$	−0.7483	−0.7483	−0.7483
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{5}{2}\right)_2$	−0.1198	−0.1346	−0.1372
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{5}{2}\right)_3$	0.3719	0.3665	0.3639
$\mathrm{Xe}(M_5 \longrightarrow N_4N_5)$	$\left(\frac{3}{2}, \frac{5}{2}\right)_4$	−0.5858	−0.5951	−0.5972
$\mathrm{Xe}(M_5 \longrightarrow N_5N_5)$	$\left(\frac{5}{2}, \frac{5}{2}\right)_0$	−1.0690	−1.0690	−1.0690
$\mathrm{Xe}(M_5 \longrightarrow N_5N_5)$	$\left(\frac{5}{2}, \frac{5}{2}\right)_2$	−0.3255	−0.3376	−0.3401
$\mathrm{Xe}(M_5 \longrightarrow N_5N_5)$	$\left(\frac{5}{2}, \frac{5}{2}\right)_4$	0.1994	0.1874	0.1839

waves can take on values of $l = 2$ and 4, while for *jj* coupling, the Auger electron states with $l = 6$ are also taken into account. Our numerical calculations for the coefficients α_2 at $l = 2$, 4, and 6 are given in Table 4. Note that the partial waves with $l = 2$ and/or $l = 4$ mainly contribute to the parameter α_2.

We calculated the angular anisotropy parameters of the Auger electrons produced when electrons are scattered by atoms "from the first principles." We used various types of coupling as the approximation: *LS*, *jj*, and intermediate couplings for single-electron and multielectron wave functions. Our calculations show that the coefficients α_2 are extremely sensitive to the method of calculations. This sensitivity can be used as a test for the method of calculations for a wide range of problems of particle scattering by atoms.

ACKNOWLEDGMENTS

This study was supported by the Integration Program (project no. L-01-02).

REFERENCES

1. W. Mehlhorn, Phys. Lett. **26**, 166 (1986).
2. H. Winick and S. Doniach, *Sinhrotron Radiation Researches* (Wiley, New York, 1980).
3. J. Eichler and W. Fritsch, J. Phys. B: At. Mol. Opt. Phys. **9**, 1477 (1976).
4. E. G. Berezhko and N. M. Kabachnik, J. Phys. B: At. Mol. Opt. Phys. **10**, 2467 (1977).
5. H. Klar, J. Phys. B: At. Mol. Opt. Phys. **13**, 4741 (1980).
6. K. Blum, B. Lohmann, and E. Taute, J. Phys. B: At. Mol. Opt. Phys. **19**, 3915 (1986).
7. B. Lohmann, J. Phys. B: At. Mol. Opt. Phys. **23**, 3147 (1990).
8. N. M. Kabachnik, H. Aksela, and S. Ricz, Phys. Rev. A **49**, 4653 (1994).
9. M. H. Chen, Phys. Rev. A **45**, 1684 (1992).
10. J. Tulkki, N. M. Kabachnik, and H. Aksela, Phys. Rev. A **48**, 1277 (1993).
11. N. M. Kabachnik, J. Tulkki, H. Aksela, and S. Ritz, Phys. Rev. A **49**, 4653 (1994).
12. B. Cleff and W. Mehlhorn, J. Phys. B: At. Mol. Opt. Phys. **7**, 593 (1974).
13. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Sci., Singapore, 1988).
14. I. I. Sobel'man, *Atomic Spectra and Radiative Transitions* (Fizmatgiz, Moscow, 1963; Springer, Berlin, 1979).
15. W. N. Asaad, Nucl. Phys. **44**, 415 (1963).
16. I. P. Grant, Adv. Phys. **19**, 747 (1970).
17. R. Karazija, *The Theory of X-ray and Electronic Spectra of Free Atoms* (Vilntus, Mokslas, 1987), p. 274.
18. *Handbook of Mathematical Functions*, 2nd ed., Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
19. A. R. Barnett, Comput. Phys. Commun. **24**, 141 (1981).
20. M. E. Rose, *Relativistic Electron Theory* (Wiley, New York, 1961).
21. R. A. Swainson and G. W. F. Drake, J. Phys. A **24**, 79 (1991).
22. L. Infeld, Phys. Rev. **59**, 737 (1941).
23. U. Hahn, J. Semke, H. Merc, and J. Kessler, J. Phys. B: At. Mol. Opt. Phys. **18**, L417 (1985).
24. N. M. Kabachnik and I. P. Sazhina, J. Phys. B: At. Mol. Opt. Phys. **21**, 267 (1988).
25. N. M. Kabachnik, I. P. Sazhina, I. S. Lee, and O. V. Lee, J. Phys. B: At. Mol. Opt. Phys. **18**, L417 (1985).
26. B. Kammerling, V. Schmidt, W. Mehlhorn, *et al.*, J. Phys. B: At. Mol. Opt. Phys. **22**, L597 (1989).
27. P. Grant, B. J. McKenzie, P. H. Norrington, *et al.*, Comput. Phys. Commun. **21**, 207 (1980).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 668–677.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 744–754.
Original Russian Text Copyright © 2003 by Bodrov, Gil'denburg, Sergeev.

ATOMS, SPECTRA, RADIATION

Quasi-Stationary Stage of a Nonequilibrium Optical Discharge in the Field of a Bessel Wave Beam

S. B. Bodrov*, V. B. Gil'denburg, and A. M. Sergeev
Institute of Applied Physics, Russian Academy of Sciences, Nizhni Novgorod, 603950 Russia
**e-mail: bosbor@mail.nnov.ru*

Received January 27, 2003

Abstract—The self-consistent stationary distributions of the field and plasma produced by Bessel wave beams in a gas with the ionization and Kerr nonlinearities are studied analytically. Using a stationary model based on the condition of a constant field amplitude in the ionized region, the structures are considered to be formed by Bessel beams of two types: with an azimuthal electric field equal to zero at the symmetry axis and a quasi-transverse field having a maximum at the axis. A specific feature of the plasma channel formed in the first case is the presence of a nonionized region in its central part (tubular discharge), whose radius is independent of the incident power. In the second case, the channel is continuous. The relation is found between the incident radiation power and the external radius of the discharge. It is shown that the Kerr nonlinearity, which is especially important at small divergence angles of the beam, enhances the maximum plasma density and reduces the discharge radius. The parameters of plasma structures produced upon focusing a Gaussian beam by a conical lens are estimated using the model proposed. © *2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The study of nonequilibrium discharges produced in the axial region of focused laser beams attracts attention due to the possibility of using them for fabrication of extended plasma waveguides on which modern projects for developing X-ray lasers and electron accelerators are based. In this respect, filamentary discharges produced upon focusing a beam by a conical lens (axicon) are very promising [1–5]. Axicon focusing provides independent distributed input of the electromagnetic energy into a plasma filament over its entire length, thereby increasing its axial stability and excluding the restrictions imposed on its length by the defocusing action of the plasma. Figure 1 schematically shows the axicon focusing of a wave beam and a discharge produced in its axial region. The characteristic longitudinal (L) and transverse (d) dimensions of the discharge are determined, when the threshold breakdown intensity I_{th} is not greatly exceeded, by the corresponding dimensions of the axial caustics of the beam; in particular, for a wavelength of $\lambda \sim 1$ μm, an input radius of the beam of $R \sim 1$ cm, and a focusing angle of $\vartheta \sim 10^{-1}$, these dimensions are equal to $L \sim R/\tan\vartheta \sim 10$ cm and $d \sim \lambda/\sin\vartheta \sim 10$ μm.

The field of a wave beam formed by the axicon is close to that of a cylindrical wave. The radial distribution of the beam is determined over a rather large length by the Bessel function (the transverse component of the electric field $E_\perp \sim J_0(k_\perp r)$, where $k_\perp$ is the transverse wave number); for this reason, the beam itself is often called a Bessel beam. Using a special phase corrector placed in front of the axicon, the beams with nonzero azimuthal indices $m = 1, 2, 3, \ldots$, can be formed, whose transverse field at the axis is zero [$E_\perp \sim J_m(k_\perp r)\exp(im\varphi)$; r and φ are polar coordinates in the cross section plane]. The dynamics and structure of discharges in Bessel beams of different types (both with a maximum and a minimum at the beam axis) were studied experimentally and by numerical simulations by several research groups (see papers [6–10] and references therein). In [6–9], studies were performed on comparatively long time scales, when the discharge dynamics is essentially determined by the heating and gas-dynamic expansion of the plasma being formed. These processes result in the formation, over the course of ~100 ps, of the required waveguide structure with the plasma density minimum at the axis [8, 9] and, as follows from the numerical solution of the self-consistent problem [9], can be accompanied by the self-capture of ionizing radiation by the eigenmodes of the plasma waveguide

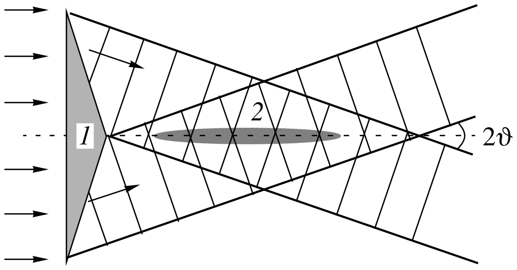

Fig. 1. Schematic view of axicon focusing and the breakdown region: (*1*) conic lens, (*2*) discharge region.

1063-7761/03/9704-0668$24.00 © 2003 MAIK “Nauka/Interperiodica”

being formed. On a time scale of 1–10 ns, the same thermal and gas-dynamic processes can produce the longitudinal self-modulation of the plasma [4–7], resulting in the destruction of a regular waveguide channel. The authors of [7] also experimentally produced a plasma structure of the waveguide type (persisting at the later stages of the process) with the density minimum at the axis, which had formed already at the initial stage of the medium breakdown in the field of a "tubular" Bessel beam whose radial structure is described by the nonzero-order Bessel function.

When the gas breakdown is produced by sufficiently intense and short (<10 ps) laser pulses, the effects of the heating and motion of the gas cannot be manifested until the pulse ends. One can expect that in this case, due to the screening produced by the plasma being formed and a rather strong dependence of the ionization rate on the field amplitude, fast (avalanche or tunneling) ionization should give rise to the establishment of a nonequilibrium quasi-stationary state, in which the gas still remains cold and immobile, while the electric-field amplitude in the plasma is reduced to the level corresponding to the ionization threshold. Such a stationary regime (which we will call a cold discharge) can be set for rather short times (<1 ps) and can be the initial stage of any discharge (before the appearance of noticeable changes in the gas temperature) in long pulses.

Experiments on the axial breakdown of a solid dielectric produced by high-power femtosecond laser pulses are described in [10]. However, the direct observation and measurement of the parameters of the quasi-stationary state in these experiments, as in experiments with long laser pulses, were complicated because of the limited time resolution of the detecting system. It is possible that such measurements can be performed by using weak probe beams.

The aim of this paper is to study analytically the possible self-consistent stationary distributions of the field and plasma in Bessel beams of two types (with a zero or a maximum of the field at the beam axis). The plasma density and the amplitude of a cylindrical wave forming the Bessel beam are treated as functions of the distance to the symmetry axis and are independent of the longitudinal coordinate. The analysis is performed using a rather simple approximate model, in which the discharge plasma is considered as a medium with a specified level of the nonlinear self-restriction of the electric-field amplitude. Such a model was proposed and used earlier for describing different types of stationary structures of microwave [11–13] and optical [14] discharges. As in [14, 15], along with ionization nonlinearity, we also take into account the Kerr (electronic) nonlinearity, which causes the bending of ray trajectories in the wave beam outside the discharge.

2. BASIC EQUATIONS AND APPROXIMATIONS

The rate w of the plasma concentration (density) N changing averaged over the field period, which is defined at each point as the difference of the local ionization rates w_i and recombination losses w_r, is in the general case a function of the electron concentration, concentration N_m of neutral particles, and amplitude $|\mathbf{E}|$ of an alternating electric field:

$$\frac{\partial N}{\partial t} = w_i - w_r = w(N, N_m, |\mathbf{E}|). \tag{1}$$

Other (nonlocal) mechanisms of the density variation, caused, for example, by the action of the ponderomotive force or the transfer of particles and heat, have no time to play any significant role under the conditions of optical breakdown of the gas by sufficiently short pulses of interest to us. In this case, the dependence $w(|\mathbf{E}|)$ is rather strong in the vicinity of a threshold value E_{th} of the amplitude, which is usually defined as the threshold breakdown field. Upon gas breakdown in a Bessel beam with the frequency ω and the transverse wavenumber $k_\perp = (\omega/c)\sin\vartheta$ (ϑ is the convergence angle of the beam), the density of the plasma being produced rather rapidly achieves the value

$$N \sim N_s = N_{\text{cr}}\sin^2\vartheta$$

($N_{\text{cr}} = m\omega^2/4\pi e^2$ is the critical concentration) at which the plasma's influence on the value and structure of the electromagnetic field becomes noticeable. Because of the screening produced by the plasma, the field amplitude in the discharge decreases, resulting in turn in a drastic reduction in the rate of increase in concentration, which virtually stops. Therefore, the electromagnetic field is in fact self-restricted in the discharge at the level E_{th}, after which the discharge remains in the quasi-stationary state for some time (until the beginning of gas-dynamic processes). In this state, the field inside the plasma can be assumed equal to the threshold one, which allows us to find the discharge structure by solving the corresponding inverse problem of electrodynamics. This problem is stated as follows: given the parameters of the incident wave, it is required to determine the distribution of the plasma density $N(r)$ for which the solution of Maxwell's equations (for the specified parameters of the incident cylindrical wave) satisfies the conditions (see also [11–14])

$$\begin{aligned} |\mathbf{E}| &= E_{\text{th}} \quad \text{in the region} \quad N > 0, \\ |\mathbf{E}| &< E_{\text{th}} \quad \text{in the region} \quad N = 0. \end{aligned} \tag{2}$$

Note that this approach provides high accuracy in calculating self-consistent stationary distributions of the field and plasma in the discharge at a sufficiently high level of recombination losses, when the dependence

$w(|\mathbf{E}|)$ has a sufficiently high steepness in the vicinity of point $|\mathbf{E}| = E_{\text{th}}$, defined by the condition $w = w_i - w_r = 0$. In the absence of recombination, comparatively narrow long-lived regions with the density exceeding the values calculated from condition (2) can be formed at the initial stage of the breakdown. However, in the case under consideration, the solution obtained using this condition also allows one to approximately calculate the basic structural characteristics of the stationary state and to find their dependence on the initial parameters of the problem.

We will write the stationary wave equation for the vector of the complex amplitude of the electric field $\mathbf{E}(\mathbf{r})\exp(-i\omega t)$ in the plasma and the surrounding neutral medium in the form

$$-[\nabla \times [\nabla \times \mathbf{E}]] + k_0^2[1 - n(1 - i\delta) + \beta|\mathbf{E}|^2]\mathbf{E} = 0. \quad (3)$$

Here, $k_0 = \omega/c$ is the wave number in vacuum, $\delta = \nu/\omega$, ν is the effective frequency of collisions of electrons,

$$n = \frac{N}{N_{\text{cr}}}(1 + \delta^2)^{-1}$$

is the dimensionless electron concentration, and β is the Kerr nonlinearity parameter—the coefficient in the expansion of the dielectric constant $\varepsilon = \varepsilon_L + \beta|\mathbf{E}|^2$, which is determined by the nonlinear polarizability of the medium (deformation of the electron orbits and the shift of the electronic levels of atoms in a strong optical field). We neglect the linear polarizability of atoms, assuming that $\varepsilon_L = 1 - n(1 - i\delta)$. The value of β in gases at atmospheric pressure is typically (1–10) × 10^{-21} cm^2/V^2. The coefficient β increases with pressure proportionally to the concentration of neutral particles. The Kerr nonlinearity changes the real angle of radiation incidence on the breakdown region, and this nonlinearity proves to be substantial at small convergence angles ϑ of the unperturbed beam.

The solution of the formulated problem is significantly simplified when a discharge has a cylindrical symmetry and the electric-field vector is represented (exactly or approximately) by one scalar projection. We consider below two such models: (1) a discharge in the field of a symmetric TE-wave with the azimuthal electric field E_φ, which vanishes at the symmetry axis (vector model), and (2) a discharge produced by the quasi-transverse field E_x of a broad Bessel beam with a maximum at the axis (scalar model). In the first case, the problem admits, within the framework of approach (2), (3), a comparatively simple exact solution for any value of the beam convergence angle ϑ. In the second case, the problem can be solved in the scalar (paraxial) approximation for $\vartheta \ll 1$. In both cases, we are interested in the solutions of Eq. (3), which can be written in cylindrical coordinates r and z in the form

$$\mathbf{E}(\mathbf{r}) = \boldsymbol{\xi}_0 E(r)\exp(ihz); \quad (4)$$

i.e., we represent an inhomogeneous traveling wave, which is inhomogeneous over the radius and has a fixed longitudinal wave number $h = k_0\cos\vartheta$ and polarization specified by a unit vector $\boldsymbol{\xi}_0$.

3. DISCHARGE IN A SYMMETRIC TE-WAVE WITH THE AZIMUTHAL ELECTRIC FIELD (EXACT SOLUTION OF THE VECTOR PROBLEM)

Consider a discharge maintained by the field of a symmetric TE-wave with an azimuthal electric field. Assuming $\boldsymbol{\xi}_0 = \boldsymbol{\varphi}_0$ in expression (4) ($\boldsymbol{\varphi}_0$ is the unit vector in the direction in which azimuthal angle φ is being measured) and representing the corresponding projection of the field in the form $E = A\exp(i\psi)$, we obtain from (2) and (3) the equation

$$\frac{\partial^2 E}{\partial^2 r} + \frac{1}{r}\frac{\partial E}{\partial r} - \frac{E}{r^2} + k_\perp^2 E + k_0^2\beta|E|^2 E = 0 \quad (5)$$

for the complex amplitude E in the nonionized region (for $n = 0$), and the equations

$$\frac{\partial^2 \psi}{\partial^2 r} + \frac{1}{r}\frac{\partial \psi}{\partial r} = -k_0^2\delta n, \quad (6)$$

$$\left(\frac{\partial \psi}{\partial r}\right)^2 = k_0^2(n_m - n) - \frac{1}{r^2} \quad (7)$$

for the real functions $\psi(r)$ and $n(r)$ in the plasma region ($n > 0$, $A = E_{\text{th}}$).

Here, the notation

$$n_m = (1 + \gamma)\sin^2\vartheta, \quad \gamma = \frac{\beta E_{\text{th}}^2}{\sin^2\vartheta}$$

is introduced. The conditions at the ionized–nonionized region interface, which are equivalent to the conditions of continuity of the tangential components of the electric and magnetic fields, have the form

$$|E| = A = E_{\text{th}}, \quad \frac{\partial A}{\partial r} = 0, \quad \left\{\frac{\partial \psi}{\partial r}\right\} = 0, \quad (8)$$

where the braces denote the jump of a quantity at the interface.

Because n is positive, Eq. (7) cannot be satisfied in the axial region and, hence, within the stationary model used here, the discharge should have an inner cavity where the gas remains nonionized. In the general case, the plasma density can differ from zero only inside some circular region $r_0 \le r \le r_1$; i.e., the stationary discharge under study represents a cylindrical tube.

In the inner region $r < r_0$, where plasma is absent, the phase ψ = const stems from the absence of any external sources at the axis. In this case, in the absence of Kerr nonlinearity (for $\beta = 0$), the solution of Eq. (5) is the first-order Bessel function $E \sim J_1(k_\perp r)$, and the radius r_0 of the inner cavity, as follows from boundary conditions (8), is determined by the position of the first maximum (i.e., by the first zero of the derivative) of this function:

$$r_0 = \bar{r}_0 \approx \frac{1.84}{k_0 \sin\vartheta}.$$

The Kerr nonlinearity causes an increase in the effective transverse wave number, thereby decreasing the internal radius of the plasma tube: $r_0 = \bar{r}_0 f(\gamma)$. The dependence $f(\gamma)$ found by numerically solving Eq. (5) is shown in Fig. 2. Note that the radius r_0 is determined only by the converging angle ϑ of a cylindrical wave and the parameter γ (r_0 is independent of the incident power and the frequency of electron collisions).

The equation for the concentration in the region $r_0 \le r \le r_1$ can be obtained from Eqs. (6) and (7) by excluding phase ψ from them:

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\sqrt{k_0^2(n_m - n) - \frac{1}{r^2}}\right) = k_0^2 \delta n. \qquad (9)$$

It follows from Eq. (7) and the condition $d\psi(r_0)/dr = 0$ that the plasma density at the internal boundary of the ionized region is

$$n_0 = n(r_0 + 0) = n_m - \frac{1}{(k_0 r_0)^2}.$$

In the absence of collisions (for $\delta = 0$), the condition $d\psi/dr = 0$ is satisfied over the entire space. In this case, Eq. (7) allows us to find function $n(r)$ in the entire region where $n \neq 0$:

$$n = n_m - \frac{1}{(k_0 r)^2}. \qquad (10)$$

In the general case, the solution of Eq. (9) for function $n(r)$ is expressed in terms of complex-order Bessel functions. Because this representation of the solution is cumbersome and not sufficiently descriptive (see Appendix), we present here only the results of its approximate analysis for some limiting cases, as well as the results of numerical calculations.

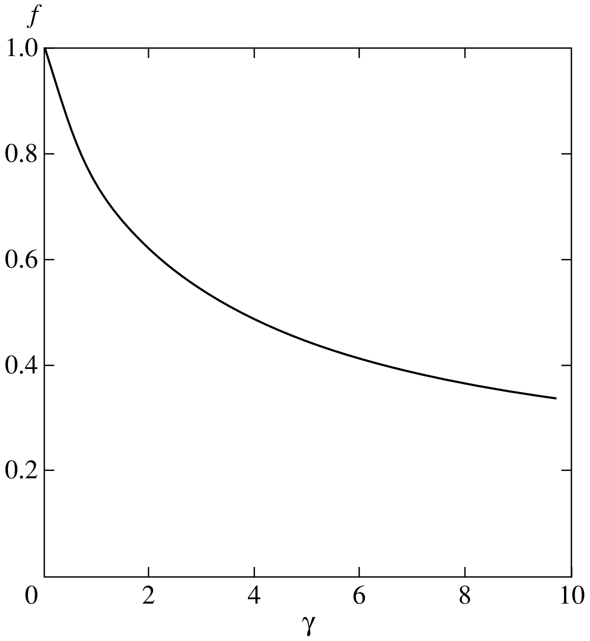

Fig. 2. Dependence of the dimensionless internal radius $f = r_0/\bar{r}_0$ of a "tubular" discharge on the Kerr nonlinearity parameter $\gamma = \beta E_{\rm th}^2/\sin^2\vartheta$ (vector model).

For low collision frequencies of electrons ($\delta \ll 1$), the solution in the region $r \ll 1/k_0\sqrt{n_m}\,\delta$ can be found by the perturbation method by calculating from (9) the first correction to zero-order approximation (10):

$$n = n_m - \frac{1}{(k_0 r)^2} - \left(\frac{\delta}{r}\right)^2\left[\frac{n_m}{2}(r^2 - r_0^2) + \frac{1}{k_0^2}\ln\left(\frac{r_0}{r}\right)\right]^2. \qquad (11)$$

Function $n(r)$ defined by this expression has the maximum

$$n_{\max} \approx \tilde{n} = n_m(1 - \delta)$$

at the point

$$r = \tilde{r} \approx \sqrt{\frac{2}{k_0^2 n_m \delta}}.$$

At large distances from the axis, the plasma density monotonically decreases for any $\delta \neq 0$ and tends to its asymptotic behavior

$$n = \frac{\sqrt{n_m}}{k_0 r \delta}$$

in the region

$$r \gg \frac{1 + \delta}{k_0 \delta \sqrt{n_m}}.$$

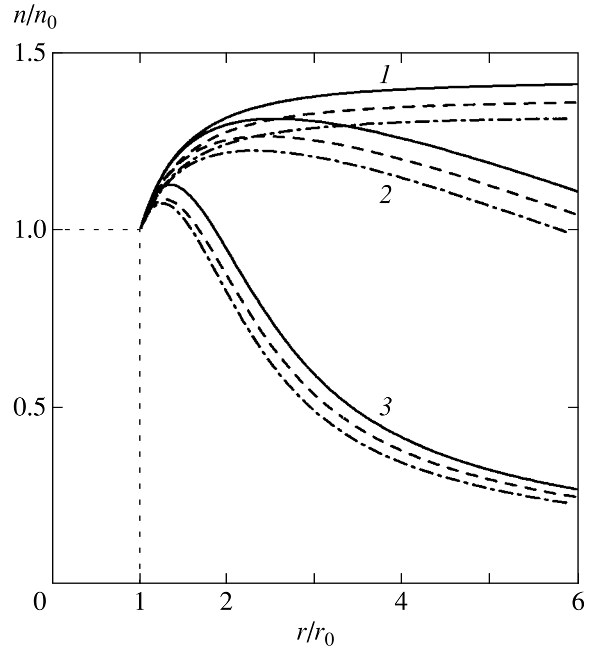

Fig. 3. Radial distributions of the plasma density in a "tubular" discharge for different values of parameters δ and γ (radius r is normalized to the internal radius of the discharge $r_0 = 1.84f/k\sin\vartheta = 0.29\lambda f/\sin\vartheta$; the values of f are determined from the plot in Fig. 2): $\delta = 0$ (*1*), 0.1 (*2*), 0.5 (*3*); $\gamma = 0$ (solid curves), 1 (dashed curves), 10 (dash-dotted curves).

The results of numerical integration of Eq. (9) for different values of δ and γ are shown in Fig. 3.

For any finite value of the specific power P of the incident wave (defined as the energy flux of the incident wave propagating through the surface of a unit-length cylinder embracing the z axis), the density distributions described above should be truncated at a point $r = r_1$ to the right of which the plasma density again vanishes $[n(r > r_1) = 0]$. Position r_1 of this external boundary of the plasma tube, like the specific power Q (per unit length) absorbed by the discharge, depend on parameters P, δ, ϑ, and γ and can be found by solving Eq. (5) with boundary conditions (8) at $r = r_1$ and the radiation condition for the reflected-wave field at infinity. In particular, in the absence of Kerr nonlinearity if the condition $k_\perp r_1 \gg 1$ is fulfilled, when the solution is expressed in terms of Bessel functions, and the latter are close to their long-range asymptotics, the field in the external region can be represented as a superposition of the incident and reflected waves:

$$E = A\exp(i\psi) = \frac{C_i}{\sqrt{r}}\exp(-ik_\perp r) + \frac{C_i}{\sqrt{r}}\exp(ik_\perp r). \qquad (12)$$

Boundary conditions (8) allow us to relate the incident-wave amplitude $|C_i|r^{-1/2}$ with the phase derivative $d\psi(r)/dr$ for $r = r_1$:

$$\frac{C_i}{\sqrt{r_1}} = \frac{1}{2}E_{\rm th}\left(\frac{1}{k_0\sin\vartheta}\frac{d\psi(r_1)}{dr} + 1\right). \qquad (13)$$

By using the expressions for the incident and absorbed powers

$$P = \frac{c}{4}|C_i|^2\sin\vartheta, \quad Q = \frac{cr_i}{4k_0}\frac{d\psi(r_1)}{dr}E_{\rm th}^2, \qquad (14)$$

and Eq. (7), which allows us to express the derivative $d\psi/dr$ in terms of the plasma density, we obtain from (13) the dependences of P and Q on the radius r_1 for $r_1 \gg r_0$:

$$P = \frac{cr_1}{16}\sin\vartheta E_{\rm th}^2 \times\left(1 + \frac{1}{\sin\vartheta}\sqrt{\sin^2\vartheta - n(r_1) - (k_0r_1)^{-2}}\right)^2, \qquad (15)$$

$$Q = \frac{cr_1}{4}E_{\rm th}\sqrt{\sin^2\vartheta - n(r_1) - (k_0r_1)^{-2}}. \qquad (16)$$

The quantity $n(r_1)$ in expressions (15) and (16) is determined from the solution of Eq. (9). To maintain a discharge with radius r_1 of the external boundary in the absence of absorption ($\delta = 0$, $d\psi/dr = 0$, $Q = 0$), an incident power of

$$P_0 = \frac{c\sin\vartheta}{16}r_1E_{\rm th}^2$$

is required. When absorption is strong ($k_0r_1\delta\sqrt{\sin\vartheta} \gg 1 + \delta$), the absorbed power is close to the incident power: $Q \approx P = 4P_0$. The threshold specific power $P_{\rm th}$ for maintaining the discharge can be found from the condition $E_{\rm max}^{(0)} = E_{\rm th}$ that the maximum amplitude of the field $E^{(0)}(r)$ unperturbed by the plasma should satisfy. In the absence of Kerr nonlinearity, when $E^{(0)} \sim J_1(rk_\perp)$, we have $P_{\rm th} = 0.12cE_{\rm th}^2/k_0$.

The Kerr nonlinearity gives small corrections to the expressions presented above if $\gamma \ll 1$. In the region $r_1k_1 \gg 1$, these corrections can be taken into account analytically. By making the replacement $Y = E\sqrt{r}$ in (5), we obtain for Y the equation of a nonlinear oscillator with a continuously varying spatial frequency:

$$Y'' + k_\perp^2Y + k_0^2\frac{\beta}{r}|Y|^2Y - \frac{3}{4}\frac{Y}{r^2} = 0. \qquad (17)$$

The boundary condition for the function Y has the form

$$Y(r_1) = E_{\mathrm{th}}\sqrt{r_1}, \quad \frac{dY}{dr}(r_1) = i\frac{d\psi}{dr}(r_1)E_{\mathrm{th}}\sqrt{r_1}. \tag{18}$$

We will seek the solution in the form of counterpropagating waves with slowly varying complex amplitudes $C_i(r)$ and $C_r(r)$,

$$Y = C_i(r)\exp(-ik_\perp r) + C_r(r)\exp(ik_\perp r).$$

At a large distance from the boundary, where the Kerr nonlinearity can be neglected, the amplitudes $|C_i(r)|$ and $|C_r(r)|$ tend to constants C_i and C_r, from which the incident power can be found by using (14). For $\gamma \ll 1$, the approximate solution of Eq. (17) can be found by retaining in the nonlinear term in the left-hand side only the fundamental harmonics $\exp(\pm ik_\perp r)$ and by neglecting higher harmonics. Taking into account boundary conditions (18), this leads to a system of algebraic equations for the amplitudes of the fundamental harmonics, whose solution in the first-order approximation in γ gives

$$|C_{i,r}| = \frac{1}{2}E_{\mathrm{th}}\sqrt{r_1}\left[1 \pm \alpha + \frac{\gamma}{16}(3 \mp \alpha + \alpha^2 \mp 3\alpha^3)\right], \tag{19}$$

where

$$\alpha = \frac{\sqrt{(1+\gamma)\sin^2\vartheta - n(r_1) - (k_0 r_1)^{-2}}}{\sin\vartheta}. \tag{20}$$

In this case, the incident power per unit length is

$$P \approx \frac{cr_1}{16}\sin\vartheta E_{\mathrm{th}}^2(1+\alpha)^2 \times\left[1 + \frac{\gamma}{8}(3 - \alpha + \alpha^2 - 3\alpha^3)\right]. \tag{21}$$

The absorbed specific power is still described by expression (16).

This method also allows one to obtain a qualitatively correct result (with accuracy to factors on the order of unity) in the case of strong Kerr nonlinearity. In particular, in the absence of collisions, we obtain by this method

$$P = \frac{cr_1}{16}\sin\vartheta E_{\mathrm{th}}^2\sqrt{1 + \frac{3}{4}\gamma}. \tag{22}$$

A result close to expression (22) can be obtained in the same quasi-plane approximation ($r_1 k_\perp \gg 1$) by another method. Let us represent the solution of Eq. (17) in the form of the elliptic Jacobian function *cn* with a slowly varying "amplitude" Y_0, "frequency" u, and parameter m:

$$Y(r) = Y_0(r)\mathrm{cn}(u(r), m(r)). \tag{23}$$

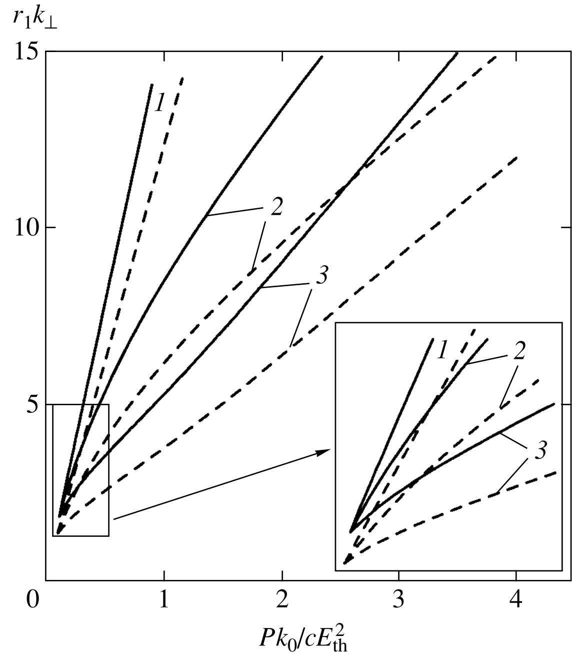

Fig. 4. Dependences of the external radius of the discharge (normalized to the transverse scale $k_\perp = 2\pi/\lambda\sin\vartheta$ of the beam) on the incident power for different parameters δ and γ for the vector model: $\delta = 0$ (*1*), 0.1 (*2*), 0.5 (*3*); $\gamma = 0$ (solid curves) and 1 (dashed curves).

In the case of a one-dimensional plane problem, this method yields an exact solution. By introducing (23) into (17), we can obtain a system of algebraic equations for Y_0, u, and m whose solution, together with boundary conditions (18), gives

$$P = \frac{cr_1}{16}\sin\vartheta E_{\mathrm{th}}^2\sqrt{1+\gamma}\left[1 - \frac{\gamma}{2(1+\gamma)}\right]^{1/4}. \tag{24}$$

For $\gamma \gg 1$, both these methods give the same dependence of the channel radius on the specific power:

$$P = \frac{cr_1}{16}E_{\mathrm{th}}^3\sqrt{\beta}\eta. \tag{25}$$

Depending on the method of solving Eq. (17), constant η is equal to 0.84 [from expression (24)] or 0.86 [from expression (22)]. The numerical calculation gives $\eta = 0.79$. Note that in the case of strong Kerr nonlinearity, the discharge radius is independent of the focusing angle.

Figure 4 shows the dependences of the external radius of the discharge on the incident power calculated numerically for different values of parameters δ and γ (solid and dashed curves). Figure 5 shows the radial

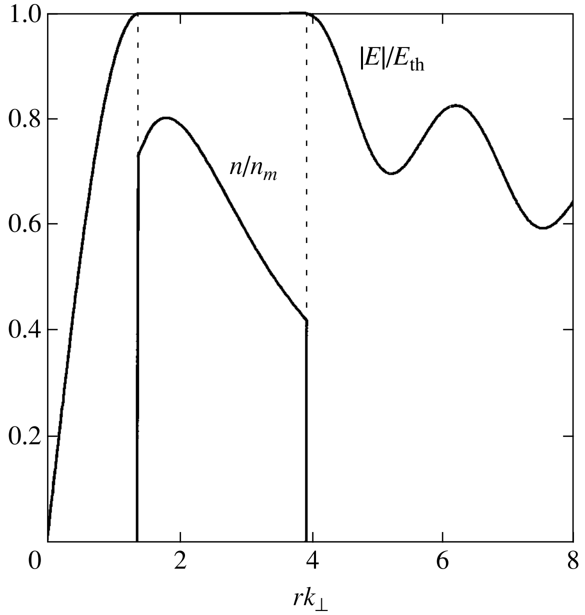

Fig. 5. Radial distributions of the plasma density and the field amplitude for the vector model with $\delta = 0.5$, $\gamma = 1$, $Pk_0/cE^2_{th} = 1$, and $P/P_{th} = 10$.

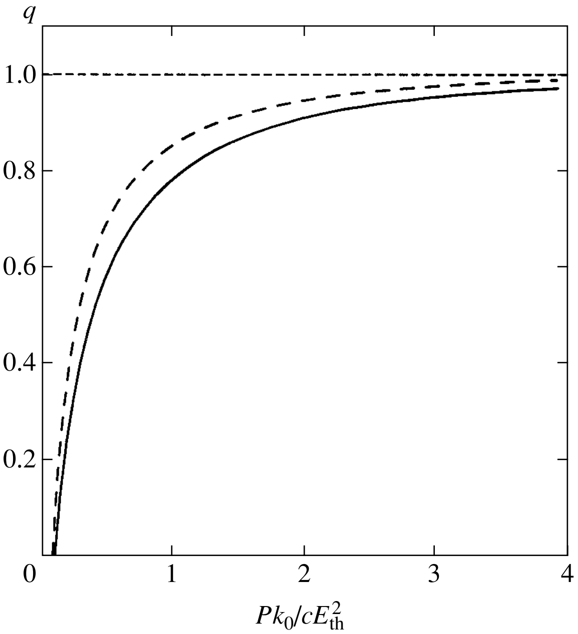

Fig. 6. Dependence of the fraction of absorbed power $q = Q/P$ on the incident power P for the vector model for $\gamma = 0$ (solid curve) and 10 (dashed curve); $\delta = 0.1$.

distributions of the plasma density and the field amplitude for $\delta = 0.5$, $\gamma = 1$, and $Pk_0/cE^2_{th} = 1$. The results of the numerical calculation of the fraction of absorbed energy $q(P) = Q/P$ and the threshold specific power $g(\gamma) = P_{th}(\gamma)/P_{th}(0)$ are presented by solid curves in Figs. 6 and 7.

4. DISCHARGE IN A PARAXIAL BESSEL BEAM WITH THE FIELD MAXIMUM AT THE AXIS (SCALAR APPROXIMATION)

When the condition $\vartheta \ll 1$, $n \ll 1$ is fulfilled, the wave beam can be described using the so-called small-angle (paraxial) approximation. We will see that the second of these conditions for the stationary states under study proves to be always satisfied when the first condition is fulfilled, so that the first condition can be considered as a sufficient condition for the applicability of this approximation. Within the framework of this approximation, the electromagnetic field is almost transverse and vector wave equation (3) is reduced to the scalar Helmholtz equation for the transverse (Cartesian) component $E_x = E(r)\exp(ihz)$ of the electric field:

$$\Delta_\perp E + k_0^2[\sin^2\vartheta - n(1 - i\delta) + \beta|E|^2]E = 0, \qquad (26)$$

where

$$\Delta_\perp = \frac{d^2}{dr^2} + \frac{1}{r}\frac{d}{dr}$$

is the transverse (radial) part of the Laplace operator. For the nonionized region, we should set $n = 0$ in this equation. The equations for phase $\psi(r)$ and plasma density $n(r)$ in the discharge, which are similar to those derived in the previous section, have the form

$$\frac{\partial^2\psi}{\partial^2 r} + \frac{1}{r}\frac{\partial\psi}{\partial r} = -k_0^2\delta n, \qquad (27)$$

$$\left(\frac{d\psi}{dr}\right)^2 = k_0^2(\sin^2\vartheta - n + \beta E^2_{th}), \qquad (28)$$

$$\frac{1}{r}\frac{\partial}{\partial r}(r\sqrt{n_m - n}) = k_0\delta n. \qquad (29)$$

Here, the same notation is used as in the previous section. Unlike the previous case, the discharge structure determined by these equations does not contain an inner cavity: the plasma occupies the region $0 \le r < r_1$, where the field amplitude is $A = E_{th}$ and the plasma density, determined from Eq. (29), is

$$n = n_m\left\{1 - \left[\frac{I_1(\delta\sqrt{n_m}k_0 r)}{I_0(\delta\sqrt{n_m}k_0 r)}\right]^2\right\}, \qquad (30)$$

where I_0 and I_1 are modified Bessel functions. The plasma density achieves the maximum value

$$n = n_m = (1 + \gamma)\sin^2\vartheta$$

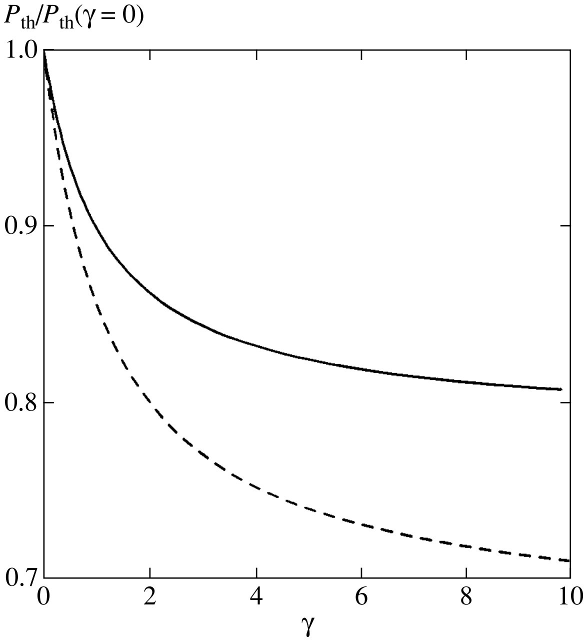

Fig. 7. Dependence of the threshold power on Kerr nonlinearity parameter γ for the vector (solid curve) and scalar (dashed curve) models.

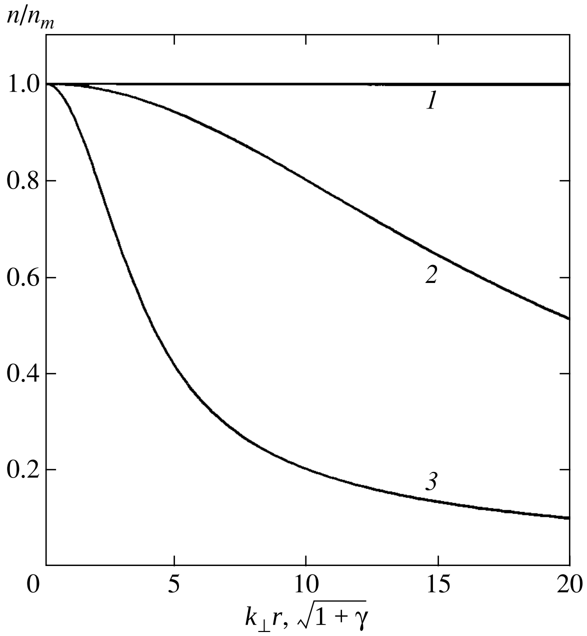

Fig. 8. Radial distributions of the plasma density for the scalar model at δ = 0 (*1*), 0.1 (*2*), and 0.5 (*3*).

for $r = 0$ and gradually decreases with a characteristic scale of $R = 1/k_0\sqrt{n_m}\,\delta$ with increasing r in the region $r < r_1$ (Fig. 8). At the boundary surface $r = r_1$, the plasma density abruptly vanishes. In the absence of collisions (for δ = 0), $n = n_m$ = const over the entire region $r < r_1$. Position r_1 of the boundary and the specific power Q absorbed by the discharge, as in the previous section, depend on the value of the incident specific power P and are determined by matching the solutions for the internal and external regions using the same boundary conditions (8). When $r_1 k_\perp \gg 1$, the quantities

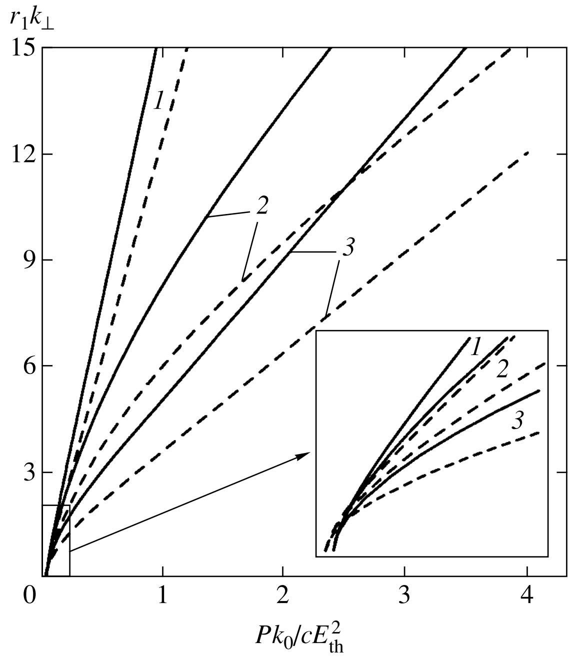

Fig. 9. Dependence of the plasma-channel radius on the incident power for the scalar model for δ = 0 (*1*), 0.1 (*2*), and 0.5 (*3*); and γ = 0 (solid curves) and 1 (dashed curves).

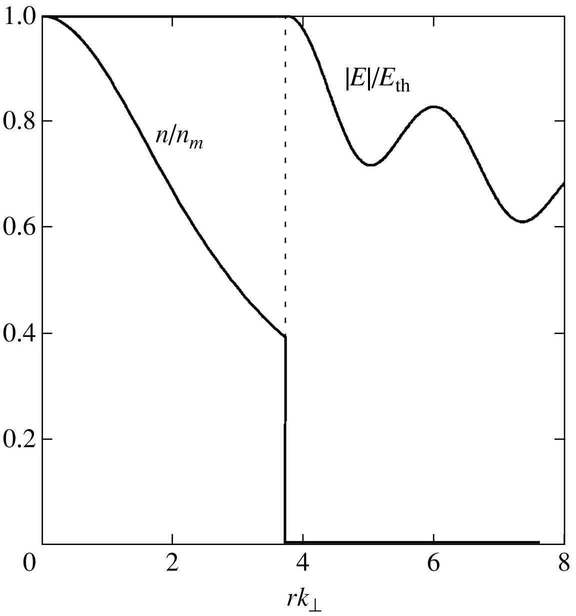

Fig. 10. Radial distributions of the plasma density and the field amplitude for the scalar model for δ = 0.5, γ = 1, $Pk_0/cE_{th}^2 = 1$, and $P/P_{th} = 30$.

r_1 and Q are related to the incident specific power by expressions that differ from similar expressions (15), (16), and (20) only by the absence of the term r^{-2} in the radicands, while expressions (22) and (24) retain their form. In this case, the threshold specific power is three times lower than in the previous case: $P_{\mathrm{th}} = 0.04cE_{\mathrm{th}}^2/k_0$. The dependence of the plasma channel radius on the incident power P for different parameters of the problem and the dependence of the threshold power P_{th} on γ are shown in Figs. 9 and 7 (dashed curves). Figure 10 presents the radial distributions of the plasma density and the field amplitude for $\delta = 0.5$, $\gamma = 1$, and $Pk_0/cE_{\mathrm{th}}^2 = 1$.

5. CONCLUSIONS

Based on the approximate model of a stationary discharge treated as a system with a specified level of nonlinear self-restriction of the field amplitude, we have calculated analytically and numerically the axially symmetric stationary distributions of the plasma density in optical discharges produced by “Bessel” wave beams of different types. In a beam with an azimuthal electric field equal to zero at the symmetry axis, a discharge appears in the form of a plasma tube, whose internal radius is determined by the convergence angle of the beam, and the external radius is determined by the specific power of a converging cylindrical wave forming the beam. Such discharges can be used as waveguide structures producing channeling of the electromagnetic radiation over rather large distances. In a beam with the field maximum located at the axis, the discharge is continuous: the plasma density in the discharge first monotonically decreases with distance from the symmetry axis and then abruptly vanishes. The maximum values of the plasma density in the discharges of both types are proportional to the square of the convergence angle of the beam. It is shown that even weak Kerr nonlinearity in the medium at small convergence angles of the beam strongly affects the maximum plasma density and the position of discharge boundaries.

The solutions found in the paper can be also used for the description of the quasi-stationary states of axicon discharges in longitudinally inhomogeneous Bessel wave beams formed in real laboratory experiments (for example, upon the axial focusing of a Gaussian beam). Under these conditions, the parameters of structures found, in particular, the external radius of the discharge, should be considered as local functions of the incident specific power, which gradually changes along the axial coordinate. Note, however, that stationary discharges are subjected under real conditions to self-focusing [16] and ionization [17] instabilities, which we have not considered here. In particular, the first of them develops in a nonionized medium [4] if the characteristic transverse size $k_\perp^{-1}$ of the beam exceeds the characteristic transverse instability scale $(k_0\sqrt{\beta E_{\max}^2})^{-1}$ [16] ($E_{\max}$ is the maximum field amplitude in the unperturbed beam), i.e., when the condition

$$P > P_{\mathrm{cr}}, \quad P_{\mathrm{cr}} \sim \frac{ck_0\sin^2\vartheta}{10\beta} \tag{31}$$

is fulfilled. If

$$\vartheta < \vartheta_{sf} \sim \sqrt{\frac{10\beta P_{\mathrm{th}}}{ck_0}} \approx \sqrt{\beta E_{\mathrm{th}}^2}$$

and $P_{\mathrm{th}} > P > P_{\mathrm{cr}}$, then gas breakdown occurs only after the development of a self-focusing instability and also can be accompanied by the formation of stationary waveguide channels, in which the defocusing action of the plasma is compensated by Kerr focusing [14, 18].

In conclusion, we estimate the characteristic parameters of a stationary discharge formed upon the axicon focusing of a Gaussian beam with an intensity (in front of the axicon lens) of

$$I_{\mathrm{in}}(r) = I_0\exp\left(-\frac{r^2}{R_0^2}\right)$$

in air at atmospheric pressure. As follows from geometrical–optical calculations [4, 5], the specific power

$$P(z) = I_{\mathrm{in}}(z\tan\vartheta)2\pi z\tan^2\vartheta$$

in a beam converging behind the axicon at a small angle ϑ slowly changes along the longitudinal z axis, achieving a maximum of $P_{\max} \approx 2.7I_0R_0\tan\vartheta$ at a distance of $z_m \approx 2^{-1/2}R_0/\tan\vartheta$. Assuming that $E_{\mathrm{th}} \approx 2.3 \times 10^8$ V cm^{-1} and $\beta = 1.4 \times 10^{-21}$ cm^2/V^2 [18], we find the value of the critical angle $\vartheta_{sf} \approx 0.01$ rad above which the self-focusing instability (before gas breakdown) will not develop, and the breakdown threshold intensity $I_{0\,\mathrm{th}}$ is determined, using the model proposed in the paper, from the expression $I_{0\,\mathrm{th}} = P_{\mathrm{th}}/2.7R_0\tan\vartheta$. For example, for a laser beam of radius $R_0 = 0.5$ cm in front of the axicon and a wavelength of $\lambda = 1$ μm, the value of $I_{0\,\mathrm{th}}$ changes from 9×10^9 to 4×10^{10} W/cm^2 when angle ϑ changes from 0.1 to 0.02 rad. For a maximum input power density of $I_0 = 10^{11}$ W/cm^2, this range of variation of focusing angel ϑ corresponds to changes in the discharge length from 10 to 31 cm, in its maximum radius r_{1m} from 1.2×10^{-3} to 10^{-3} cm, and in the maximum plasma density $N = N_c\sin^2\vartheta$ from 10^{17} to 4×10^{15} cm^{-3}. When radiation is focused at an angle of $\vartheta < \vartheta_{sf}$, the self-focusing instability will develop, resulting in local increase in the field and, hence, in gas breakdown. For example, if $\vartheta = 0.005$ rad and the input power density lies in the range $8 \times 10^{10} < I_0 < I_{0\,\mathrm{th}} = 1.3 \times 10^{11}$ W/cm^2, then gas breakdown can occur only due to the development of this instability. If the input intensity exceeds the threshold value, the gas will be ionized in the absence of instability as well; however, the appearance of instability will result in an earlier breakdown (closer to the axicon) and in an increase in the discharge length. In the latter case, the discharge radius in the region where the specific power exceeds the threshold one ($P > P_{\text{th}}$) and the discharge radius r_1 can be found using the model proposed in the paper. For $I_0 = 2 \times 10^{11}$ W/cm^2, the maximum discharge radius is $r_{1m} = 4 \times 10^{-4}$ cm. The plasma concentration in this case can be found from the expression

$$N = N_c(\sin^2\vartheta + \beta E_{\text{th}}^2) \approx N_c\beta E_{\text{th}}^2 = 1.3\times 10^{15}\ \frac{1}{\text{cm}^3}.$$

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research, project no. 02-02-17271.

APPENDIX

The solution of Eq. (9) in the general case has the form

$$n = n_m - \frac{1}{(k_0 r)^2} - \left(\frac{\partial\psi}{\partial r}\right)^2, \tag{A.1}$$

where

$$\frac{\partial\psi}{\partial r} = k_0\sqrt{n_m}\frac{MI_{i\delta+1}(x) - K_{i\delta+1}(x)}{MI_{i\delta}(x) - K_{i\delta}(x)} + \frac{i}{r}; \tag{A.2}$$

M is a constant determined from the boundary condition $\partial\psi/\partial r = 0$ at the inner boundary r_0 of the discharge:

$$M = \frac{k_0 r_0\sqrt{n_m}K_{i\delta+1}(x_0) - iK_{i\delta}(x_0)}{k_0 r_0\sqrt{n_m}I_{i\delta+1}(x_0) + iI_{i\delta}(x_0)}, \tag{A.3}$$

$$x = \delta\sqrt{n_m}rk_0, \quad x_0 = \delta\sqrt{n_m}r_0k_0,$$

$I_\nu(x)$ and $K_\nu(x)$ are the modified Bessel functions of the first and second kinds.

REFERENCES

1. J. N. McLeod, J. Opt. Soc. Am. **44**, 592 (1954).
2. B. Ya. Zel'dovich, B. F. Mul'chenko, and N. F. Pilipetskiĭ, Zh. Éksp. Teor. Fiz. **58**, 794 (1970) [Sov. Phys. JETP **31**, 425 (1970)].
3. F. V. Bunkin, V. V. Korobkin, Yu. A. Kurinyĭ, *et al.*, Kvantovaya Élektron. (Moscow) **10**, 443 (1983).
4. N. E. Andreev, Yu. A. Aristov, L. Ya. Polonskiĭ, and L. N. Pyatnitskiĭ, Zh. Éksp. Teor. Fiz. **100**, 1756 (1991) [Sov. Phys. JETP **73**, 969 (1991)].
5. N. E. Andreev, L. Ya. Margolin, I. V. Pleshanov, and L. N. Pyatnitskiĭ, Zh. Éksp. Teor. Fiz. **105**, 1232 (1994) [JETP **78**, 663 (1994)].
6. L. Ya. Margolin, L. Ya. Polonskiĭ, and L. N. Pyatnitskiĭ, Pis'ma Zh. Tekh. Fiz. **13**, 18 (1987) [Sov. Tech. Phys. Lett. **13**, 89 (1987)].
7. S. S. Bychkov, S. V. Gorlov, L. Ya. Margolin, *et al.*, Kvantovaya Élektron. (Moscow) **26**, 229 (1999).
8. H. M. Milchberg, T. R. Clark, C. G. Durfee, and T. M. Antonsen, Phys. Plasmas **3**, 2149 (1996).
9. J. Fan, E. Parra, and H. M. Milchberg, Phys. Rev. Lett. **84**, 3085 (2000).
10. A. A. Babin, A. M. Kiselev, K. I. Pravdenko, *et al.*, Usp. Fiz. Nauk **169**, 80 (1999) [Phys. Usp. **42**, 74 (1999)].
11. V. B. Gil'denburg and S. V. Golubev, Zh. Éksp. Teor. Fiz. **67**, 89 (1974) [Sov. Phys. JETP **40**, 46 (1975)].
12. V. B. Gil'denburg, *Nonlinear Waves* (Nauka, Moscow, 1981), p. 87.
13. V. B. Gildenburg, T. A. Petrova, and A. D. Yunakovsky, Physica D (Amsterdam) **87**, 335 (1995).
14. D. Anderson, A. V. Kim, M. Lisak, *et al.*, Phys. Rev. E **52**, 4564 (1995).
15. L. Berge and A. Couairon, Phys. Plasmas **7**, 210 (2000).
16. V. I. Bespalov and V. I. Talanov, Pis'ma Zh. Éksp. Teor. Fiz. **3**, 471 (1967) [JETP Lett. **3**, 307 (1967)].
17. V. B. Gil'denburg and A. V. Kim, Zh. Éksp. Teor. Fiz. **74**, 141 (1978) [Sov. Phys. JETP **47**, 72 (1978)].
18. A. Braun, G. Korn, X. Liu, *et al.*, Opt. Lett. **20**, 73 (1995).

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 678–687.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 755–765.
Original Russian Text Copyright © 2003 by Grigoryan, Leonov, Manykin, Rudenko, Sitnikov, Starostin.

ATOMS, SPECTRA, RADIATION

The Near Infrared (0.8–2.6 μm) Absorption Spectrum of a Dense Sodium Vapor and Possible Mechanisms of the Spectrum Formation

G. G. Grigoryan[a], A. G. Leonov[b],*, É. A. Manykin[a], A. A. Rudenko[b], M. G. Sitnikov[a], and A. N. Starostin[c]

[a]*Kurchatov Institute, Russian State Scientific Center, Moscow, 123182 Russia*
[b]*Moscow Institute of Physics and Technology (State University), Dolgoprudnyĭ, Moscow oblast, 141700 Russia*
[c]*Troitsk Institute of Innovation and Fusion Research, State Scientific Center of the Russian Federation, Troitsk, Moscow oblast, 142190 Russia*
**e-mail: leonov@ape.relarn.ru*

Received February 20, 2003

Abstract—The absorption coefficient of a dense sodium vapor ($N_0 \sim 10^{17}$–10^{18} cm^{-3}) in the near infrared spectral range (0.8–2.6 μm) was measured in a homogeneously heated isolated cell. In the range of parameters studied, the sample exhibits significant absorption. Neither the observed spectral features nor the measured absorption coefficients can be explained using the existing notions of the possible absorption mechanisms (absorption due to a forbidden intercombination transition, collision-induced processes, the trimer vapor component, and many-particle effects) and the available data. *© 2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The optical properties of a dense alkali metal vapor in the near-infrared (NIR) spectral range have been extensively studied for a long time [1–16]. However, no commonly accepted mechanism has been proposed that could unambiguously explain the significant absorption experimentally observed [1–10, 12–16] outside the region of allowed electron-vibrational-rotational bands of alkali metal dimers (such dimers are present in a heated metal vapor in a sufficiently large proportion, reaching up to several percent of the atomic density component). In this part of the IR spectral range, the alkali metal vapor is usually considered as transparent. Indeed, optical absorption due to the allowed IR vibrational transitions in homonuclear dimers is absent because of the symmetry of these molecules. It is obvious that, in the absence of excitation of a vapor, the absorption cannot be related to the transitions from any excited electron states of, for example, a sodium atom or dimer, because these levels are characterized by very small thermal populations.

A significant optical absorption of an alkali metal vapor in the IR region outside the well-known singlet band $A^1\Sigma_u^+ \longleftarrow X^1\Sigma_g^+$ was reported for the first time by Chertoprud [1], who measured the transmission of a cell containing potassium vapor in a wavelength interval from 1 to 2.5 μm. The observed phenomenon was explained assuming excitation of the intercombination transition $a^3\Sigma_u^+ \longleftarrow X^1\Sigma_g^+$ in K$_2$ dimers. Later, the absorption spectra of potassium, cesium, and sodium in the interval from 0.8 to 2.5 μm were thoroughly studied by Bhaskar *et al.* [2–4] using optical cells of the “heat pipe” type [17, 18] in the presence of helium as a buffer gas. The whole body of new experimental data [2–4] represented evidence against the hypothesis of intercombination transitions formulated in [1]. As alternative mechanisms, Bhaskar *et al.* [2–4] suggested absorption due to the free-bound transition $b^3\Sigma_g^+ \longleftarrow a^3\Sigma_u^+$ and the absorption in trimers and more complicated alkali metal polymers. However, attempts to check for the latter possibility by determining the activation energy spectrum gave rather contradictory results [2–4]. The possibility of absorption by microdroplets of a fog which can, in principle, form in the aforementioned cells as a result of condensation of the axial vapor flows from the central (heated) zone to the peripheral (cooled) parts of the cell was rejected: the authors of [2–4] believed that the presence of helium excluded the formation of a fog (which was nevertheless observed when helium was replaced by a heavier buffer gas such as neon or argon).

Subsequent investigations [5–11], using more perfect methods for the calculation of potential curves of alkali metal dimers, solved the problem of determining the position of a longwave boundary of the band of the free-bound transition $b^3\Sigma_g^+ \longleftarrow a^3\Sigma_u^+$. In sodium vapor, the quasiclassical [5, 7, 9] and quantum-mechanical [11] calculations of the absorption coefficient taking into account all the possible low-lying transitions in

Na_2 dimers showed good coincidence with experimental data [5, 7–9] obtained in a broad spectral range including both the UV and IR regions. These investigations reliably established that the longwave boundary of the triplet band due to the transition $b^3\Sigma_g^+ \longleftarrow a^3\Sigma_u^+$ for the sodium dimer occurs at $\lambda \approx 0.92$ μm (Fig. 1). Therefore, absorption in the range of $\lambda > 0.92$ μm cannot be related to this transition. It was suggested [7, 8] that absorption observed in the IR region outside the tripled band is related, first, to trimers and other polymers of sodium (as in [2–4]) and, second, to the presence of a microdroplet condensate. The latter possibility was confirmed, at first glance, by the results of high-precision measurements [9] performed in a specially designed cell ensuring homogeneous heating of a vapor in the absence of a buffer gas: no absorption in the region of 0.92–1.075 μm was observed to within the experimental error (on the order of 10^{-3} cm^{-1}) under the conditions of a strongly overheated vapor excluding condensate formation. However, strong overheating significantly reduced the partial pressure of a molecular component in the vapor and, accordingly, decreased the IR absorption coefficient related to this component. For a light atom such as sodium, the absorption due to the intercombination transition $a^3\Sigma_u^+ \longleftarrow X^1\Sigma_g^+$ in the first approximation is rather small [4, 7, 20].

A significant thermal emission and absorption in the IR range was also observed in our investigations of sodium vapor [12–16] performed both in cells of the "heat pipe" type and in a cell with homogeneous heating [16]. It was demonstrated [15] that the formation of a condensate, even when helium used as a buffer gas, can significantly influence the spectral characteristics of sodium vapor in the IR range, but this effect cannot explain the whole body of experimental data. It should be noted that we have also put forward a hypothesis [13, 15] concerning the possible influence of many-particle effects on the shape of far wings of the atomic and molecular transitions corresponding to large detunings from resonance, $\Delta\omega \sim \omega_0$, where $\Delta\omega = \omega - \omega_0$, ω is the current frequency, and ω_0 is the resonance transition frequency. This problem remains insufficiently studied, although some aspects have been touched upon, for example, in [21–23]. Such a mechanism can play a significant role at large pressures of both sodium vapor and a buffer gas. A model example of the distribution of ion microfields in a dense plasma [13, 15] showed that many-particle effects may, in principle, account for a slower decay of the line contour as compared to that according to the nearest-neighbor approximation, thus significantly increasing the absorption in a dense medium at large detunings from the resonance.

It should be also noted that intense emission in the NIR range has been known for a long time in the well-studied spectra of high-pressure gas-discharge sodium lamps [24, 25]. These IR losses, which can account for up to 20% of the total energy balance, have been called "mysterious" [24] and still have not receive even qualitative explanation within the framework of any existing theoretical models.

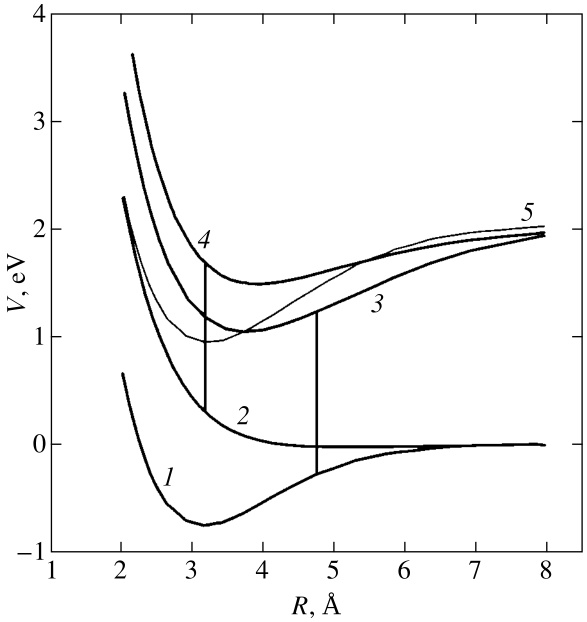

Fig. 1. Potential curves of the lower levels of a Na_2 dimer: (*1*) $X^1\Sigma_g^+$; (*2*) $a^3\Sigma_u^+$; (*3*) $A^1\Sigma_u^+$; (*4*) $b^3\Sigma_g^+$; (*5*) $a^3\Pi_u$. Vertical lines show the positions of classical satellites of the *1–3* and *2–4* transitions.

Thus, the problem of explaining significant optical absorption and intense emission in the IR spectral range for a dense alkali metal vapor remains open. On the other hand, there is an urgent need for refining the mechanisms of absorption in these systems. This is required for testing various models of the interaction potential in dimers and trimers of an alkali metal, which are employed in the descriptions of various phenomena related to ultracold collisions of alkali metal atoms, diagnostics of high-pressure gas-discharge plasma, etc.

In this context, we have performed for the first time a series of experiments for determining the absorption coefficient of a dense sodium vapor in a spectral interval from 0.8 to 2.6 μm in a cell with "hot" windows excluding the process of vapor condensation.

2. EXPERIMENTAL SETUP

The optical absorption spectra of a dense metal vapor were measured using the setup schematically depicted in Fig. 2. The main part of this setup is a sealed-off high-temperature leucosapphire cell containing sodium, which was analogous to a cell described previously [26]. In cells of this and analogous types (see, e.g., [27, 28]), the pressure of a saturated metal vapor is determined by the coldest point. In the rest of the cell volume, the temperature is higher and, hence, the vapor is overheated. This prevents condensation and

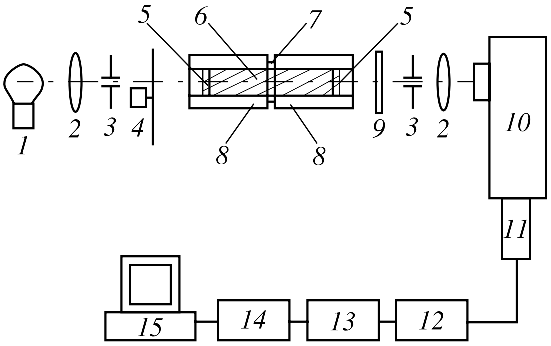

Fig. 2. A schematic diagram of the experimental setup: (*1*) halogen lamp; (*2*) objective lenses; (*3*) diaphragms; (*4*) modulator; (*5*) sapphire windows; (*6*) leucosapphire cell; (*7*) heat exchanger; (*8*) heaters; (*9*) optical filter; (*10*) MDR-23 monochromator; (*11*) detector; (*12*) lock-in preamplifier; (*13*) amplifier; (*14*) analog-to-digital converter; (*15*) computer.

the formation of a microdroplet component capable of distorting the results of absorption measurements.

This situation is essentially different from that in cells of the “heat pipe” type, where the maximum saturated vapor pressure is determined by the most heated (usually, central) region of the cell, while cold points occur at the cooled windows. Taking into account a considerable temperature gradient (and the related saturated vapor pressure gradient) in the axial direction, a stationary vapor pressure is established as a result of the vapor transfer by diffusion and convective flows from a hot region to the cooled ends of the cell. The vapor pressure at the periphery is significantly greater than the saturated vapor pressure corresponding to the temperature in this region, which gives rise to intensive condensation with the formation of sufficiently large microparticles [15].

The cell in our setup was a leucosapphire tube with a length of L = 200 mm and an internal diameter of 15 mm. Sapphire windows glued (with a high-temperature glue) to the cell edges were transparent up to a wavelength of 6 μm. We employed no buffer gas, and the cell was prepared by evacuation to a residual pressure not exceeding 10^{-4} Torr. In order to avoid deterioration of the vacuum level at elevated temperatures, the cell was provided with a titanium getter. The cell was sealed off and placed into a furnace composed of two parts (see Fig. 2). Each part comprised a quartz tube with a nichrome heating coil. The tube and the coil were coated with a thermoinsulating layer of chamotte and glass composite ribbon. The temperature of the windows and at the center of the cell was measured by chromel–alumel thermocouples. The maximum temperature of the windows (1170 K) was determined by the thermal stability of the glue. The temperature at the center of the cell (cold point) was maintained at a level about 100 K below that of the windows. This was provided by a chamotte ring in the gap between the two parts of the furnace, which acted as a heat exchanger.

The optical absorption coefficient of a dense sodium vapor was measured using a collimated beam of radiation from a halogen lamp possessing a sufficiently smooth continuous emission spectrum in a wavelength interval from 0.4 to 3 μm. The intensity of radiation transmitted through the sodium vapor in the axial region of the cell was measured as follows. The incident radiation was modulated by a chopper operating at a frequency of about 800 Hz. The modulated beam passing through the absorbing vapor column in the cell was focused at the entrance slit of an MDR-23 monochromator with a reciprocal linear dispersion of 2.6 nm/mm (for a 600 lines/mm grating used in the region of wavelengths below 1.8 μm) and 5.2 nm/mm (300 lines/mm grating for the wavelengths above 1.8 μm). The slit width in most cases was set at 2.2 mm. The wavelength range studied was scanned by rotating the diffraction grating of the monochromator controlled by a mechanical system driven with an electric step motor. The output radiation intensity was measured by detectors placed in the focal plane of the monochromator: a germanium diode of the FD-3 type operating in a photovoltaic regime and a photoresistor of the FPU 1235 GA4UBR type used in the range of wavelengths above 1.8 μm. The signal from the detector output was fed to a narrowband lock-in preamplifier tuned to the modulation frequency, amplified, digitized by an analog-to-digital converter, and processed by a personal computer.

In order to determine the optical absorption in a heated vapor, we measured the intensity I_0 of the radiation transmitted through the cell at room temperature (i.e., in the absence of a metal vapor) and the intensity I transmitted through the same cell containing sodium vapor heated to various temperatures: $I = I_r - I_v$, where I_r is the recorded signal intensity and I_v is the intrinsic emission intensity (in all experiments, the intrinsic emission from the vapor was negligibly small as compared to I_r). The experimental absorption coefficient was calculated by the formula

$$K_{\exp} = \frac{1}{L}\ln\frac{I_0}{I}. \tag{1}$$

The I_0 values measured before and after heating of the cell were reproduced to within 1%, which corresponds to an error in determining the absorption coefficient of below 5×10^{-4} cm^{-1}.

3. EXPERIMENTAL IR ABSORPTION SPECTRA OF A DENSE SODIUM VAPOR AND DISCUSSION OF RESULTS

Figure 3 shows the typical absorption spectra of sodium vapor measured in a wavelength interval from 0.8 to 2.6 μm and, Fig. 4 presents more detailed spectra recorded in a wavelength interval from 0.8 to 1.4 μm for various temperatures at the cold point of the cell. According to these data, sodium vapor exhibits significant absorption in the IR range, the absorption band structure being generally analogous to that observed in cells of the heat pipe type [4]. The spectra in Figs. 3 and 4 also exhibit a sharp absorption boundary at a wavelength of 0.82 μm, which is related to the longwave edge of the singlet band $A^1\Sigma_u^+ \longleftarrow X^1\Sigma_g^+$. The spectra reveal clearly pronounced absorption peaks (indicated by vertical lines in Fig. 4) at $\lambda \approx 0.89$ and 0.85 μm, which are attributed to a longwave satellite of the triplet band $b^3\Sigma_g^+ \longleftarrow a^3\Sigma_u^+$ [7–9, 25] and the related secondary maximum [9, 25]. The relative intensity of these peaks decreases with increasing temperature.

However, instead of a sharp drop in absorption at the longwave edge of the triplet band in the region of $\lambda \sim$ 0.9 μm observed in [9], our spectra exhibit a plateau extending up to about 1.02 μm. Note that, according to the results of quantum-mechanical calculations, the coefficient of absorption in the triplet at 0.92 μm and a temperature of T = 1000 K is on the order of $0.3 \times 10^{-39} N_0^2$ cm^5 (where N_0 is the number density of sodium atoms) [11]. Under our experimental conditions, N_0 is determined by the saturated vapor pressure at a temperature of the cold point of the cell and, for the temperature indicated above, $N_0 \approx 1.32 \times 10^{18}$ cm^{-3} [29]. Then, the calculated absorption coefficient is 0.5×10^{-3}, which is much smaller than the experimentally measured value amounting to about 0.15 cm^{-1} at nearly the same temperature of T = 1003 K (Fig. 4). Thus, the results of our measurements at first glance appear to contradict the experimental data reported in [9]. As the wavelength further increases, the absorption initially decreases and then increases again, thus forming a broad band with a maximum at about 1.6 μm (Fig. 3).

Under the experimental conditions studied, that is, without condensation (and the related formation of microdroplets) and in the absence of a buffer gas, the absorption observed in the 0.9–2.6 μm range can only be related to the presence of sodium dimers and trimers. In a strongly overheated vapor under quasiequilibrium conditions in the cell employed in our experiments, the role of higher polymers is insignificant because of their low abundance [30].

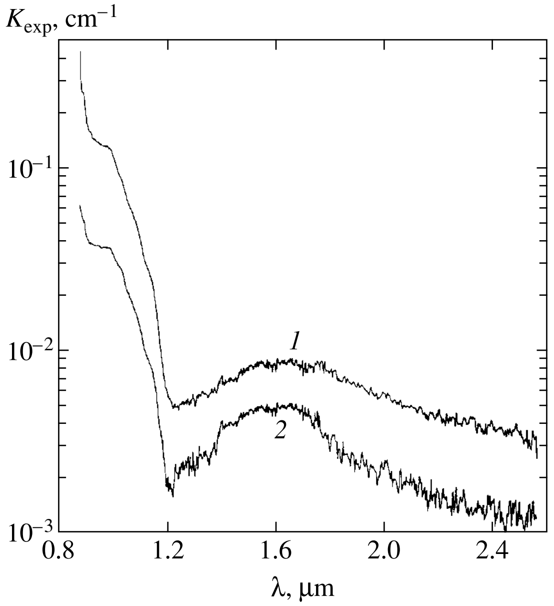

Fig. 3. Absorption spectra of sodium vapor in the cell with a cold point temperature of T = 1003 (*1*) and 954 K (*2*).

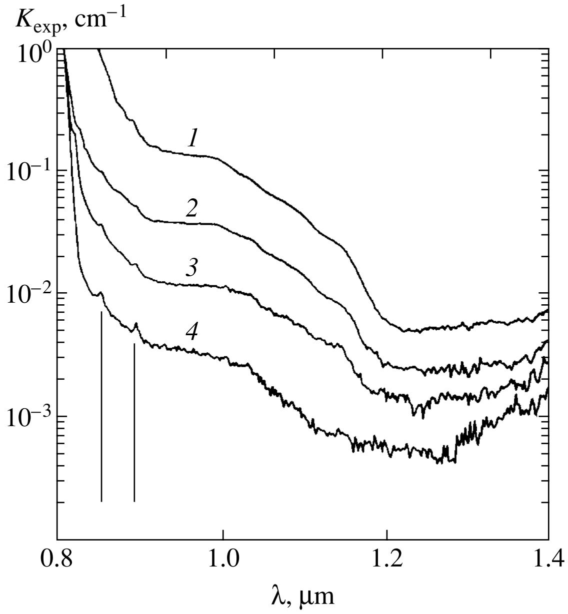

Fig. 4. Absorption spectra of sodium vapor in a wavelength interval from 0.8 to 1.4 μm, measured in the cell with a cold point temperature of T = 1003 (*1*), 954 (*2*), 905 (*3*), and 856 K (*4*). Vertical lines indicate the positions of absorption peaks corresponding to a longwave satellite of the triplet band $b^3\Sigma_g^+ \longleftarrow a^3\Sigma_u^+$ and the related secondary maximum.

First, let us consider the possible role of Na$_2$ dimers in the IR absorption. As pointed out above, it has been reliably established that sodium dimers are characterized by a longwave boundary of the triplet band $b^3\Sigma_g^+ \longleftarrow a^3\Sigma_u^+$ at $\lambda \approx 0.92$ μm and, hence, absorption in the range of $\lambda > 0.92$ μm cannot be related to this transition. Another possibility, also mentioned above, is the absorption due to the intercombination transition $a^3\Sigma_u^+ \longleftarrow X^1\Sigma_g^+$. To the best of our knowledge, no papers have been reported so far where the dipole moment of this transition would be calculated or at least estimated.

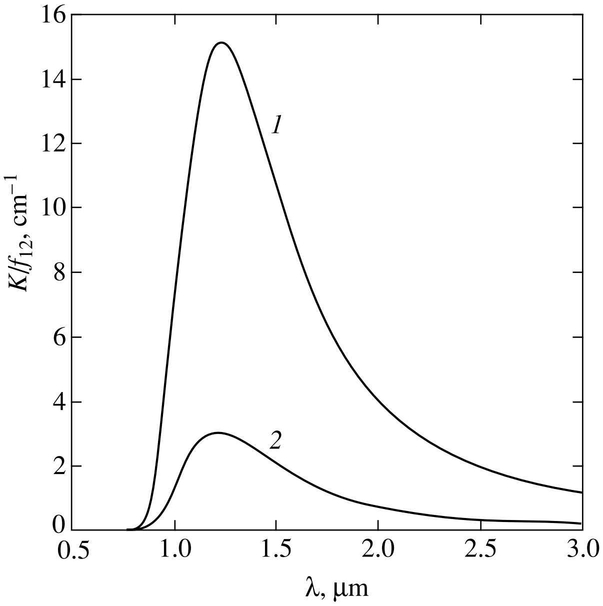

Fig. 5. Plots of the absorption coefficient $K_f(\lambda) = K(\lambda)/f_{12}$ (normalized to the oscillator strength) versus wavelength for the intercombination transition $a^3\Sigma_u^+ \longleftarrow X^1\Sigma_g^+$ calculated for two values of the cold point temperature T = 1003 (*1*) and 905 K (*2*).

However, for a light atom such as sodium, this quantity is expected to be very small.

Assuming that the dipole moment and the related oscillator strength f_{12} of this transition weakly depend on the internuclear distance (which is generally a quite satisfactory approximation for the allowed transitions in sodium dimers), we can estimate the wavelength dependence of the absorption coefficient $K(\lambda)$ for the intercombination transition using the well-known formula obtained in the quasistatic approximation [21, 31],

$$K(\lambda) = \frac{\pi^2 e^2}{2mc} f_{12} g_1 N_0^2 R^2 \left|\frac{dR}{d\nu_{12}}\right| \exp\left[-\frac{V_1(R)}{kT}\right]. \qquad (2)$$

Here,

$$\nu_{12} = \nu_{12}(R) = c/\lambda$$

is the frequency of the transition between ground state 1 (singlet term $X^1\Sigma_g^+$) and state 2 (triplet term $a^3\Sigma_u^+$); R is the internuclear distance; e and m are the electron charge and mass, respectively; c is the speed of light in vacuum; $g_1 = 1$ is the statistical weight of the ground state; and $V_1(R)$ is the potential energy of the ground state, such that

$$V_2(R) - V_1(R) = h\nu_{12},$$

where $V_2(R)$ is the energy of the triplet state and h is the Planck constant.

According to Eq. (2), the wavelength dependence of $K(\lambda)$ is determined by the exponential term $\exp(-V_1(R)/kT)$ and reaches maximum at the minimum of the potential $V_1(R)$. Figure 5 shows the plots of $K_f(\lambda) = K(\lambda)/f_{12}$ calculated for two values of the cold point temperature (T = 905 and 1003 K) using formula (2) and the known potential curves of sodium dimers [19]. In the calculation, it was assumed that the density of the atomic vapor component corresponds to the saturated vapor pressure (calculated using the data from [29]) at the given cold point temperature, while the absorption coefficient was determined at the window temperature (for the vapor occurring in the overheated state in most part of the cell). As can be seen from the data in Fig. 5, the calculated absorption spectra of the intercombination transition (with the oscillator strength independent of the distance R) are qualitatively different from those measured in the experiment and exhibit maxima in the region of 1.25 μm, where the experimental curves are at a minimum (cf. Figs. 3, 4, and 5).

It should be noted that the dipole moment of the intercombination transition can strongly depend on the internuclear distance [21], which, in principle, will lead to the calculated absorption spectrum with a shape close to that measured in experiment. Nevertheless, assuming that the absorption coefficient at a given wavelength is determined by the intercombination transition, we can estimate the "effective" oscillator strength of the transition for this wavelength using formula (2) and evaluate the temperature dependence of this value:

$$f_{12\text{eff}}(\lambda, T) = \frac{K_{\text{exp}}(\lambda, T)}{K_f(\lambda, T)}.$$

A significant variation of the $f_{12\text{eff}}$ value with the temperature would exclude the mechanism of absorption related to the intercombination transition, because the real oscillator strength of a molecular transition has to be independent of the temperature.

According to the results of our calculations, $f_{12\text{eff}}$ at a wavelength of 1.25 μm equals, on the average, $f_{12\text{eff}} \approx (3.6 \pm 0.3) \times 10^{-4}$ and is virtually (to within the measurement error) independent of the temperature in the 850–1000 K interval. This result does not contradict the hypothesis concerning the role of the intercombination transition. Note that the above value of the "effective" oscillator strength refers to the minimum of K_{exp} and the maximum of the calculated spectrum of K. In other regions of the spectrum, calculations give significantly greater values of $f_{12\text{eff}}$, for example, $f_{12\text{eff}} \sim 10^{-2}$ at $\lambda = 1.0$ μm. It is hardly probable for an intercombination transition to possess such a large value of the oscillator strength.

The absorption related to the other forbidden transitions, $a^3\Pi_u \longleftarrow a^3\Sigma_u^+$ and $A^1\Sigma_u^+ \longleftarrow a^3\Sigma_u^+$, can be ignored because of the very small thermal population of the triplet term $a^3\Sigma_u^+$. Analogous calculations of the absorption coefficients for these transitions give values comparable to the experimental data, only assuming that the oscillator strengths are as large as $f \sim 1$, which is unreal. The region of possible absorption for another forbidden transition, $a^3\Pi_u \longleftarrow X^1\Sigma_g^+$, falls within the visible spectral range.

Besides the direct optical absorption due to forbidden transitions, there is an alternative possibility of absorption related to the collision-induced processes [32] accompanying the interaction between two sodium dimers or between a dimer and an atom, the probabilities of which can be significant provided that the metal vapor pressure is sufficiently high. Such processes were studied in sufficient detail for the absorption related to the vibrational–rotational transitions in some molecules [32] and the forbidden transitions in atoms (see, e.g., [21, 33, 34]). In the latter case, the values of reduced absorption coefficients

$$K_r(\lambda) = \frac{K(\lambda)}{N_a N_p}$$

(where N_a and N_p are the number densities of the absorbing and perturbing particles) can be rather large, on the order of 10^{-38} cm^5 [33, 34]. Some investigations were devoted to the collision-induced absorption involving forbidden molecular electron transitions, predominantly in the molecule of oxygen [35–37], which is related to the role of this phenomenon in the formation of the absorption spectrum of the Earth's atmosphere. The values of the absorption coefficients of O_2 molecules for the triplet–singlet transitions between ground state terms in the IR range can be on the order of 10^{-45} cm^5 [35, 36].

To our knowledge, no investigations of the collision-induced absorption related to the intercombination transition in the sodium dimer were reported in the literature (the calculation of such processes represents a very cumbersome quantum-mechanical problem). In particular, the assumption that the dipole moment of the collision-induced transition is independent of the distance R_c between the interacting particles is by no means valid in this case, since this quantity tends to zero for $R_c \longrightarrow \infty$. However, let us assume that the collision-induced absorption in sodium dimers (with allowance for the difference in the polarizabilities of molecules) is on the same order of magnitude as that in molecular oxygen. For the sodium dimer, the polarizability amounts to 39.2 Å^3 [38], while that of the O_2 molecule is 1.571 Å^3 [39]. Taking into account that the dipole moment of a transition is proportional to the square of polarizability [40, 41], we may expect that the reduced absorption coefficient of sodium dimers can be about 600 times that of oxygen and reach a level on the order of 10^{-42} cm^5. At $T \sim 1000$ K, the number density of sodium dimers in a saturated vapor is $N_d \sim 10^{17}$ cm^{-3}. This yields an absorption coefficient on the level of 10^{-8} cm^{-1} for the Na_2–Na_2 collisions, and on the level of 10^{-7} cm^{-1} for the Na_2–Na collisions. As can be readily seen (Fig. 3), these estimates are three to four orders of magnitude below the values observed in experiment. Note also that, if we assume that a broad peak in the region of 1.6 μm is related to the collision-induced absorption, the peak width must be on the order of the reciprocal time of flight of a perturbing particle moving at a thermal velocity near an absorbing particle. Taking a radius of intermolecular forces on the order of 10^{-7} cm, we estimate the width at about 0.1 μm, which is several times smaller than the measured value (about 0.4 μm in Fig. 3).

One of the main features of the collision-induced absorption is a quadratic dependence of the absorption coefficient on the density (this corresponds to the pair collisions) and a weak dependence on the temperature [32]. The latter dependence consists mostly in a slight increase of the absorption band width with the temperature, which is related to a decrease in the collision time caused by increasing relative velocity of the interacting particles:

$$v \propto T^{1/2}.$$

For the conditions of our experiments, this factor can be ignored in the first approximation. Figure 6 shows the temperature variation of the reduced absorption coefficients $K_{\exp}/N_0N_d$ and $K_{\exp}/N_d^2$ at a wavelength of 1.25 μm. The number densities of sodium dimers were calculated using data on the saturated sodium vapor pressure reported in [29] and the spectroscopic constants of Na_2 molecules tabulated in [42]. As can be seen from Fig. 6, the reduced absorption coefficients in both cases significantly vary with the temperature. This contradicts the hypothesis about on collision-induced character of absorption of sodium vapor in the spectral range under consideration.

Now let us consider the possible role of Na_3 trimers in IR absorption. The absorption spectra of trimers were studied both experimentally and theoretically in sufficient detail (see, e.g. reviews [43–45] and references therein). However, the experimental investigations of the optical absorption of small clusters of various metals were restricted to the visible spectral range, whereas data on the IR absorption are not available. This is related to the fact that such spectra cannot be directly measured by transillumination of the medium, since the cluster beams presently used for the spectro-

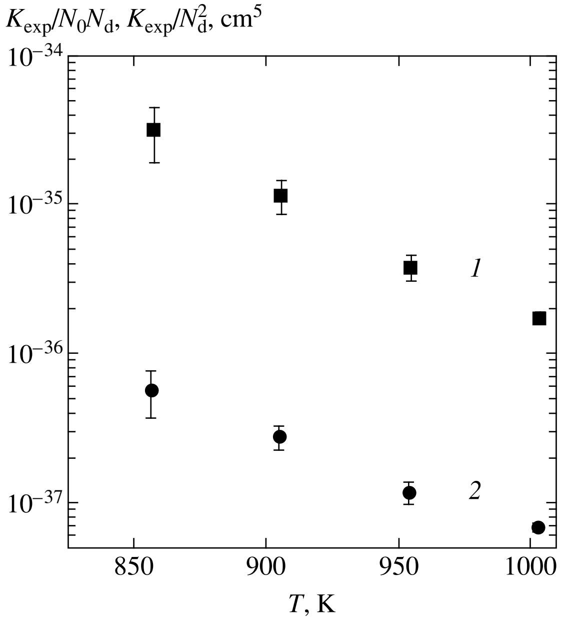

Fig. 6. Temperature variation of the reduced absorption coefficients K_{exp}/N_0N_d (*1*) and K_{exp}/N_d^2 (*2*) at a wavelength of 1.25 μm.

scopic investigations have a very low density. For this reason, the absorption spectra of clusters (including sodium trimers) are obtained by various indirect methods based on the use of dye lasers. Unfortunately, the tuning range of such lasers is restricted to the region of wavelength below 0.8–0.9 μm.

In our opinion, the most thorough experimental investigation of the optical absorption of sodium trimers in a wavelength interval from 0.35 to 0.85 μm was reported in [46], where the spectra reveal seven clearly pronounced absorption bands. Simultaneously published results of theoretical calculations [47] of the energies and oscillator strengths of the electron transitions in Na_3 molecule showed good coincidence with the experimental data [46]. The positions of measured absorption peaks generally correspond (with a slight shortwave shift by 0.04–0.1 eV) to the calculated excitation energies, while the peak amplitudes correlate well with the calculated oscillator strengths of the transitions.

In addition to the results for the transitions in the visible spectral range, the excitation energies were calculated [47] for the three lowest levels of sodium trimer with the wavelengths of vertical transitions from the ground state falling within the IR range: $\lambda_1 = 2.383$ μm, $\lambda_2 = 1.158$ μm, and $\lambda_3 = 0.932$ μm, the oscillator strengths f_{12} of these transitions being 0.0119, 0.0092, and 0.0115, respectively. In principle, wavelengths λ_2 and λ_3 may correspond to the absorption observed in our experiments at 0.9–1.2 μm (see Fig. 4), whereas the broad absorption peak observed in the region of 1.6 μm (Fig. 3) is too strongly shifted from the value of λ_1 toward shorter wavelengths. It should be noted that the experimentally measured widths ΔE of the absorption peaks observed in the visible range fall within 0.07–0.17 eV [46] and we may suggest that the width of IR transitions should be on the same order of magnitude.

Using the above data, it is possible to calculate the absorption cross sections at maxima of the IR absorption peaks for sodium trimers, proceeding from the Lorentzian line shape [48]:

$$\sigma = \frac{4\pi e^2}{mc} f \frac{1}{\Delta\omega_{12}}, \tag{3}$$

where $\Delta\omega = \Delta E/\hbar$. Assuming for estimation purposes that $\Delta E \approx 0.1$ eV and using formula (3), we obtain the cross sections $\sigma = 8.3 \times 10^{-18}$, 6.4×10^{-18}, and 8.0×10^{-18} cm^2 for λ_1, λ_2, and λ_3, respectively.

In order to evaluate the absorption coefficients corresponding to the above peaks, we have calculated the equilibrium densities of the trimer vapor component N_{tr} using recent theoretical data on the fundamental vibrational frequencies ($\omega_1 = 295.3$ cm^{-1}, $\omega_2 = 56.9$ cm^{-1}, and $\omega_3 = 102.9$ cm^{-1}) [49] and the dissociation energies (about 3000 cm^{-1}) calculated in [50–52]. The moments of inertia of the trimer molecule were calculated using the structural data from [43]. The calculations were performed assuming (as it was done previously) that the density of the atomic vapor component corresponds to the saturated vapor pressure at the given cold point temperature, while the trimer number density refers to the window temperature (since the vapor occurs in the overheated state in the greater part of the cell).

The results of calculating the absorption coefficients K_{calcd} for the IR absorption peaks of sodium trimers at various temperatures are presented in the table. As can be seen from these data, the calculated absorption coefficients at all wavelengths and temperatures are significantly smaller than the experimentally measured ones. The differences, while being relatively small (to within a factor of 2–5) for λ_1, increase to one order of magnitude for λ_2 and reach up to two orders of magnitude for λ_3. It should be noted that subsequent refinement of both experimental and calculated data on the parameters of sodium trimers may lead to some change in the calculated number density of sodium trimers in the vapor and, hence, to the possible approach of K_{calcd} to K_{exp}.

Still another possible factor contributing to the optical absorption of sodium vapor in the IR range is the absorption related to many-particle effects, that is, to the interaction of an emitting atom or molecule with a large number of perturbing particles. As was demonstrated (see, e.g., [21, 53]), such interactions may lead to the appearance of additional satellites in the colli-

Calculated absorption coefficients K_{calc} for the IR absorption peaks of sodium trimers at various temperatures

T, K	N_{tr}, cm^{-3}	$\lambda_1 = 2.383$ μm $\sigma = 8.3 \times 10^{-18}$ cm^2		$\lambda_2 = 1.158$ μm $\sigma = 6.4 \times 10^{-18}$ cm^2		$\lambda_3 = 0.932$ μm $\sigma = 8.0 \times 10^{-18}$ cm^2	
		K_{calcd}, cm^{-1}	K_{exp}, cm^{-1}	K_{calcd}, cm^{-1}	K_{exp}, cm^{-1}	K_{calcd}, cm^{-1}	K_{exp}, cm^{-1}
1003	2.65×10^{14}	2.2×10^{-3}	3.9×10^{-3}	1.7×10^{-3}	1.7×10^{-2}	2.1×10^{-3}	1.4×10^{-1}
954	8.86×10^{13}	7.4×10^{-4}	1.5×10^{-3}	5.7×10^{-4}	6.1×10^{-3}	7.1×10^{-4}	3.8×10^{-2}
905	2.60×10^{13}	2.2×10^{-4}	–	1.7×10^{-4}	2.7×10^{-3}	2.1×10^{-4}	1.2×10^{-3}
856	6.60×10^{12}	5.5×10^{-5}	–	4.2×10^{-5}	5.6×10^{-4}	5.3×10^{-5}	3.5×10^{-3}

sion-broadened absorption spectra. The distance from each subsequent satellite (satellites-on-satellites [21, p. 1140]) to the main spectral line is approximately a multiple of the distance from the first satellite. The many-particle effects were experimentally observed, for example, in the spectra of cesium resonance absorption lines broadened by inert buffer gases [53] and in some other mixtures at relatively small densities of the perturbing gas (about 5×10^{18}–5×10^{19} cm^{-3}).

Under the experimental conditions studied here, such additional satellites of the sodium 3S–3P resonance line may, in principle, appear as a result of collisions involving three atoms, which can be interpreted as the collisions between sodium atoms and dimers. In this case, we may expect that the additional satellites will represent the satellites-on-satellites of the singlet–singlet ($A^1\Sigma_u^+ \longleftarrow X^1\Sigma_g^+$, $\lambda = 0.804$ μm [5, 9, 11]) and triplet–triplet ($b^3\Sigma_g^+ \longleftarrow a^3\Sigma_u^+$, $\lambda = 0.89$ μm, see Fig. 4) molecular transitions in sodium dimers. Then, the satellites must correspond to wavelengths of 1.1 and 1.82 μm and, in principle, can be identified with the experimentally observed broad absorption bands at 0.9–1.2 and 1.6–1.8 μm (see Figs. 3 and 4). However, neither correct justification of this possibility nor evaluation of the corresponding coefficients are possible without detailed calculations.

4. CONCLUSIONS

We have measured the optical absorption coefficients in the IR spectral interval from 0.8 to 2.6 μm for a dense sodium vapor ($N_0 \approx 10^{17}$–10^{18} cm^{-3}) in a homogeneously heated sealed-off cell. The experimental data show evidence of significant absorption in the range of parameters studied. An analysis of the possible absorption mechanisms (absorption due to a forbidden intercombination transition, collision-induced processes, trimer vapor component, and many-particle effects) and estimates of the corresponding contributions showed that neither the spectral features observed nor the values of the absorption coefficients measured can be explained based on these mechanisms and the available published data. Many-particle effect can play an important role; they can, in principle, explain the experimental results. Unfortunately, no numerical estimates are available for these effects. For a reliable interpretation of the experimental data and justified choice of one or several factors contributing to the observed absorption in the spectral range studied, it necessary to perform a detailed quantum-mechanical calculation of the aforementioned mechanisms in application to the particular cases of sodium and other alkali metals.

As was indicated in the Introduction, no absorption (to within the experimental error) was observed previously [9] in the wavelength interval from 0.92 to 1.075 μm for sodium vapor in a homogeneously heated cell, which is at variance with our data. Note, however, that the experiments described in [9] were performed under the conditions of strong overheating, whereby a temperature difference between the reservoir with sodium vapor and the sapphire cell reached about 500 K. Under these conditions, the number densities of dimer and trimer vapor components were significantly lower as compared to those in our experiments. Indeed, the number densities of sodium dimers and trimers calculated for the conditions studied in [9] did not exceed 3×10^{15} and 2×10^{12} cm^{-3}, respectively. Under the conditions of our experiments, the corresponding values even at a minimum temperature were 2×10^{17} and 6.6×10^{12} cm^{-3}, respectively. This is evidence, in particular, for the fact that the observed absorption is due to the molecular vapor component.

It should be also pointed out that the pattern of optical absorption observed in the IR spectral range approximately corresponds to a considerable emission from sodium vapor previously reported in the spectral interval from 2 to 3 μm [12, 14]. Indeed, using the Kirchhoff law [49], it is possible to determine the spectral intensity of emission from sodium vapor under the conditions of our experiments,

$$\varepsilon(\lambda) = F(\lambda, T)K(\lambda), \tag{4}$$

where $F(\lambda, T)$ is the Planck function. Then, the total emission intensity is given by the integral from $\varepsilon(\lambda)$

over the wavelength interval studied. Upon extrapolating the measured wavelength dependence of the absorption coefficient to $\lambda = 3$ μm at $T = 1003$ K, we obtain the following value of the intensity of emission in the 2–3 μm interval: 1.7×10^4 erg/(s cm^2 sr). The experimental value of the emission intensity reported for the same interval is 2.6×10^4 erg/(s cm^2 sr) [12, 14].

In conclusion, it should be noted that analogous behavior of the optical absorption coefficient in the IR spectral range can be expected for the vapor of other alkali metals.

ACKNOWLEDGMENTS

The authors are grateful to N.V. Znamenskiĭ, V.S. Lisitsa, and E. A. Petrenko for fruitful discussions of the experimental results and to D. I. Chekhov for his help in conducting the measurements.

This study was supported by the Russian Foundation for Basic Research, project nos. 00-15-96539 and 02-02-16758a.

REFERENCES

1. V. E. Chertoprud, Teplofiz. Vys. Temp. **14**, 216 (1976).
2. N. D. Bhaskar, E. Zouboulis, T. McClelland, and W. Happer, Phys. Rev. Lett. **42**, 640 (1979).
3. E. Zouboulis, N. D. Bhaskar, A. Vasilakis, and W. Happer, J. Chem. Phys. **72**, 2356 (1980).
4. A. Vasilakis, N. D. Bhaskar, and W. Happer, J. Chem. Phys. **73**, 1490 (1980).
5. J. P. Woerdman and J. J. De Groot, Chem. Phys. Lett. **80**, 220 (1981).
6. S. Shandin, B. Wellegehausen, and Z. G. Ma, Appl. Phys. B **29**, 195 (1982).
7. J. Huennekens, S. Schaefer, M. Ligare, and W. Happer, J. Chem. Phys. **80**, 4794 (1984).
8. M. Palle, S. Milosevic, D. Veza, and G. Pichler, Opt. Commun. **57**, 394 (1986).
9. J. Schlejen, C. J. Jalink, J. Korving, *et al.*, J. Phys. B **20**, 2691 (1987).
10. M. Ligare and J. B. Edmonds, J. Chem. Phys. **95**, 3857 (1991).
11. H.-K. Chung, K. Kirby, and J. F. Babb, Phys. Rev. A **63**, 032516 (2001).
12. A. G. Leonov, A. N. Starostin, D. I. Chekhov, *et al.*, in *Proceedings of 14th International Conference on Spectral Line Shapes*, Ed. by R. M. Herman (Woodbury, New York, 1998), Vol. 10, p. 400.
13. A. N. Starostin, I. I. Yakunin, A. G. Leonov, *et al.*, in *Proceedings of 15th International Conference on Spectral Line Shapes*, Ed. by J. Seidel (Melvill, New York, 2001), Vol. 11, p. 16.
14. Yu. K. Zemtsov, A. Yu. Sechin, A. N. Starostin, *et al.*, Zh. Éksp. Teor. Fiz. **114**, 135 (1998) [JETP **87**, 76 (1998)].
15. A. G. Leonov, A. A. Rudenko, A. N. Starostin, *et al.*, Zh. Éksp. Teor. Fiz. **122**, 282 (2002) [JETP **95**, 242 (2002)].
16. A. A. Rudenko, Candidate's Dissertation in Physics and Mathematics (Moscow Inst. of Physics and Technology, Moscow, 2002).
17. S. R. Vidal and J. Cooper, J. Appl. Phys. **40**, 3370 (1969).
18. W. Demtroder, *Laser Spectroscopy: Basic Concepts and Instrumentation* (Springer, New York, 1981; Nauka, Moscow, 1985).
19. D. D. Konowalow, M. E. Rosenkrantz, and M. L. Olson, J. Chem. Phys. **72**, 2612 (1980).
20. M. A. El'yashevich, *Atomic and Molecular Spectroscopy* (Nauka, Moscow, 1962).
21. N. Allard and J. Kielkopf, Rev. Mod. Phys. **54**, 1103 (1982).
22. D. A. Evensky and K. M. Sando, Phys. Rev. A **31**, 772 (1985).
23. J. Kielkopf and N. Allard, in *DAMOP 2000 Meeting Program* (Storrs, CT, 2000), D 9.45.
24. D. O. Wharmby, IEE Proc. **127**, 165 (1980).
25. J. J. de Groot and J. A. J. M. van Vliet, *The High-Pressure Sodium Lamp* (Macmillan, Basingstoke, 1986).
26. D. H. Sarkisyan and A. V. Papoyan, Appl. Opt. **35**, 3207 (1996).
27. J. Schlejen, J. Post, J. Korving, *et al.*, Rev. Sci. Instrum. **58**, 768 (1987).
28. M. Shurgalin, W. H. Parkinson, K. Yoshino, *et al.*, Meas. Sci. Technol. **11**, 730 (2000).
29. O. D. Kuznetsova and A. M. Semenov, Teplofiz. Vys. Temp. **38**, 30 (2000).
30. B. M. Smirnov, Usp. Fiz. Nauk **170**, 495 (2000) [Phys. Usp. **43**, 453 (2000)].
31. S. Y. Chen and M. Takeo, Rev. Mod. Phys. **29**, 20 (1957).
32. L. Frommhold, *Collisional-Induced Absorption in Gases* (Cambridge Univ. Press, Cambridge, 1993).
33. A. Gallagher and T. Holstein, Phys. Rev. A **16**, 2413 (1977).
34. E. Bichoutskaia, A. Devdariani, K. Ohmori, *et al.*, J. Phys. B **34**, 2301 (2001).
35. G. S. Tabisz, E. J. Allin, and H. L. Welsh, Can. J. Phys. **47**, 2859 (1969).
36. G. D. Greenblatt, J. J. Orlando, J. B. Burkholder, and A. R. Ravishankara, J. Geophys. Res. **95**, 18577 (1990).
37. L. Biennier, D. Romanini, A. Kachanov, *et al.*, J. Chem. Phys. **112**, 6309 (2000).
38. G. Tikhonov, V. Kasperovich, K. Wong, and V. V. Kresin, Phys. Rev. A **64**, 063202 (2001).
39. A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Atomizdat, Moscow, 1980; Springer, Berlin, 1985).
40. V. A. Kas'yanov and A. N. Starostin, Zh. Éksp. Teor. Fiz. **48**, 295 (1965) [Sov. Phys. JETP **21**, 193 (1965)].
41. V. A. Astapenko, L. A. Bureeva, and V. S. Lisitsa, Usp. Fiz. Nauk **172**, 155 (2002) [Phys. Usp. **45**, 149 (2002)].

42. *Physical Quantities. Handbook*, Ed. by I. S. Grigor'ev and E. Z. Meĭlikhov (Énergoatomizdat, Moscow, 1991).
43. V. Bonavic-Koutecky, P. Fantucci, and J. Koutecky, Chem. Rev. **91**, 1035 (1991).
44. W. A. de Heer, Rev. Mod. Phys. **65**, 611 (1993).
45. B. M. Smirnov, Usp. Fiz. Nauk **167**, 1169 (1997) [Phys. Usp. **40**, 1117 (1997)].
46. C. R. C. Wang, S. Pollack, D. Cameron, and M. M. Manfred, J. Chem. Phys. **93**, 3787 (1990).
47. V. Bonacic-Koutecky, P. Fantucci, and J. Koutecky, J. Chem. Phys. **93**, 3802 (1990).
48. Ya. B. Zel'dovich and Yu. P. Raĭzer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, 2nd ed. (Nauka, Moscow, 1966; Academic, New York, 1966 and 1967), Vols. 1 and 2.
49. B. Kendric, Phys. Rev. Lett. **79**, 2431 (1997).
50. R. L. Martin and E. R. Davidson, Mol. Phys. **35**, 1713 (1978).
51. J. L. Martins, R. Car, and J. Buttet, J. Chem. Phys. **78**, 5646 (1983).
52. T. C. Thompson, G. Izmiran, S. J. Lemon, *et al.*, J. Chem. Phys. **82**, 5597 (1985).
53. J. Kielkopf, in *Proceedings of 5th International Conference on Spectral Line Shapes*, Ed. by B. Wende (Walter de Gruyter, Berlin, 1981), Vol. 4, p. 665.

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 688–701.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 766–780.
Original Russian Text Copyright © 2003 by Makarov, Letokhov.

ATOMS, SPECTRA, RADIATION

Spontaneous Decay in a System of Two Spatially Separated Atoms (One-Dimensional Case)

A. A. Makarov* and V. S. Letokhov**
Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow oblast, 142190 Russia
**e-mail: amakarov@isan.troitsk.ru*
***e-mail: letokhov@isan.troitsk.ru*

Received April 10, 2003

Abstract—The problem of the dynamics and the spectrum of spontaneous radiation is solved for a system of two atoms in one-dimensional space. In order to single out, to the maximum possible degree, phenomena associated with the influence of spatially separated atoms on each other via the radiation field, the present analysis is performed precisely for the one-dimensional case. As a result, two effects are revealed and considered in detail: (i) the existence of stable (metastable) entangled superposition states at specific distances between the atoms and (ii) a considerable distinction between the spectra of photons emitted in two opposite directions from the system where only one of the atoms is initially excited. The possibilities of observing these effects are discussed. *© 2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

A quantum description of spontaneous radiation from an atom dates back to the classic study of Weisskopf and Wigner [1]. It includes three basic effects (see, for example, [2]):

(i) the exponential decay of an excited state prepared at the initial instant of time;

(ii) the shift of the observed central frequency in the spectrum of emitted photons with respect to the transition frequency obtained by solving the problem of energy levels of an isolated atom (radiative or Lamb shift);

(iii) the Lorentzian shape of the emitted-photon spectrum with respect to the central frequency.

Properties (i) and (iii) are tightly related (see, for example, [3]), but they are, strictly speaking, approximate. The degree of deviation from the exponential law and, hence, from a Lorentzian contour is controlled by the ratio of the spontaneous-decay rate Γ to a frequency $\tilde{\omega}$ that is a characteristic scale at which the density of electromagnetic-field modes changes in the vicinity of the transition frequency ω_0. For an atom in a free space, we have $\tilde{\omega} \approx \omega_0$, and the ratio Γ/ω_0 in the optical (visible) range is not more than 10^{-8} to 10^{-7} even for strong transitions.

From the aforesaid, it follows that means for governing the dynamics and the spectrum of spontaneous radiation must rely on how the density of electromagnetic-field modes in the vicinity of the transition frequency is structured. A number of specific examples were considered in Bykov and Shepelev's monograph [4], which anticipated in part subsequent studies devoted to radiation from an atom in a resonator [5], in a periodic structure (photon crystals [6, 7]), in the vicinity of a dielectric sphere with a size about that of the wavelength [8–10], and in some other cases. Influence at the microscopic level is also possible if, for example, one places, in the vicinity of an excited atom (within a distance about that of the wavelength λ), an atom having a close transition frequency. In the particular case (which, at the same time, is the most interesting) of identical atoms separated by a distance much shorter than λ, we have a system that exhibits the Dicke superradiance effect [11] or, more specifically, at least three basic cooperative effects. For the purpose of illustration, we restrict our consideration to the case of two atoms whose ground and excited states ($|g\rangle$ and $|e\rangle$, respectively) are coupled by the off-diagonal matrix element d_{eg} of the dipole-moment operator.

The first effect consists in what is referred to as superradiance—namely, if, at the initial instant of time, both atoms are in the excited state, in which case the system as a whole is in the $|ee\rangle$ state, the probability of finding a specific atom in the excited state decreases faster than in the case where there is no neighboring atom.

As a matter of fact, the second effect explains the mechanism of superradiance: in the dynamics of the process, the symmetric superposition $(|eg\rangle + |ge\rangle)/\sqrt{2}$ of two states, where one atom is excited and the other is not excited, is the intermediate state for the transition to the ground state $|gg\rangle$ of the entire system. The dipole moments of transitions between this superposition and both upper $|ee\rangle$ and lower $|gg\rangle$ states are greater than d_{eg} by a factor of $\sqrt{2}$.

Finally, the third effect consists in that the antisymmetric superposition $(|eg\rangle - |ge\rangle)/\sqrt{2}$ of the same states is stable since the dipole moment for the transition to the ground state $|gg\rangle$ (and to the excited state $|ee\rangle$ as well) is equal to zero. In the following, we will be interested primarily in this effect, which we will refer to as subradiance.

These cooperative effects, which are well known in the theory, have not yet been observed experimentally in a pure form—that is, for the case of two or a few atoms in a volume whose dimensions are much smaller than λ. Superradiance has been investigated predominantly in large ensembles of atoms or molecules in extended media (for an overview, see, for example, [12, 13]). Only quite recently did experiments with simple systems result in the observation of manifestations of both superradiance and subradiance. In the first of these experiments, which was reported in [14], two Ba^+ ions were captured in an ion trap and localized, by using the laser-cooling technique, at positions separated by a distance of about 2λ, which was slightly varied. A short laser pulse generated a superposition (antisymmetric in the case being discussed) of two states of the system, $|eg\rangle$ and $|ge\rangle$, whereupon the experiment involved measuring the time of spontaneous decay. In this way, a weak modulation was revealed in the dependence of this time on the distance between the ions (actually, the measurements were performed only for three distances).

In another experiment (see [15]), which also employed a Ba^+ ion in an ion trap, the mirror image of the localized ion played the role of the second identical particle. Spontaneous emission along the normal to the mirror was detected, and a sizable modulation of the signal with period λ/2 was observed in the dependence on the distance between the ion and the mirror. Naturally, an explanation in terms of interference suggests itself—this is true, in principle, but the authors of that study preferred to gain a deeper insight into the nature of the phenomenon, claiming that they were dealing with a manifestation of a cooperative effect. In order to validate their point of view, they performed an additional experiment. Two ions, 1 and 2, were captured into a trap, and observation was performed along the direction connecting ion 1 and the image of ion 2 (or vice versa). Again, a modulation of the signal was observed.

In the present study, the formulation of the problem is close to that in the experiment reported in [15]. Specifically,

(i) we consider two atoms that are spatially separated to such an extent that their dipole–dipole interaction is insignificant;

(ii) the atoms radiate only along the line passing through their centers (physically, such a one-dimensional situation can be realized in a waveguide whose transverse dimension is about λ).

In this one-dimensional model, the cooperative effect eventually manifests itself to the maximum possible degree not only as a modulation of the fluorescence signal in the experiment described in [15] but also in the dynamics of decay and in the radiation spectrum. The subradiance effect, which consists in that superposition (meta)stable excited states of the system exist at specific values of the interatomic spacing, is of particular interest.

2. FORMULATION OF THE ONE-DIMENSIONAL PROBLEM

We consider two identical two-level atoms, 1 and 2, located at an arbitrary distance from each other. For the definiteness, we place atom 1 at the origin of coordinates and denote the coordinate of atom 2 by $\mathbf{R}$. At the initial instant $t = 0$, atom 1 is in the excited state $|e\rangle$, while atom 2 is in the ground state $|g\rangle$. An electromagnetic field propagates only along vector $\mathbf{R}$; at the initial instant $t = 0$, the field is assumed to be in the vacuum state, $|0\rangle$. Thus, the initial state of the entire system is $|eg0\rangle$. As an initial basis, we take the eigenfunctions of the Hamiltonian in the zero-order approximation, where there is no interaction between atoms and the radiation field, which is quantized in an infinite one-dimensional space. We denote by A_1 the amplitude of the initial state of the "atom 1 + atom 2 + field" system. Further, we take into consideration continuum states where there is a single photon of arbitrary frequency ω and where both atoms are in the ground state. These continuum states are doubly degenerate in the direction of photon propagation. We denote them by $|gg1_\omega^{(+)}\rangle$ and $|gg1_\omega^{(-)}\rangle$ for the directions to the right and to the left, respectively; we also denote the corresponding amplitudes by $B_\omega^{(+)}$ and $B_\omega^{(-)}$. Finally, we include, in our consideration, the state $|eg0\rangle$, which interacts with the original state via continuum states, and denote its amplitude by A_2. The set of equations for the amplitudes A_1, A_2, $B_\omega^{(+)}$, and $B_\omega^{(-)}$ can be derived in a standard way from the time-dependent Schrödinger equation. The result is given by

$$
\begin{aligned}
i\frac{dA_1}{dt} &= \int_0^\infty V(\omega)B_\omega^{(+)}e^{-i(\omega-\omega_0)t}d\omega \\
&\quad + \int_0^\infty V(\omega)B_\omega^{(-)}e^{-i(\omega-\omega_0)t}d\omega, \\
i\frac{dA_2}{dt} &= \int_0^\infty V(\omega)B_\omega^{(+)}e^{-i[(\omega-\omega_0)t-k_\omega R]}d\omega \\
&\quad + \int_0^\infty V(\omega)B_\omega^{(-)}e^{-i[(\omega-\omega_0)t+k_\omega R]}d\omega,
\end{aligned}
\tag{1}
$$

$$i\frac{dB_\omega^{(+)}}{dt} = V(\omega)A_1e^{i(\omega-\omega_0)t} + V(\omega)A_2e^{i[(\omega-\omega_0)t-k_\omega R]},$$

$$i\frac{dB_\omega^{(-)}}{dt} = V(\omega)A_1e^{i(\omega-\omega_0)t} + V(\omega)A_2e^{i[(\omega-\omega_0)t+k_\omega R]},$$

where the matrix element $\langle e0|\hat{V}|g1_\omega\rangle$ of the operator $\hat{V}$ representing the interaction of an atom with the field is denoted by $V(\omega)$ for the atom at the origin of coordinates. For simplicity, we assume (without any loss of generality) that this matrix element is real-valued. For the atom at point R, the matrix element in question must therefore be multiplied by the factor $\exp(\pm ik_\omega R)$, where the sign corresponds to the direction of wave vector $\mathbf{k}_\omega$ of the $|1_\omega^\pm\rangle$ photon. Also, we have denoted by ω_0 the frequency of the $e \longrightarrow g$ transition.

A method for exactly solving the set of Eqs. (1) is given in Appendix A. This is a generalization of the method proposed in [16–18] for analogous problems involving a continuous spectrum. In the main body of the text, we will restrict our consideration to a well-known approximation that consists in the following. The matrix element of interaction, V, is assumed to be independent of the emitted-photon frequency ω and is set to $V(\omega_0)$. Further, integration in the first two equations of the set of Eqs. (1) is extended to the entire infinite axis. In the case of a single atom, it is the approximation in which one arrives at a precisely exponential decay of the excited-state amplitude and at a precisely Lorentzian spectrum of emitted photons (see the discussion at the beginning of the Introduction). Yet, information about that part of the Lamb shift which depends formally on the form of function $V(\omega)$ is lost in this case (see, for example, [2]), but we assume here that the Lamb shift was originally taken into account in the value of transition frequency ω_0, a quantity that is known from experimental data.

The meaning of the above approximation and the consequences that it entails become clear if one formally integrates the third and the fourth equation in the set of Eqs. (1) for the initial condition $B_\omega^{(\pm)}(0) = 0$ and substitutes the functions $B_\omega^{(\pm)}(t)$, obtained as a result of integration, into the first two equations. This yields the set of two integro-differential equations

$$\frac{dA_1}{dt} = -2\int_0^t \mathcal{K}_0(t-\tau)A_1(\tau)d\tau$$
$$-\int_0^t [\mathcal{K}_+(t-\tau) + \mathcal{K}_-(t-\tau)]A_2(\tau)d\tau,$$
$$\frac{dA_2}{dt} = -2\int_0^t \mathcal{K}_0(t-\tau)A_2(\tau)d\tau \qquad (2)$$
$$-\int_0^t [\mathcal{K}_+(t-\tau) + \mathcal{K}_-(t-\tau)]A_1(\tau)d\tau,$$

where the kernels $\mathcal{K}_0(t-\tau)$ and $\mathcal{K}_\pm(t-\tau)$ can be represented in an integral form; that is,

$$\mathcal{K}_0(x) = \int_0^\infty V^2(\omega)e^{-i(\omega-\omega_0)x}d\omega,$$
$$\mathcal{K}_\pm(x) = \int_0^\infty V^2(\omega)e^{\pm ik_\omega R}e^{-i(\omega-\omega_0)x}d\omega. \qquad (3)$$

We can see that, if one sets $V(\omega) = V(\omega_0)$ and replaces in (3) the lower limit of integration by $-\infty$, the kernel $\mathcal{K}_0$ reduces to a delta function, with the result that the derivatives of amplitudes A_1 and A_2 at the instant t do not depend on the values of amplitudes A_1 and A_2, respectively, at previous instants of time. The situation around the kernels $\mathcal{K}_\pm$ is subtler. For them to be also reducible to delta functions, it is necessary, in addition to invariability of $V^2(\omega)$, that the medium be nondispersive—that is, the frequency dependence of the wave vector k_ω must be linear. It is obvious that this requirement is much more significant physically than the preceding one.[1] In the main body of the text, we assume that the requirement in question is satisfied, but, in Appendix A, we prove that nonzero dispersion does not violate qualitative effects considered in the present study.

Thus, we assume that $k_\omega \propto \omega$ and denote by t_0 the time it takes for light to travel from atom 1 to atom 2 and in the opposite direction. We then have

$$\mathcal{K}_+(t-\tau) \propto e^{-i\omega_0 t_0}\delta(t-\tau-t_0),$$
$$\mathcal{K}_-(t-\tau) \propto e^{-i\omega_0 t_0}\delta(t-\tau+t_0). \qquad (4)$$

Upon substitution of relations (4) into Eqs. (2), we see that the term involving $\mathcal{K}_-$ drops out since the relevant delta function is concentrated off the integration interval $(0, t)$. Relating the rate of the spontaneous decay of the excited state, Γ, to the quantity $V^2(\omega_0)$ as

$$\Gamma = 4\pi V^2(\omega_0) \qquad (5)$$

and introducing the notation $\gamma = \Gamma/2$, we obtain the set of equations

$$\frac{dA_1(t)}{dt} = -\gamma A_1(t) - \gamma e^{i\omega_0 t_0}A_2(t-t_0),$$
$$\frac{dA_2(t)}{dt} = -\gamma A_2(t) - \gamma e^{i\omega_0 t_0}A_1(t-t_0), \qquad (6)$$

---

[1] In particular, this requirement is satisfied exactly in a vacuum, but it can also be satisfied for the so-called leading wave in a waveguide whose cross section is multiply connected [19].

where the derivative of amplitude A_1 depends on the A_2 value delayed by time t_0 and vice versa.

Further, we consider the dynamics of the system in time (Section 3) and the spectrum of radiation in each of the two directions (Section 4).

3. TIME EVOLUTION OF THE SYSTEM

As was discussed above, we are interested in the situation where, at the initial instant $t = 0$, atom 1 is excited, while atom 2 is not excited. For the initial conditions $A_1(0) = 1$ and $A_2(0) = 0$ that correspond to this case, the set of Eqs. (6) can easily be solved by induction. The result, which can be straightforwardly verified by means of a direct substitution into the original equations, has the form

$$A_1(t) = \sum_{n=0}^{\infty} \frac{\theta(t-2nt_0)}{(2n)!} \times [\gamma e^{i\omega_0 t_0}(t-2nt_0)]^{2n} e^{-\gamma(t-2nt_0)}, \tag{7}$$

$$A_2(t) = -\sum_{n=0}^{\infty} \frac{\theta[t-(2n+1)t_0]}{(2n+1)!} \times \{\gamma e^{i\omega_0 t_0}[t-(2n+1)t_0]\}^{2n+1} e^{-\gamma[t-(2n+1)t_0]}, \tag{8}$$

where θ is the Heaviside theta function. At each instant of time, the series on the right-hand sides of (7) and (8) are finite. New terms come into play at instants that are integral multiples of t_0. For $\exp(i\omega_0 t_0) \neq \pm 1$, each new term appears with a phase that differs from the phase of the preceding term; therefore, it is quite natural that, for $t \longrightarrow \infty$, the amplitudes $A_1(t)$ and $A_2(t)$ tend to zero, as they must. However, at $\exp(i\omega_0 t_0) = \pm 1$—that is, in the case where $\omega_0 t_0 = m\pi$, where m is an arbitrary integer—all terms in the expression for A_1 are positive, while all terms in the expression for A_2 have a sign opposite to the sign of $\exp(i\omega_0 t_0)$. It turns out that, in this case, the absolute values of both amplitudes tend to a constant for $t \longrightarrow \infty$:

$$\lim_{t\to\infty} A_1(t) = -e^{-i\omega_0 t_0} \lim_{t\to\infty} A_2(t) = \frac{1}{2(1+\gamma t_0)} \quad (\omega_0 t_0 = m\pi). \tag{9}$$

The proof of formula (9) is given in Appendix B [see (B.16) and comments to this formula]. The dynamics of the transient process for the probabilities $\mathcal{P}_{1,2}(t) = |A_{1,2}(t)|^2$ of finding the atoms in the excited states is illustrated in Fig. 1 for various values of parameter γt_0.

The case where parameter γt_0 appearing in (9) is much less than unity is the most realistic. The total probability of finding the system in states $|eg\rangle$ and $|ge\rangle$, where one of the atoms is excited, while the other is not excited, then tends to a value approximately equal to 1/2. This fact means that stable superpositions of states $|eg\rangle$ and $|ge\rangle$ can exist. In the case where the interatomic spacing is equal to an integral number of wavelengths corresponding to the transition frequency, the antisymmetric superposition $(|eg\rangle - |ge\rangle)/\sqrt{2}$ is stable, while, in the case where the interatomic spacing is equal to a half-integer number of wavelengths, the symmetric superposition $(|eg\rangle + |ge\rangle)/\sqrt{2}$ is stable. We note that, if the interatomic spacing is formally set to zero, we arrive at the classic case considered by Dicke [11], where it was precisely the antisymmetric superposition that was stable.

However, the situation where the atoms are spatially separated differs qualitatively from the Dicke system, where the stable superposition is an eigenstate of the "atom + field" system. In the case considered here, a more complicated superposition involving, in addition to states $|eg0\rangle$ and $|ge0\rangle$, continuum states $|gg1_\omega^{(+)}\rangle$ and $|gg1_\omega^{(-)}\rangle$ forming a wave packet spatially bounded between the atoms is an eigenstate. Quantitatively, this fact will be interpreted below upon analyzing the radiation spectrum of our system. In this section, we will additionally consider, by way of example, only the dynamics of the system where a symmetric superposition of states $|eg0\rangle$ and $|ge0\rangle$ is initially prepared and where the interatomic spacing is close to a half-integer number of wavelengths. If this is so and if the equality $\omega_0 t_0 = (2n+1)\pi$ holds exactly, the dynamics reduces to a transient process where the above stable state is formed, but, if this equality holds only approximately, a relatively fast transient process is accompanied by a relatively slow decay of the state formed, which must now be treated as metastable.

Thus, we consider set of Eqs. (6) with the initial conditions $A_1(0) = A_2(0) = 1/\sqrt{2}$. The solution has the form

$$A_1(t) = A_2(t) = \frac{1}{\sqrt{2}} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \times \theta(t-nt_0)[\gamma e^{i\omega_0 t_0}(t-nt_0)]^n e^{-\gamma(t-nt_0)}. \tag{10}$$

It was mentioned above that, at $\exp(i\omega_0 t_0) = -1$, the total probability $\mathcal{P}(t) = |A_1(t)|^2 + |A_2(t)|^2$ of finding the atoms in the excited states tends to a constant for $t \longrightarrow \infty$ (illustrative examples are given in Fig. 2):

$$\lim_{t\to\infty} \mathcal{P}(t) = \frac{1}{(1+\gamma t_0)^2}. \tag{11}$$

For $\gamma t_0 \ll 1$, this constant is close to unity. Any small

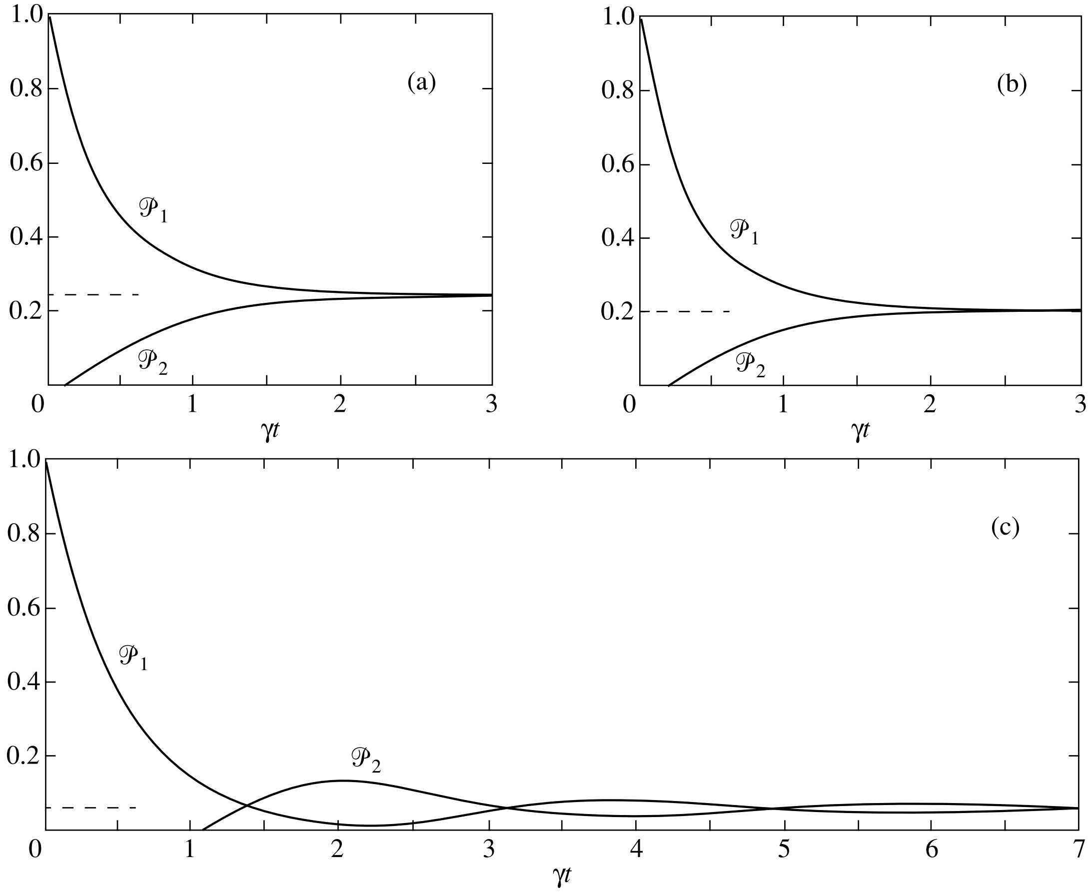

Fig. 1. Dynamics of the transient process from the initial state where one of the atoms is excited to a stable superposition of two states where amplitudes A_1 and A_2 of states $|eg0\rangle$ and $|ge0\rangle$, respectively, are given by formula (9). The quantity $\mathcal{P}_1$ ($\mathcal{P}_2$) is the probability of finding atom 1 (2) in the excited state. The parameter γt_0 was set to (a) 0.01, (b) 0.1, and (c) 1.

deviation of the quantity $\exp(i\omega_0 t_0)$ from –1 leads to an irreversible decay of the excited states, but the rate of this decay decreases as $\exp(i\omega_0 t_0)$ tends to –1. The illustration in Fig. 3 corresponds to the parameter value of $\gamma t_0 = 0.1$, in which case the limit in (11) is approximately equal to 0.83. This figure displays three dependences $\mathcal{P}(t)$ at various small values of the parameter $\beta = |\omega_0 t_0 - (2n+1)\pi|$. The scales for these dependences are significantly different, but they become virtually identical if we choose $\beta^2 t$ for a variable.

The last result can be interpreted in terms of Fermi's well-known golden rule for transitions to a continuous spectrum. We are interested here in the rate of decay of a symmetric superposition of states $|eg\rangle$ and $|ge\rangle$; in order to apply the corresponding golden-rule formula $\mathcal{W} = 2\pi|F(\omega_0)|^2$ (see, for example, [20]) to this case, matrix element $F(\omega)$ of the interaction of this superposition with states of the doubly degenerate continuous spectrum must be expressed in terms of the matrix element $V(\omega)$ appearing in the original set of Eqs. (1); that is,

$$F(\omega) = \frac{V}{\sqrt{2}}(1 + e^{\pm i\omega t_0}). \tag{12}$$

Using this relation and Eq. (5), we can express the rate $\mathcal{W}$ in terms of γ as

$$\mathcal{W} = \gamma\left|1 + e^{i\omega_0 t_0}\right|^2 \approx \beta^2\gamma. \tag{13}$$

In the second, approximate, equality, we have taken into account the proximity of the quantity $\omega_0 t_0$ to $(2n + 1)\pi$ and used the small parameter β, which was introduced above.

The decay curve $\mathcal{P}(t) \approx \exp(-\beta^2\gamma t)$ corresponding to the golden rule is given in Fig. 3 along with one of the exact solutions. Of course, it does not describe the transient process in question. Moreover, the exact solution leads to a somewhat slower decrease than that which is prescribed by the golden rule. The reason is that there is a small effective shift of the central frequency of the transition from the metastable superposition of states.

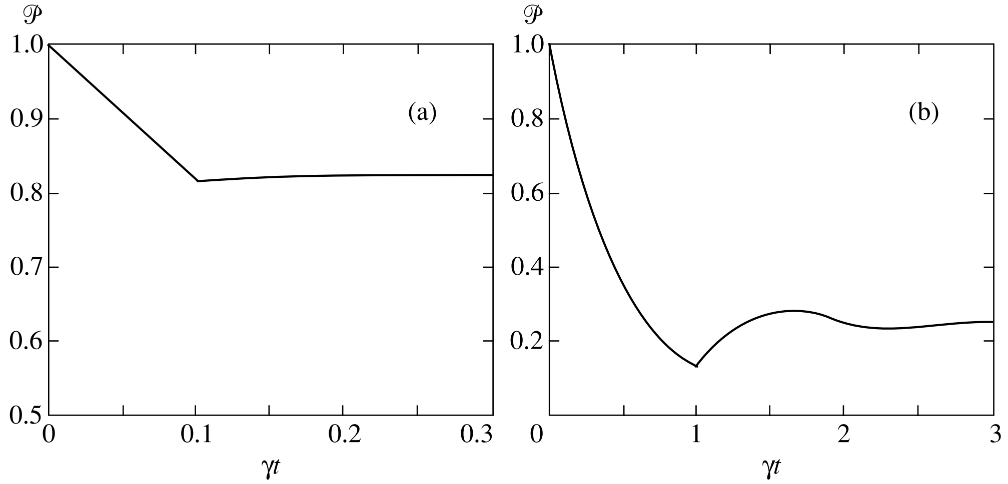

Fig. 2. Dynamics of the decay of part of the population $\mathcal{P}$ of the upper level for the system of two atoms in the case where the interatomic spacing is exactly equal to a half-integer number of wavelengths corresponding to the atomic-transition frequency and where a symmetric superposition of states $|eg\rangle$ and $|ge\rangle$ is initially excited. The decay proceeds from the value of unity to the value given by (11). The parameter γt_0 was set to (a) 0.1 and (b) 1.

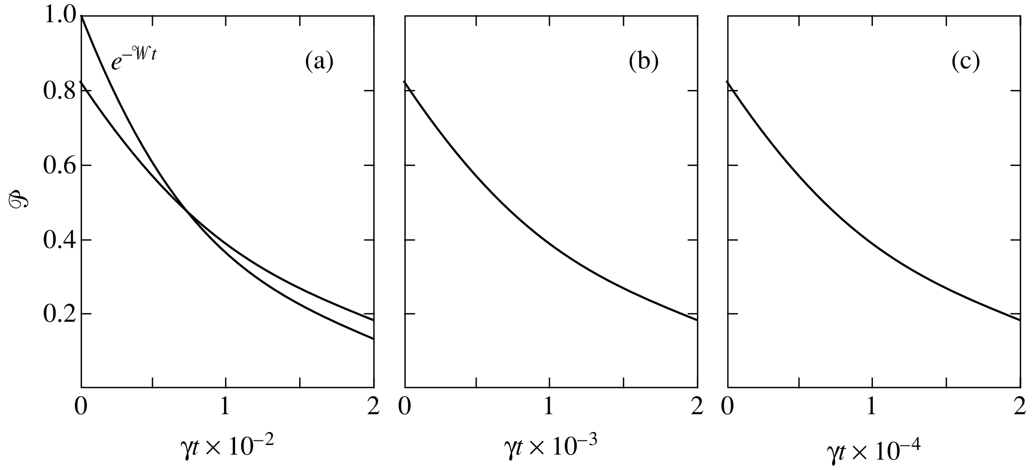

Fig. 3. Dynamics of the decay of a metastable symmetric superposition of states $|eg\rangle$ and $|ge\rangle$. The quantity $\mathcal{P}$ is the total probability of finding the atoms in the excited state. The quantity γ is defined in the main body of the text. The parameter γt_0 is set to 0.1. The interatomic spacing was taken to be close to a half-integer number of wavelengths corresponding to the atomic-transition frequency ω_0, so that the parameter $\beta = |\omega_0 t_0 - (2n+1)\pi|$ is small: (a) $\beta^2 = 10^{-2}$, (b) $\beta^2 = 10^{-3}$, and (c) $\beta^2 = 10^{-4}$. The one-exponential reference curve $\exp(-\mathcal{W}t)$, where $\mathcal{W}$ is given by (13), is presented in Fig. 3a for the sake of comparison. A much faster transient process—that within which $\mathcal{P}$ decreases from unity to the value in (11)—is shown in Fig. 2.

We will discuss this effect upon considering the radiated-photon spectrum.

4. RADIATION SPECTRUM

It is reasonable to begin by addressing the simpler problem of the two considered in the preceding section—namely, the problem of the decay of an initially prepared symmetric (antisymmetric) superposition of states $|eg0\rangle$ and $|ge0\rangle$. The simplicity of this problem consists in that the spectra of radiated photons are identical in the two propagation directions. At the same time, the spectra in the two directions may be significantly different in the problem where only one of the two atoms is initially excited. This effect will be analyzed below.

Thus, we consider the solution in (10) for amplitudes $A_1(t)$ and $A_2(t)$ describing states $|eg0\rangle$ and $|ge0\rangle$, respectively. In order to obtain the spectrum of emitted photons, it is necessary to substitute (10) into the last two equations in set of Eqs. (1) and to integrate them[2]

---

[2] The further procedure is correct if amplitudes $A_1(t)$ and $A_2(t)$ tend to zero for $t \longrightarrow \infty$. On the contrary, more sophisticated calculations are required (see Appendix B) if $\exp(i\omega_0 t_0) = -1$ and, accordingly, $\lim_{t \to \infty} A_{1,2}(t) \neq 0$.

from the initial instant $t = 0$ to the final instant of observation (we set it to $t = \infty$). For a normalization to be convenient, we define the spectrum $\mathcal{S}_+(\omega)$ (the plus sign means that, in this case, a symmetric superposition of states $|eg0\rangle$ and $|ge0\rangle$ is initially prepared) as the sum of the probabilities of the emission of photons of the given frequency in the two directions:

$$\mathcal{S}_+(\omega) = \left|B_\omega^{(+)}\right|^2 + \left|B_\omega^{(-)}\right|^2.$$

As an intermediate result, we have

$$\begin{aligned} B_\omega^{(\pm)}(t = \infty) &= -iV(\omega)(1 + e^{\mp i\omega t_0}) \\ &\times \frac{1}{\sqrt{2}}\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}(\gamma e^{i\omega_0 t_0})^n \mathcal{J}_n, \end{aligned} \tag{14}$$

where

$$\begin{aligned} \mathcal{J}_n &= \int_{nt_0}^{\infty}(t - nt_0)^n e^{-\gamma(t - nt_0)} e^{-i(\omega - \omega_0)} dt \\ &= \frac{n!e^{in(\omega - \omega_0)t_0}}{[\gamma - i(\omega - \omega_0)]^{n+1}}. \end{aligned} \tag{15}$$

The further procedure involves substituting (15) into (14), summing a geometric progression, and squaring the absolute values of $B_\omega^{(+)}$ and $B_\omega^{(-)}$. In this way, we reduce the expression for the emission spectrum to the form

$$\begin{aligned} \mathcal{S}_+(\omega) &= \frac{1}{\pi} \\ &\times \frac{\gamma(1 + \cos\omega t_0)}{\gamma^2(1 + \cos\omega t_0)^2 + (\omega - \omega_0 - \gamma\sin\omega t_0)^2}, \end{aligned} \tag{16}$$

where, for simplicity, we replaced, in accordance with (5), the slowly varying function $V^2(\omega)$ by the constant $\Gamma = 2\gamma$. Without going into details, we would also like to note that the result for spectrum $\mathcal{S}_-(\omega)$ corresponding to an initially prepared antisymmetric superposition of states $|eg0\rangle$ and $|ge0\rangle$,

$$\begin{aligned} \mathcal{S}_-(\omega) &= \frac{1}{\pi} \\ &\times \frac{\gamma(1 - \cos\omega t_0)}{\gamma^2(1 - \cos\omega t_0)^2 + (\omega - \omega_0 + \gamma\sin\omega t_0)^2}, \end{aligned} \tag{17}$$

differs from formula (16) only in the signs in front of the cosine and the sine.

Further, expressions (16) and (17) would be coincident in form with a Lorentzian contour if the arguments of the sine and the cosine were constants. But in the case of $\gamma t_0 \ll 1$, which is of greatest interest to us (see the preceding section), this is indeed so: argument ωt_0 changes insignificantly within the characteristic spectral-contour width, which, in any case, is less than 2γ. For the case of $\gamma t_0 \ll 1$, it is therefore quite natural to introduce the effective central frequency of the contour,

$$\tilde{\omega}_0 = \omega_0 \pm \gamma\sin\omega_0 t_0 \quad (\gamma t_0 \ll 1), \tag{18}$$

where the upper (lower) sign corresponds to the symmetric (antisymmetric) initial condition. Accordingly, the effective half-width of the contour is

$$\tilde{\gamma} = \gamma(1 \pm \cos\tilde{\omega}_0 t_0) \quad (\gamma t_0 \ll 1), \tag{19}$$

while the shape of the emitted-photon spectrum can be approximately represented in the canonical Lorentzian form

$$\mathcal{S}_\pm(\omega) \approx \frac{1}{\pi}\frac{\tilde{\gamma}}{\tilde{\gamma}^2 + (\omega - \tilde{\omega}_0)^2} \quad (\gamma t_0 \ll 1). \tag{20}$$

This result is quite consistent with the conclusions drawn in Section 3 that concern the dynamics of the radiative decay of an initially prepared symmetric superposition state. By way of example, we indicate that, for $\mathcal{S}_+(\omega)$ taken in the case where the above parameter $\beta = |\omega_0 t_0 - (2n + 1)\pi|$ (see Fig. 3 and explanations in the main body of the text) is much less than unity, the effective width of the emitted-photon spectrum is $\tilde{\Gamma} = 2\tilde{\gamma} \approx \beta^2\gamma(1 - 2\gamma t_0)$ in accordance with (19) (here, we retained only the first two terms in the expansion in the small parameter γt_0). In the zero-order approximation in γt_0, this width is exactly equal to decay rate $\mathcal{W}$ (13), corresponding to the golden rule, while, with allowance for the correction, it is somewhat less, which follows from the graphs in Fig. 3.

Let us now examine the case where only one of the atoms is initially excited. In order to obtain the radiation spectrum, we use the solutions in (7) and (8) for amplitudes $A_1(t)$ and $A_2(t)$, respectively, and substitute them into set of Eqs. (1). As a result, we obtain the probabilities for the emission of photons of the given frequency in the two opposite directions:

$$\begin{aligned} \mathcal{S}_\odot(\omega) &= \frac{1}{4\pi(1 + \sin^2\omega t_0)} \\ &\times \left[\frac{(1 + 2\sin^2\omega t_0)\gamma(1 + \cos\omega t_0)}{\gamma^2(1 + \cos\omega t_0)^2 + (\omega - \omega_0 - \gamma\sin\omega t_0)^2}\right. \\ &- \frac{\sin\omega t_0(\omega - \omega_0 - \gamma\sin\omega t_0)}{\gamma^2(1 + \cos\omega t_0)^2 + (\omega - \omega_0 - \gamma\sin\omega t_0)^2} \\ &+ \frac{(1 + 2\sin^2\omega t_0)\gamma(1 - \cos\omega t_0)}{\gamma^2(1 - \cos\omega t_0)^2 + (\omega - \omega_0 + \gamma\sin\omega t_0)^2} \\ &+ \left.\frac{\sin\omega t_0(\omega - \omega_0 + \gamma\sin\omega t_0)}{\gamma^2(1 - \cos\omega t_0)^2 + (\omega - \omega_0 + \gamma\sin\omega t_0)^2}\right] \end{aligned} \tag{21}$$

in the “free” direction and

$$\begin{aligned}\mathcal{S}_{\otimes}(\omega) = {} & \frac{1}{4\pi(1+\sin^2\omega t_0)} \\ & \times\left[\frac{\gamma(1+\cos\omega t_0)}{\gamma^2(1+\cos\omega t_0)^2+(\omega-\omega_0-\gamma\sin\omega t_0)^2}\right. \\ & + \frac{\sin\omega t_0(\omega-\omega_0-\gamma\sin\omega t_0)}{\gamma^2(1+\cos\omega t_0)^2+(\omega-\omega_0-\gamma\sin\omega t_0)^2} \\ & + \frac{\gamma(1-\cos\omega t_0)}{\gamma^2(1-\cos\omega t_0)^2+(\omega-\omega_0+\gamma\sin\omega t_0)^2} \\ & \left. - \frac{\sin\omega t_0(\omega-\omega_0+\gamma\sin\omega t_0)}{\gamma^2(1-\cos\omega t_0)^2+(\omega-\omega_0+\gamma\sin\omega t_0)^2}\right]\end{aligned} \qquad (22)$$

in the direction from the initially excited atom through the initially nonexcited atom.

It obviously follows from (21) and (22) that the total spectrum $\mathcal{S}_{\odot}(\omega) + \mathcal{S}_{\otimes}(\omega)$ of photons emitted in the two directions is equal, as it must be, to $[\mathcal{S}_{+}(\omega) + \mathcal{S}_{-}(\omega)]/2$—that is, the half-sum of the spectra corresponding to the symmetric and antisymmetric initial conditions. However, spectra $\mathcal{S}_{\odot}(\omega)$ and $\mathcal{S}_{\otimes}(\omega)$ themselves may be significantly different (see the two examples in Fig. 4). First of all, the total probabilities of photon emission in the two directions are not equal to each other: the integral of $\mathcal{S}_{\odot}(\omega)$ is always greater than the integral of $\mathcal{S}_{\otimes}(\omega)$. Introducing the ratio

$$\mathcal{R} = \frac{\int_{-\infty}^{\infty}\mathcal{S}_{\odot}(\omega)d\omega}{\int_{-\infty}^{\infty}\mathcal{S}_{\otimes}(\omega)d\omega} \qquad (23)$$

and calculating, by a method similar to that used to evaluate the integral in (B.11) in Appendix B, the relevant integrals of various terms appearing in expressions (21) and (22), we obtain

$$\mathcal{R} = \frac{2 - \dfrac{\cos(2\omega_0 t_0)}{\cosh(2\sqrt{2}\gamma t_0)} + \dfrac{5}{4}\sqrt{2}\tanh(2\sqrt{2}\gamma t_0)}{1 + \dfrac{3}{4}\sqrt{2}\tanh(2\sqrt{2}\gamma t_0)}. \qquad (24)$$

It follows from formula (24) that, for $\gamma t_0 \ll 1$, the ratio in (23) takes a maximum value of $\mathcal{R} \approx 3$ at interatomic distances differing from an integral number of half-wavelengths by λ/4. The corresponding small-scale dependences $\mathcal{R}(\omega_0 t_0)$ are shown in Fig. 5 for $\gamma t_0 = 0.1$ and $\gamma t_0 = 0.5$. It can be seen that, as parameter γt_0 increases, the dependence of $\mathcal{R}$ on the interatomic spacing is smoothed out over the wavelength scale, tending to the constant $\mathcal{R}|_{\gamma t_0 \gg 1} \approx 2\sqrt{2} - 1 \approx 1.83$.

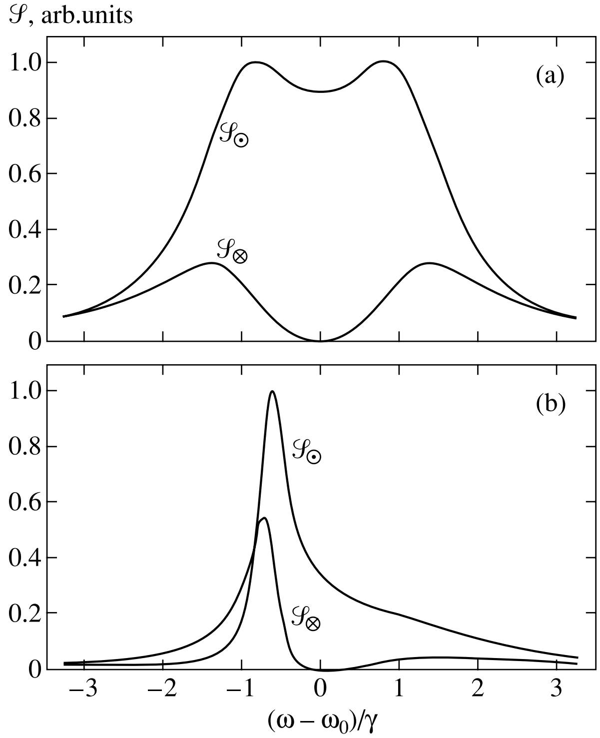

Fig. 4. Central part of the spectrum of photons emitted in two opposite directions for the case where one of the atoms is initially excited. The two examples demonstrate distinctions between spectra $\mathcal{S}_{\odot}(\omega)$ and $\mathcal{S}_{\otimes}(\omega)$. The normalization was performed with respect to the maximum of function $\mathcal{S}_{\odot}(\omega)$. Parameter γt_0 was set to 0.1 for both examples. The interatomic spacing differs from an integral number of half-wavelengths by (a) λ/4 and (b) λ/8.

To conclude this section, we would like to note the fact (see, in particular, Fig. 4) that function $\mathcal{S}_{\otimes}(\omega)$ is exactly equal to zero at transition frequency ω_0. Thus, we can say that a nonexcited atom is like an ideal mirror for a photon emitted by an identical atom strictly at the atomic-transition frequency.

5. SOME PROPERTIES OF STABLE STATES

In Section 2, we indicated that a nondecaying superposition of the two states $|eg\rangle$ and $|ge\rangle$ (it is symmetric if the interatomic spacing is equal to a half-integer number of wavelengths and antisymmetric if this spacing is equal to an integral number of wavelengths) is “supported” by a wave packet bounded between the

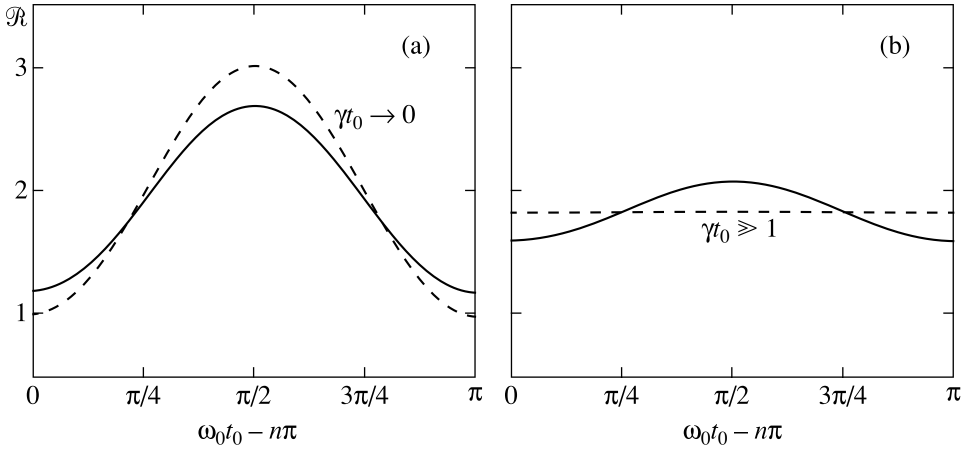

Fig. 5. Ratio (24) of the total probabilities of photon emission in two opposite directions in the case where one of the atoms is excited initially as a function of the interatomic spacing over the scale $\lambda/2$: (a) evolution of ratio $\mathcal{R}$ as parameter γt_0 grows from extremely small values (dashed curve) to a value of 0.1 (solid curve) and (b) subsequent evolution of $\mathcal{R}$ as γt_0 increases to 0.5 (solid curve) and further to greater limiting values (dashed curve).

atoms. This means that, for the "atoms + field" system, the superposition

$$\begin{aligned} |\Upsilon\rangle &= C_1|eg0\rangle + C_2|ge0\rangle \\ &+ \int C_\omega^{(+)}|gg1_\omega^{(+)}\rangle d\omega + \int C_\omega^{(-)}|gg1_\omega^{(-)}\rangle d\omega, \end{aligned} \tag{25}$$

where $C_2 = \pm C_1$ and where the relation between $C_{1,2}$ and $C_\omega^{(\pm)}$ can be derived from the corresponding formulas of Appendix A, is the corresponding stable eigenstate that is embedded in the continuous spectrum and is "discrete" in the sense that it is normalized to unity rather than to a delta function (for more details, see Appendix A). In particular, the relation between $C_{1,2}$ and $C_\omega^{(\pm)}$ indicates which fraction of energy is associated with the excitation of the atoms and which is concentrated in the "supporting" field.

Here, we would like to extend somewhat the issue in question as follows. By way of example, we examine the situation where a symmetric superposition of states $|eg\rangle$ and $|ge\rangle$ is excited at $t = 0$ and where the interatomic spacing is exactly equal to a half-integer number of wavelengths and address the problem of assessing the following probabilities:

(i) that of finding an atom in the excited state for $t \longrightarrow \infty$ ($\mathcal{P}_{\text{at}}$);

(ii) that of recording a photon by a detector placed at infinity (the probability of emission into a free space, $\mathcal{P}_{\text{free}}$);

(iii) that of finding a photon in the supporting field between the atoms for $t \longrightarrow \infty$ ($\mathcal{P}_{\text{support}}$).

The first probability is given by formula (11). The answers to the questions associated with the second and third probabilities can be obtained by two methods. One of these (indirect) employs the relation between expansion coefficients $C_{1,2}$ and $C_\omega^{(\pm)}$ in eigenstate Υ (25) of our system (see Appendix A). We will rely on the other, clearer, method, which follows from the proof of formula (11) in Appendix B. There, one uses, first of all, the general normalization condition (B.3) (as a matter of fact, this is the equality $\mathcal{P}_{\text{at}} + \mathcal{P}_{\text{free}} + \mathcal{P}_{\text{support}} = 1$) and then determines the integral of the total probability of finding the system in the state $|gg1_\omega^{(\pm)}\rangle$ with respect to the frequency ω. This integrated probability involves three terms. The first contains the integral $\mathcal{Q}_1$ (B.9) and is proportional to the quantity $\mathcal{P}_{\text{at}}$ (or A_∞^2 in the notation adopted in Appendix B); hence, it is $\mathcal{P}_{\text{support}}$. The second term, which contains the integral $\mathcal{Q}_2$ (B.10), tends to zero for $t \longrightarrow \infty$; it describes the interference between two fields—the field "emitted into a free space" and the supporting field. The remaining third term, which contains the integral $\mathcal{Q}_3$ (B.11), is $\mathcal{P}_{\text{free}}$. By using formulas (11), (B.9), and (B.15), we arrive at the final expressions

$$\begin{gathered} \mathcal{P}_{\text{at}} = \frac{1}{(1+\gamma t_0)^2}, \quad \mathcal{P}_{\text{free}} = \frac{\gamma t_0}{1+\gamma t_0}, \\ \mathcal{P}_{\text{support}} = \frac{\gamma t_0}{(1+\gamma t_0)^2}. \end{gathered} \tag{26}$$

We can see the following:

(i) If condition $\gamma t_0 \ll 1$, which was used as a rule in this article for the purpose of illustration, is satisfied, the initial excitation remains predominantly concentrated in the atoms, while the probability of finding a photon in the "radiated" field is approximately equal to the probability of finding a photon in the supporting field (both probabilities are about γt_0).

(ii) At $\gamma t_0 = 1$, part of the excitation energy is carried away, with a probability of 1/2, in the form of radiation to a free space, while the remaining part of the excitation energy is distributed in equal shares (1/4) between the atoms and the supporting field.

(iii) Under condition $\gamma t_0 \gg 1$, the probability of finding a photon in the supporting field is a small quantity on the order of $(\gamma t_0)^{-1}$, while the probability of finding an atom in the excited state is even smaller, on the order of $(\gamma t_0)^{-2}$—that is, the energy of the initial excitation of the system is predominantly carried away to a free space.

6. CONCLUSIONS

In this study, special features of the dynamics and of the spectrum of radiation from a system consisting of two spatially separated identical atoms have been investigated in detail for the one-dimensional case. The most interesting effect here is subradiance, which occurs if the atoms are separated by a distance approximately equal to an integral or a half-integer number of wavelengths corresponding to the atomic-transition frequency.

At first glance, it seems that the above situation is idealized and is far from experimental capabilities. However, this is not quite so. Within the last decade, so-called holey optical waveguides [21, 22] were created, where a high-quality one-mode regime can be realized. Moreover, initial experiments were performed [23] that demonstrated the possibility of channeling an atom into a hollow waveguide of a transverse dimension about λ.

As to localization of atoms, there are also here a number of possibilities following from already available experimental realizations: for example, the localization of an ion in an electromagnetic or in a radio-frequency trap, as in the experiments reported in [14, 15] and mentioned in the Introduction, and the localization of an atom at the nodes of a standing light wave by means of laser cooling of atoms (optical lattices in one-, two-, and three-dimensional cases [24–26]).

Presently, ion traps make it possible to achieve localization within a few nanometers. Light traps (see the first experiment in [27] and review article [28]) cannot yet ensure so high a degree of localization (the localization there is approximately one order of magnitude poorer than in ion traps), but we would like to mention in this connection a new method of localization precisely in a waveguide, that which was recently proposed in [29] and which can be realized on the basis of geometry where the potential of the gradient force of a light field is concentrated in the transverse direction over a distance much shorter than the wavelength.

One should also bear in mind the tight localization of an atom in condensed media. It is especially appealing in this sense to follow the trend of nanophysics which dictates the use of quantum dots rather than atoms.

In conclusion, we will briefly list possible applications.

The discovered stable superposition state of two atoms belongs to the class of entangled states. It is well known that such states have been assumed to play a special role in future quantum calculations.[3] It should be emphasized that the entangled states considered in the present study can be long-lived in the optical (visible) range, and this is of particular interest; also, a transition from a symmetric to an antisymmetric state can readily be implemented by changing the position of one of the atoms by λ/2 or by slightly changing the medium index of refraction.

An extremely high sensitivity of the effects considered in this study (in particular, of the time of superposition-state decay, as well as of the spectra of photons emitted in two opposite directions) to the interatomic spacing and, possibly, to the surroundings, which may weakly affect the atomic-transition frequency, is favorable for harnessing these effects in precision measurements and diagnostics.

ACKNOWLEDGMENTS

We are grateful to A.V. Kamchatnov and V.P. Krainov for enlightening discussions on some mathematical problems in the context of this study.

This work was supported in part by the Russian Foundation for Basic Research (project no. 00-15-96612).

APPENDIX A

Let us consider a general method for solving the set of Eqs. (1) without simplifying assumptions made in the main body of the text. First, we note[4] that set of Eqs. (1), which consists of four integro-differential equations, can be broken down into two independent sets of two equations each by means of the substitution

$$\begin{aligned} A_g &= \frac{1}{\sqrt{2}}(A_1 + A_2), \quad A_u = \frac{1}{\sqrt{2}}(A_1 - A_2), \\ B_\omega^{(g)} &= \frac{1}{\sqrt{2}}(B_\omega^{(+)} e^{ik_\omega R/2} + B_\omega^{(-)} e^{-ik_\omega R/2}), \qquad \text{(A.1)} \\ B_\omega^{(u)} &= -\frac{i}{\sqrt{2}}(B_\omega^{(+)} e^{ik_\omega R/2} - B_\omega^{(-)} e^{-ik_\omega R/2}). \end{aligned}$$

---

[3] Here, it is worthwhile to mention the review article of Bagratin *et al.* [30], which contains a detailed list of references on the subject and which places special emphasis precisely on entangled states, quoting this term in the title.

[4] It should be borne in mind, however, that the partition of our specific problem (of the “two states and two continua,” type according to the classification proposed in [16]) into two simpler problems (each belonging to the “one state and one continuum” type) is not crucial for the possibility of deriving an analytic solution, but calculations are more cumbersome in the general case.

If $A_u = 0$, then A_g is actually the amplitude of a symmetric superposition of states $|eg\rangle$ and $|ge\rangle$, and vice versa—that is, if $A_g = 0$, A_u is the amplitude of an antisymmetric superposition of the same states. Upon making the substitution specified in (A.1), we obtain the set of equations

$$\begin{aligned} i\frac{dA_g}{dt} &= \int_0^\infty V_g(\omega)B_\omega^{(g)}e^{-i(\omega-\omega_0)t}d\omega, \\ i\frac{dB_\omega^{(g)}}{dt} &= V_g(\omega)A_g e^{i(\omega-\omega_0)t} \end{aligned} \tag{A.2}$$

for the “symmetric” part of the problem, where

$$V_g = 2V(\omega)\cos\frac{k_\omega R}{2}, \tag{A.3}$$

and the set of equations

$$\begin{aligned} i\frac{dA_u}{dt} &= \int_0^\infty V_u(\omega)B_\omega^{(u)}e^{-i(\omega-\omega_0)t}d\omega, \\ i\frac{dB_\omega^{(u)}}{dt} &= V_u(\omega)A_u e^{i(\omega-\omega_0)t} \end{aligned} \tag{A.4}$$

for the “antisymmetric” part, where

$$V_u = 2V(\omega)\sin\frac{k_\omega R}{2}. \tag{A.5}$$

In accordance with the formulation of our problems, $V(\omega)$ is a slowly varying function that decreases at infinity.

Further, the substitutions

$$\begin{aligned} A_{g,u}(t) &= a(\varepsilon)e^{-i(\varepsilon-\omega_0)t}, \\ B_\omega^{(g,u)}(t) &= b_\omega(\varepsilon)e^{-i(\varepsilon-\omega)t} \end{aligned} \tag{A.6}$$

reduce both sets of equations, that in (A.2) and that in (A.4), to an eigenvalue problem, with ε and $\|a(\varepsilon), b_\omega(\varepsilon)\|$ appearing as eigenvalues and eigenvectors, respectively; that is,

$$\begin{aligned} (\varepsilon-\omega_0)a(\varepsilon) &= \int_0^\infty G(\omega)b_\omega(\varepsilon)d\omega, \\ (\varepsilon-\omega)b_\omega(\varepsilon) &= G(\omega)a(\varepsilon), \end{aligned} \tag{A.7}$$

where the index (g or u) is omitted for brevity and where function $G(\omega)$, which describes the interaction of a discrete level with a continuous spectrum, is either V_g (A.3) or V_u (A.5), respectively.

Set of Eqs. (A.7) can be diagonalized by the well-known method that, for various problems involving a continuous spectrum, was given in [16] (autoionization states of atoms), [17] (laser-induced transitions from a discrete spectrum to a band of levels), and [18] (intramolecular vibrational relaxation), as well as in the monograph of Cohen-Tannoudji *et al.* [31], which was published later. In those studies, an analytic solution was obtained for an arbitrary (in principle) form of function $G(\omega)$ (of course, it is assumed that $G(\omega)$ is a continuous function). In the context of the present study, there is, however, a subtle importance consisting in that the case where function $G(\omega)$ vanishes at some points has special features not covered in previous studies. Below, we fill this gap.

It was proven in [17] that the spectrum of eigenvalue problem (A.7) is continuous from the boundary below which function $G(\omega)$ vanishes (without loss of generality, we assume that this lower boundary is equal to zero[5]) to infinity and can also include a discrete eigenvalue[6] $\varepsilon_- < 0$ if $G(0) \neq 0$. It turns out, however, that if there exists a point where $G(\omega) = 0$ in the vicinity of $\omega = \omega_0$, this does not exhaust the solution to the problem—namely, there can exist an additional discrete eigenvalue that is as though embedded in the continuous spectrum.

In order to prove the last statement, we explicitly write, as a function of the eigenvalue ε, the eigenvector of the problem specified by Eq. (A.7); that is,

$$\begin{aligned} a(\varepsilon) &= \frac{G(\varepsilon)}{\sqrt{\pi^2G^4(\varepsilon) + [\varepsilon-\omega_0-\mathcal{D}(\varepsilon)]^2}}, \\ b_\omega(\varepsilon) &= \frac{1}{\sqrt{\pi^2G^4(\varepsilon) + [\varepsilon-\omega_0-\mathcal{D}(\varepsilon)]^2}} \\ &\quad\times\left\{G(\varepsilon)G(\omega)\mathcal{P}\frac{1}{\varepsilon-\omega} + [\varepsilon-\mathcal{D}(\varepsilon)]\delta(\omega-\varepsilon)\right\}, \end{aligned} \tag{A.8}$$

where we used the standard conventions for distributions (see, for example, [33]), a delta function and the function $\mathcal{P}\frac{1}{x}$, which is an analogue of the function $1/x$, and introduced the notation

$$\mathcal{D}(\varepsilon) = \mathrm{Vp}\int_0^\infty \frac{G^2(\omega)}{\varepsilon-\omega}d\omega \equiv \int_0^\infty G^2(\omega)\mathcal{P}\frac{1}{\varepsilon-\omega}d\omega, \tag{A.9}$$

with the symbol Vp implying that the integral in question is taken in the sense of the principal-value pre-

---

[5] In a free one-dimensional space, $V(\omega) \propto \sqrt{\omega}$, but, in a waveguide, the lower boundary of $V(\omega)$ may be different from zero. Also, the very form of function $V(\omega)$ depends both on the shape of the waveguide cross section and on the position of the atom with respect to the waveguide axis [32].

[6] In the context of the present study, this discrete value—even if it formally exists—does not play any noticeable role since the decay rate is much less than the transition frequency ($\Gamma \ll \omega_0$, see the Introduction); therefore, so far an edge of the energy spectrum makes a negligible contribution to the dynamics of the system.

scription in accordance with the definition of the function $\mathcal{P}\frac{1}{x}$.

By means of direct substitution, it can easily be verified that (A.8) is indeed a solution to the problem specified by Eq. (A.7). In any quantum-mechanical problem dealing with a continuous spectrum, it is common practice to normalize an eigenvector to a delta function; in our case, we have

$$a(\varepsilon)a(\varepsilon') + \int b_\omega(\varepsilon)b_\omega(\varepsilon')d\omega = \delta(\varepsilon-\varepsilon').$$

However, it is obvious that formulas (A.8) may lead to indefiniteness at a punctured point ε_0 of the spectrum if the following two equalities hold simultaneously at this point:

$$G(\varepsilon_0) = 0, \quad \mathcal{D}(\varepsilon_0) = \varepsilon_0. \qquad \text{(A.10)}$$

This case requires a dedicated investigation. The result is the following. In contrast to the remaining part of the spectrum, where one and only one eigenvector (A.8) corresponds to a given eigenvalue ε, eigenvalue ε_0 is doubly degenerate—that is, two linearly independent solutions to set of Eqs. (A.7) exist for an eigenvalue ε_0 that possesses the properties specified in (A.10). It is convenient to choose the corresponding pair of orthogonal eigenvectors in the form

$$a^{(1)}(\varepsilon_0) = 0, \quad b_\omega^{(1)}(\varepsilon_0) = \delta(\omega-\varepsilon_0), \qquad \text{(A.11)}$$

$$a^{(2)}(\varepsilon_0) = C,$$

$$b_\omega^{(2)}(\varepsilon_0) = \begin{cases} C\dfrac{G(\omega)}{\varepsilon_0-\omega}, & \omega \neq \varepsilon_0 \\ 0, & \omega = \varepsilon_0. \end{cases} \qquad \text{(A.12)}$$

The first of the eigenvectors in (A.11) is merely a nonexcited state of the initially continuous spectrum; by virtue of the first of the conditions in (A.10), $G(\varepsilon_0) = 0$, this state does not interact with other levels. In contrast to all of the remaining eigenvalues, the second eigenvalue is normalized to unity under a natural additional assumption concerning the form of the function $G(\omega)$ in the vicinity of point $\omega = \varepsilon_0$; the normalization constant appearing in (A.12) is given by

$$C = \left[1 + \int_0^\infty \frac{G^2(\omega)}{(\omega-\varepsilon_0)^2}d\omega\right]^{-1/2}.$$

If $G(\omega \longrightarrow \varepsilon_0) \sim (\omega-\varepsilon_0)^\alpha$, where $\alpha > 1/2$, the constant C is a well-defined value.

The functions $G(\omega)$ that are of interest in the context of this study—namely, functions V_g (A.3) and V_u (A.5)—are products of the slowly varying function V and the quickly oscillating sine or cosine of a function that is linear in the absence of dispersion ($k_\omega \propto \omega$), but which is nonlinear in general. However, function G has zeros in the vicinity of ω_0 in any case. Also, it can easily be noticed that fulfillment of the second equality in (A.10) is simultaneously ensured by appropriately choosing parameter R (interatomic spacing) appearing in expressions (A.3) and (A.5). Considering eigenvalue ε_0 as a function of R, we do indeed see that, under condition $G(\varepsilon_0) = 0$, the quantity $\mathcal{D}(\varepsilon_0)$ (A.9) depends on R only slightly: by and large, the value $\mathcal{D}(\varepsilon_0) \approx \mathcal{D}_0$ is determined by two factors, the asymmetry of function $V(\omega)$ with respect to the atomic-transition frequency ω_0 and dispersion, which, in the present case, leads to the asymmetry of functions $\cos^2(k_\omega R/2)$ and $\sin^2(k_\omega R/2)$ with respect to their local zeros. Therefore, an R value that satisfies the second condition in (A.10) is chosen in such a way that the zero of function G is shifted from frequency ω_0 by $\mathcal{D}_0$.

In the main body of the text of this article, we disregarded the asymmetry of the function $V(\omega)$, assuming that it is constant from $-\infty$ to $+\infty$; we also disregarded dispersion. Therefore, both conditions in (A.10) were automatically satisfied if distance R was equal either to an integral (for $G = V_u$) or a half-integer (for $G = V_g$) number of wavelengths corresponding to atomic-transition frequency ω_0. In a realistic situation, R values at which the stable superposition states $(|eg\rangle \pm |ge\rangle)/\sqrt{2}$ exist will slightly change, but, as has been shown in this appendix, the presence of dispersion or the dependence of V on ω does not destroy the very existence of them.

APPENDIX B

In order to adequately treat the properties of a stable excited entangled state of two atoms, as well as the properties of radiation supporting this state and the properties of radiation emitted during the corresponding transient process, it is necessary to calculate the limit

$$A_\infty = \lim_{t\to\infty}\sum_{n=0}^{\infty}\frac{1}{n!}\theta(t-nt_0)[\gamma(t-nt_0)]^n e^{-\gamma(t-nt_0)}. \qquad \text{(B.1)}$$

In particular, it can be seen that, if the equality

$$e^{i\omega_0 t_0} = -1 \qquad \text{(B.2)}$$

holds, the limiting value of amplitudes $A_1(t)$ and $A_2(t)$ for the solution in (10) is $A_\infty/\sqrt{2}$, so that the probability of finding the atoms in excited states tends to A_∞^2 for

$t \longrightarrow \infty$ in the corresponding case of the symmetric initial condition.[7]

In order to calculate the limit in (B.1), we use the normalization condition

$$|A_1(t)|^2 + |A_2(t)|^2 + \int \left|B_\omega^{(+)}(t)\right|^2 d\omega + \int \left|B_\omega^{(-)}(t)\right|^2 d\omega = 1. \tag{B.3}$$

The amplitudes $B_\omega^{(+)}(t)$ and $B_\omega^{(-)}(t)$ are expressed in terms of $A(t) \equiv A_1(t) = A_2(t)$ by integrating the last two equations in (1). The solution in (10) formally corresponds to the case where the matrix element $V(\omega)$ of the interaction is a constant—that is, integration in (B.3) extends over the entire infinite axis. Passing to the limit, we then obtain

$$A_\infty^2 = 1 - \frac{2\gamma}{\pi} \lim_{t \to \infty} \int_{-\infty}^{\infty} (1 + \cos\omega t_0) |\mathcal{J}_\omega(t)|^2 d\omega, \tag{B.4}$$

where we used relation (5) and introduced the notation

$$\mathcal{J}_\omega(t) = \int_0^t A(\tau) e^{i(\omega - \omega_0)\tau} d\tau. \tag{B.5}$$

Upon integration of (B.5) by parts, we have

$$\mathcal{J}_\omega(t) = \frac{i}{\omega - \omega_0} \times \left[\frac{1}{\sqrt{2}} - A(t) e^{i(\omega - \omega_0)t} + \int_0^t A'(\tau) e^{i(\omega - \omega_0)\tau} d\tau\right]. \tag{B.6}$$

Further, we use (10), with the additional condition (B.2), and the explicit form of the derivative $A'(t)$ to calculate the integral in (B.6) with the upper limit tending to infinity. For $\mathcal{J}_\omega$, we then arrive in this way at an asymptotic expression in the form of the sum of an oscillating and a constant term:

$$\mathcal{J}_\omega = \frac{1}{\sqrt{2}} \times \left[-iA_\infty \frac{e^{i(\omega - \omega_0)t}}{\omega - \omega_0} + \frac{1}{\gamma(1 + e^{i\omega t_0}) - i(\omega - \omega_0)}\right]. \tag{B.7}$$

Upon substituting (B.7) into (B.4), we find that the sought quantity A_∞ satisfies the equation

$$A_\infty^2\left(1 + \frac{\gamma}{\pi}\mathcal{Q}_1\right) - \frac{2\gamma}{\pi} A_\infty \operatorname{Im}\mathcal{Q}_2 = 1 - \frac{1}{\pi}\mathcal{Q}_3, \tag{B.8}$$

where $\mathcal{Q}_{1,2,3}$ [after the change of variable $x = \omega - \omega_0$ and with allowance for condition (B.2)] is a generic notation for three integrals: of these, the first is tabular,

$$\mathcal{Q}_1 = \int_{-\infty}^{\infty} \frac{1 - \cos x t_0}{x^2} dx = \pi t_0; \tag{B.9}$$

the second,

$$\mathcal{Q}_2 = \int_{-\infty}^{\infty} \frac{(1 - \cos x t_0) e^{ixt}}{\gamma(1 - e^{ixt_0}) - ix} dx \tag{B.10}$$

tends to zero for $t \longrightarrow \infty$; and the third,

$$\mathcal{Q}_3 = \int_{-\infty}^{\infty} \frac{\gamma(1 - \cos x t_0)}{[\gamma(1 - \cos x t_0)]^2 + [x + \gamma \sin x t_0]^2} dx \tag{B.11}$$

is not quite trivial.[8] In order to calculate the integral in (B.11), we transform it to

$$\mathcal{Q}_3 = \operatorname{Re}\left[\operatorname{Vp}\int_{-\infty}^{\infty} \frac{dx}{ix + \gamma - \gamma e^{-ixt_0}}\right]. \tag{B.12}$$

Further, we express the principal value appearing in (B.12) in terms of an integral along contour $\mathcal{C}$ from $-\infty$ to $+\infty$ where the pole of the integrand at $x = 0$ is circumvented in the counterclockwise direction in the complex plane:

$$\operatorname{Vp}\int_{-\infty}^{\infty} \frac{dx}{ix + \gamma - \gamma e^{-ixt_0}} = -\frac{\pi}{1 + \gamma t_0} + \int_{\mathcal{C}} \frac{dx}{ix + \gamma - \gamma e^{-ixt_0}}. \tag{B.13}$$

After this, it remains only to calculate the integral along the contour $\mathcal{C}$. In order to do this, we notice that, upon elimination of the pole at $x = 0$, the inequality $\left|\gamma e^{-ixt_0}\right| < |ix + \gamma|$ holds everywhere on contour $\mathcal{C}$. Therefore, the integrand in (B.12) can be expanded in a geometric progression,

$$\frac{1}{ix + \gamma - \gamma e^{-ixt_0}} = \frac{1}{ix + \gamma} + \sum_{n=1}^{\infty} \frac{\gamma^n e^{-inxt_0}}{(ix + \gamma)^{n+1}}. \tag{B.14}$$

Upon integration along contour $\mathcal{C}$, the first term explicitly isolated in the expansion on the right-hand side

---

[7] It was indicated above that the quantity A_∞^2 is the right-hand side of relation (11).

[8] The idea of the method used here to calculate this integral came from V.P. Krainov.

of (B.14) yields a value of π, while the terms of the sum vanish since the corresponding functions decay exponentially in the lower half-plane of the complex plane and hence have no poles there. Taking this and relation (B.13) into account, we obtain

$$\mathfrak{Q}_3 = \pi - \frac{\pi}{1+\gamma t_0} = \frac{\pi\gamma t_0}{1+\gamma t_0}. \qquad (B.15)$$

Substituting $\mathfrak{Q}_1$ (B.9) and $\mathfrak{Q}_3$ (B.15) into Eq. (B.8), we then arrive at the final result for the sought limit; that is,

$$A_\infty = \frac{1}{1+\gamma t_0}, \qquad (B.16)$$

which is exactly twice as great as the limiting values (9) of amplitudes A_1 and A_2 in the case where the initial condition corresponds to the excitation of one of the atoms and is $\sqrt{2}$ times greater than $A_{1,2}$ in the case of the “symmetric” initial condition.

REFERENCES

1. V. Weisskopf and E. Wigner, Z. Phys. **63**, 54 (1930).
2. R. Loudon, *The Quantum Theory of Light* (Clarendon Press, Oxford, 1973; Mir, Moscow, 1976), Chap. 8.
3. N. F. Mott and I. N. Sneddon, *Wave Mechanics and Its Applications* (Dover, New York, 1963; Nauka, Moscow, 1966), Sect. 48.
4. V. P. Bykov and G. V. Shepelev, *Emission of Atoms in the Vicinity of Material Bodies* (Nauka, Moscow, 1986).
5. S. Haroche and J. M. Raimond, in *Cavity Quantum Electrodynamics*, Ed. by P. Berman (Academic, New York, 1994), p. 123.
6. E. Yablonovich, Phys. Rev. Lett. **58**, 2059 (1987).
7. S. John, Phys. Rev. Lett. **58**, 2486 (1987).
8. V. V. Klimov, M. Ducloy, and V. S. Letokhov, J. Mod. Opt. **43**, 2251 (1996).
9. V. V. Klimov and V. S. Letokhov, Phys. Rev. A **54**, 4408 (1996).
10. V. V. Klimov, M. Ducloy, and V. S. Letokhov, Kvantovaya Élektron. (Moscow) **31**, 569 (2001).
11. R. H. Dicke, Phys. Rev. **93**, 99 (1954).
12. M. S. Feld and J. C. MacGillivray, in *Coherent Nonlinear Optics*, Ed. by M. S. Feld and V. S. Letokhov (Springer, Berlin, 1980); Top. Curr. Phys. **21**, 7 (1980).
13. A. V. Andreev, V. I. Emel’yanov, and Yu. A. Il’inskiĭ, *Cooperative Phenomena in Optics* (Nauka, Moscow, 1988).
14. R. G. DeVoe and R. G. Brewer, Phys. Rev. Lett. **76**, 2049 (1996).
15. J. Eschner, Ch. Raab, F. Schmidt-Kaler, and R. Blatt, Nature **413**, 495 (2001).
16. U. Fano, Phys. Rev. **124**, 1866 (1961).
17. A. A. Makarov, V. T. Platonenko, and V. V. Tyakht, Zh. Éksp. Teor. Fiz. **75**, 2075 (1978) [Sov. Phys. JETP **48**, 1044 (1978)].
18. A. A. Makarov, in *Laser Spectroscopy of Vibrationally–Excited Molecules*, Ed. by V. S. Letokhov (Nauka, Moscow, 1980), Appendix 3.A, p. 77.
19. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media*, 2nd ed. (Nauka, Moscow, 1982; Pergamon Press, Oxford, 1984).
20. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 2nd ed. (Fizmatgiz, Moscow, 1963; Pergamon, New York, 1977).
21. J. C. Knight, T. A. Birks, P. St. J. Russel, and D. M. Atkin, Opt. Lett. **21**, 1547 (1996).
22. A. M. Zheltikov, Usp. Fiz. Nauk **170**, 1203 (2000) [Phys. Usp. **43**, 1125 (2000)].
23. M. J. Renn, E. A. Donley, E. A. Cornell, *et al.*, Phys. Rev. A **53**, R648 (1996).
24. V. S. Letokhov, Pis’ma Zh. Éksp. Teor. Fiz. **7**, 348 (1968) [JETP Lett. **7**, 272 (1968)].
25. V. S. Letokhov, in *Proceedings of Conference on Methods of Spectroscopy without Doppler Broadening of Excited Levels of Simple Molecules, Aussois, France, 1973* (CNRS, Paris, 1974), p. 127.
26. V. S. Letokhov, Laser Unconv. Opt., No. 43, 3 (1973).
27. P. S. Jessen, C. Gerz, P. D. Lett, *et al.*, Phys. Rev. Lett. **69**, 49 (1992).
28. V. I. Balykin, V. G. Minogin, and V. S. Letokhov, Rep. Prog. Phys. **63**, 1429 (2000).
29. V. I. Balykin, V. V. Klimov, and V. S. Letokhov, Pis’ma Zh. Éksp. Teor. Fiz. **78**, 11 (2003) [JETP Lett. **78**, 8 (2003)].
30. I. V. Bagratin, B. A. Grishanin, and V. N. Zadkov, Usp. Fiz. Nauk **171**, 625 (2001) [Phys. Usp. **44**, 597 (2001)].
31. C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Atom–Photon Interactions: Basic Processes and Applications* (Wiley, New York, 1992), p. 64.
32. M. A. Rippin and P. L. Knight, J. Mod. Opt. **43**, 807 (1996).
33. V. S. Vladimirov, *Equations of Mathematical Physics* (Nauka, Moscow, 1967; Marcel Dekker, New York, 1971), Chap. 2.

Translated by A. Isaakyan

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 702–710.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 781–791.
Original Russian Text Copyright © 2003 by Volkova, Popov, Tikhonova.

ATOMS, SPECTRA, RADIATION

Dynamics of Electronic and Nuclear Motion in a Molecular Hydrogen Ion in a Strong Laser Field

E. A. Volkova, A. M. Popov*, and O. V. Tikhonova
Skobel'tsyn Nuclear Physics Research Institute, Moscow State University,
Vorob'evy gory, Moscow, 119992 Russia
**e-mail: popov@mics.msu.su*

Received April 10, 2003

Abstract—The dynamics of the electron and ion subsystems of a molecular hydrogen ion in the field of femtosecond titanium–sapphire laser radiation is investigated by using numerical simulation. The dependence of the probability of ionization and excitation of various electron states of H_2^+ on the radiation intensity and on the orientation of the molecule axis relative to the direction of polarization of the electric field of the wave is investigated. It is shown that the population of all electronic states of the molecule (except the ground state and the first excited state) is negligibly low in a wide range of radiation intensity. In such a situation, if the probability of system ionization is also low, the dynamics of vibration–rotation motion of nuclei in the molecule can be considered in the approximation of two potential surfaces (energy levels). The possibility of the effective orientation of the molecular axis along the direction of the electric field of the wave in the absence of dissociation of the system is considered. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The development of high-power femtosecond lasers with a pulse duration comparable to the characteristic time scale of nuclear motion in a molecule paves new ways for controlling chemical reactions and for studying their mechanisms [1]. One of the main processes, which have been studied intensely in recent years in this respect, is the possibility of aligning (orientation) of the molecular axis in a given direction in the field of a high-power laser pulse [2–10]. The possibility of such a process was repeatedly discussed both from the experimental [3–6] and the theoretical [7–9] points of view. However, it is important for practical applications that the "oriented" molecule remains in the bound and not the dissociated state. At the same time, in the experiments that have been reported so far [5, 10], the fact of orientation of molecules is monitored, as a rule, from the angular distribution of the dissociation products formed as a result of ionization of the molecular system. In addition, the physics of the process is not quite clear in some cases: the molecule is orientated in a strong laser field and then undergoes ionization and dissociation, or the observed angular dependence of scattering of the ionization products stems from the fact that only the molecules oriented in a certain direction (in the direction of the polarization vector of the electric field of the wave) are predominantly ionized [4].

It is therefore clear that an analysis of the dynamics of the electron and nuclear motion in molecules is a problem of interest. Since such a problem presumes the analysis of a multidimensional Schrödinger equation even for the simplest molecules, it is important to analyze the possibility of independent treatment of the evolution of the electronic and nuclear subsystems in a laser field. Such a separation of the electronic and nuclear degrees of freedom in a molecule forms the basis of the Born–Oppenheimer approximation and is a consequence of different temporal and energy scales of motion of the electron and nuclear subsystems. However, in strong electromagnetic fields, within the Born–Oppenheimer approximation, a large number of electron energy levels (including, above all, those describing the ionization state of the system) must be taken into account under the conditions of strong ionization and considerable rearrangement of the structure of the electron states of the molecule by the optical field. Such a problem is difficult for simulation. On the other hand, it is undoubtedly interesting to study the vibration–rotation motion of a molecule in a high-intensity electromagnetic field in the case when the probability of ionization of the molecule is negligibly low; hence, it is sufficient to take into account a small number of electron energy levels in analyzing the molecular dynamics. This paper is devoted to a numerical analysis of this problem for a model two-dimensional (planar) molecular hydrogen ion in the field of Ti:Sa laser radiation (the energy of a quantum is $\hbar\omega = 1.55$ eV). The conditions permitting the analysis of vibration–rotation dynamics of H_2^+ in the two-term approximation are determined and the possibility of the effective alignment of the molecular axis along the polarization direction of the electric field of the wave in the absence of dissociation is demonstrated.

2. MODEL OF A MOLECULAR HYDROGEN ION

It is well known that the total wave function of a molecule in the adiabatic approximation can be represented in the form of an expansion in electron energy levels:

$$\Psi(\mathbf{r}, \mathbf{R}, t) = \sum \Phi_n(\mathbf{R}, t)\varphi_n(\mathbf{r}, \mathbf{R}). \tag{1}$$

Here, **r** and **R** are the sets of coordinates of the electronic and nuclear subsystems; Φ_n is the nuclear wave function describing the vibration–rotation motion of the nuclear subsystem, corresponding to the *n*th electron energy level; and $\varphi_n(\mathbf{r}, \mathbf{R})$ is the electron wave function determining the structure of the *n*th energy level. In the absence of an external action, the motion occurs independently and an electromagnetic field induces transitions between the energy levels. In this case, summation in Eq. (1) also includes the integral over the continuum of the electron states corresponding to the ionized state of the system. The stationary states $\varphi_n(\mathbf{r}, \mathbf{R})$ can be determined from the solution of the eigenvalue problem for the Hamiltonian of the electronic subsystem of the molecule,

$$\hat{H}_e(\mathbf{r}, R)\varphi_n(\mathbf{r}, R) = E_n(R)\varphi_n(\mathbf{r}, R). \tag{2}$$

Here, $E_n(R)$ is the energy of the stationary state, corresponding to function $\varphi_n(\mathbf{r}, R)$.

In the two-dimensional approximation, the Hamiltonian of the electronic subsystem of the molecular ion H_2^+ can be written in the form

$$\hat{H}_e = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) - \frac{e^2}{\sqrt{(x - R/2)^2 + y^2 + \alpha^2}} - \frac{e^2}{\sqrt{(x + R/2)^2 + y^2 + \alpha^2}}. \tag{3}$$

Here, it is assumed that the *x* axis is directed along the molecular axis, *R* is the distance between the nuclei, and α is the smoothing parameter selected so that the characteristics of our model system are as close as possible to the experimentally observed characteristics.[1]

The numerical solution of eigenvalue problem (2) for Hamiltonian (3) was used to determine the wave functions $\varphi_n(x, y, R)$ of the stationary states of the electronic subsystem, which depend parametrically on their nuclear spacing *R*, and to construct the system of electron energy levels of the molecule,

$$V_{\mathrm{eff}}^{(n)}(R) = \frac{e^2}{R} + E_n(R), \tag{4}$$

where *n* is the number of the electronic state. In our calculations, we analyzed 12 electronic states ($n = 1, \ldots, 12$). The potential curves for two lower states obtained with the help of Eq. (4) are shown in Fig. 1. It can be seen from the figure that the equilibrium distance between the nuclei in our model is $R = R_0 = 1.2$ Å. The energies of several first electron states calculated for $R = R_0$ are given in Table 1, and the electron density distribution for the ground and the first excited states for $R = R_0$ and $R = 2$ Å are shown in Figs. 2 and 3. It should be noted that, in contrast to the structure of the excited state, the electron density distribution $|\varphi_1(x, y, R)|^2$ is essentially different for the two indicated values of *R*. For example, for $R = R_0$, the distribution $|\varphi_1(x, y, R_0)|^2$ is approximately isotropic and the peak of the electron density is in the middle of the spacing between the nuclei. An increase in the nuclear spacing to 2 Å leads to the emergence of a dichotomic structure of the wave function with the electron density peaks localized approximately in the region where the nuclei are located. As a result, an increase in the nuclear spacing results in anisotropy in the electron density, leading to peculiarities in the dynamics of the elec-

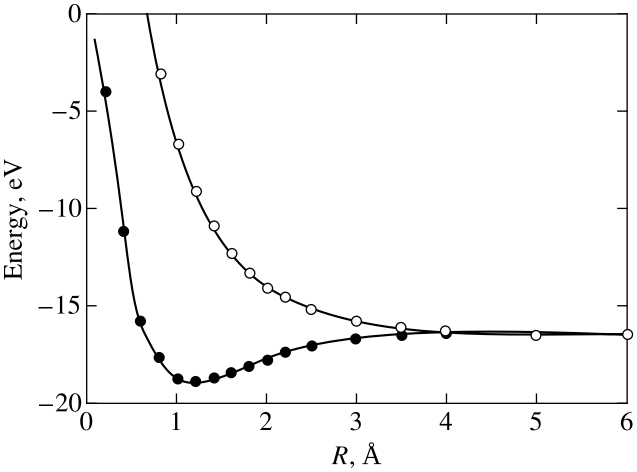

Fig. 1. Potential curves corresponding to two lower electronic states of the two-dimensional H_2^+ ion.

Table 1. Energies of the lower electronic states of the model H_2^+ ion for the equilibrium value of the nuclear spacing $R_0 = 1.2$ Å

State	Energy, eV
$\|1\rangle$	−29.8
$\|2\rangle$	−19.7
$\|3\rangle$	−14.7
$\|4\rangle$	−10.7
$\|5\rangle$	−8.4

---

[1] In our calculations, it is assumed that $\alpha = 0.35$ Å.

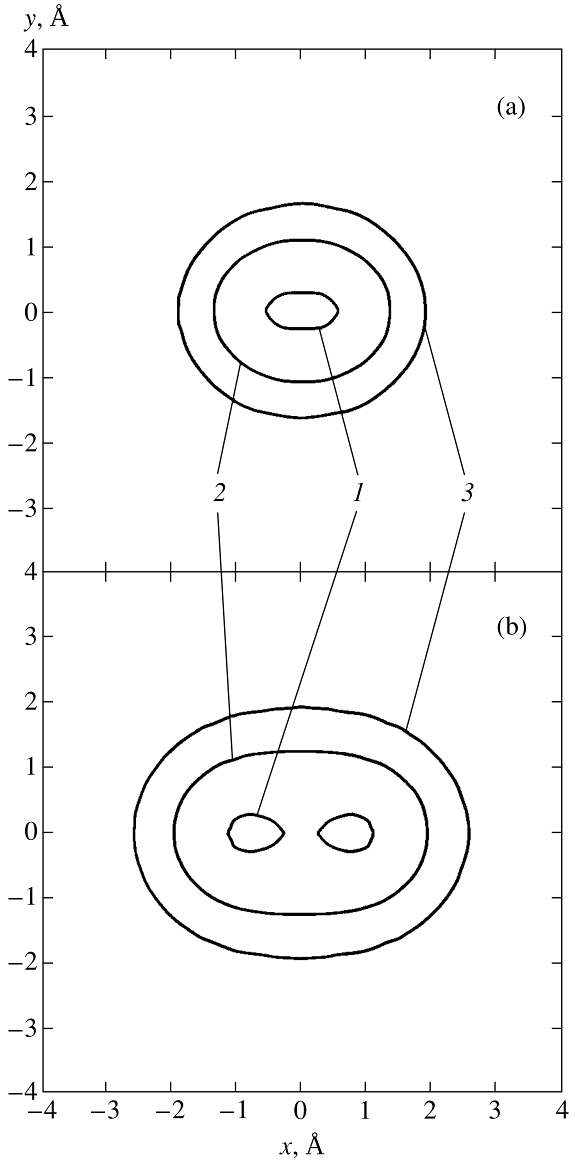

Fig. 2. Electron density distribution in the ground state of the molecule for the nuclear spacing $R = 1.2$ (a) and 2.0 Å (b). The level lines (in relative units) 1 (*1*), 0.1 (*2*), and 0.01 (*3*).

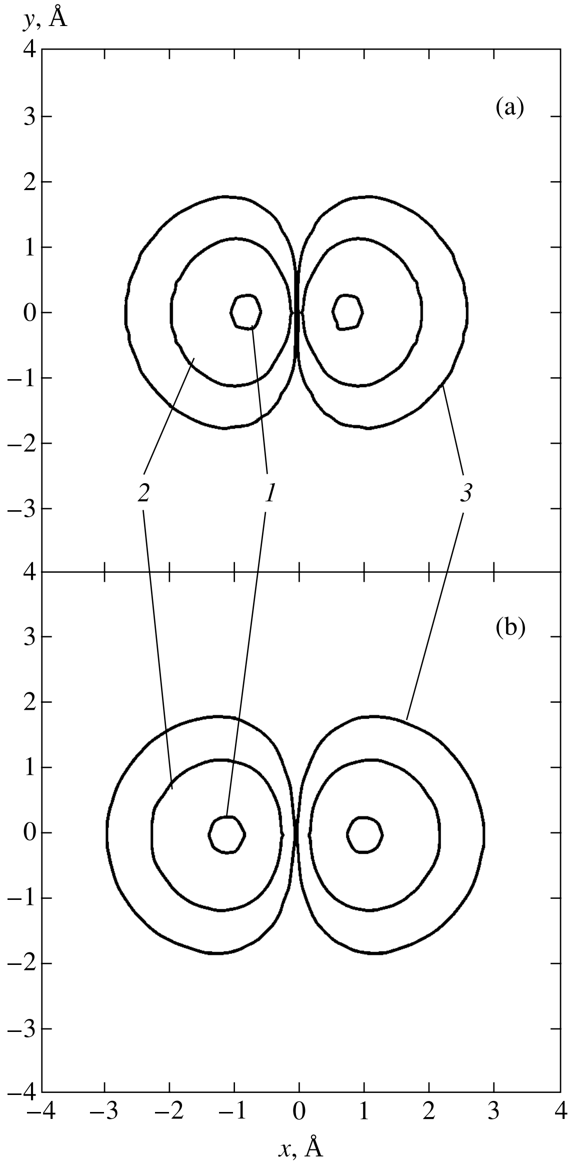

Fig. 3. Electron density distribution in the first excited state of the molecule for the nuclear spacing $R = 1.2$ (a) and 2.0 Å (b). The level lines (in relative units) 1 (*1*), 0.1 (*2*), and 0.01 (*3*).

tronic and nuclear motion depending on the angle of orientation of the molecular axis relative to the direction of the electric field of the wave.

3. IONIZATION OF THE H_2^+ ION IN A STRONG FIELD

It is usually assumed that the evolution of the electronic subsystem occurs on a much smaller time scale as compared to the nuclear subsystem. For this reason, in the first approximation, the process of ionization of the molecule in a short pulse can apparently be considered in the approximation of “frozen” nuclei, i.e., by fixing the nuclear spacing and the orientation of the molecular axis relative to the direction of action of the electric field of the wave. In this case, assuming that the x axis is directed along the molecular axis, we can write the following expression for the wave function of the electronic subsystem:

$$i\hbar\frac{\partial\psi_e(x, y, R, t)}{\partial t} = (\hat{H}_e - e(x\cos\theta + y\sin\theta)\varepsilon(t)\cos(\omega t))\psi_e(x, y, R, t). \quad (5)$$

Here, $\varepsilon(t)$ is the envelope of a laser pulse and θ is the angle between the direction of the electric field and the molecular axis.

In our calculations of the dynamics of the electronic subsystem of the molecule, we assumed that the pulse has a smoothed trapezoidal shape (see [11]) with front durations of $\tau_f = 2T$ and a generation plateau duration of $\tau_{pl} = 5T$ ($T = 2\pi/\omega$ is the duration of an optical cycle). We also assumed that the system at the initial instant is in the ground electronic state. Solving Eq. (5), we can determine the probabilities of ionization and excitation of various states of the system depending on the laser pulse parameters, as well as the orientation of the molecular axis relative to the direction of polarization of the electric field of the wave and the nuclear spacing.

Equation (5) was solved numerically by the finite element method in the Cartesian frame in a region $|x, y| \le L$, $L = 40$ Å in size. In order to prevent the reflection of the probability flux from the boundaries of the spatial region, we introduced the electron density absorption in the boundary region, which ensured the effective "erosion" of the wave function.

The probabilities of various electronic states being excited were determined as projections of calculated function $\psi_e(x, y, t)$ onto wave functions $\varphi_n(x, y)$ of stationary states of the electronic subsystem at an instant corresponding to the end of the laser pulse ($t = t_{\text{fin}}$):

$$W_n = \left|\langle \psi_e(x, y, t_{\text{fin}}) | \varphi_n(x, y) \rangle\right|^2. \tag{6}$$

The probability of ionization was calculated using several methods. First, this quantity was calculated as

$$W_i = 1 - \sum_n W_n, \tag{7}$$

where 12 different electronic bound states of the molecule were taken into account. Second, it is natural to assume that the bound states of the molecule correspond to the electron localization in the region of small values of $r = \sqrt{x^2 + y^2}$. Consequently, the probability of finding the system in a bound state can be determined by integrating the electron density over the region

$$r = \sqrt{x^2 + y^2} \le r_0.$$

Accordingly, for the ionization probability, we have

$$W_i = 1 - \int_{r \le r_0} |\psi_e(x, y)|^2 dxdy. \tag{8}$$

Finally, integrating the electron density flux absorbed in the vicinity of the boundaries of the spatial computation region with respect to time, we can also obtain the ionization probability for the system. In the latter case, it is very important that computations continue after the action of the laser pulse in view of a time delay in the electron density absorption near the boundaries of the computation region relative to the ionization of the system during the action of the pulse. Obviously, this delay is determined by the time of electron motion within the spatial computation region and depends on the size L of this region and the characteristic velocity of a photoelectron in the continuum.

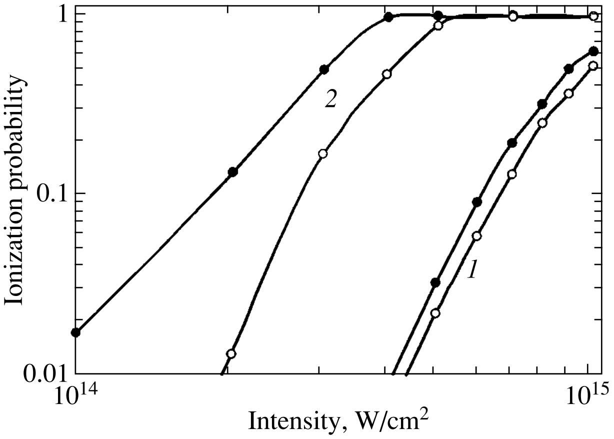

Fig. 4. Dependences of the ionization probability for H_2^+ on the radiation intensity in a laser pulse with $\tau_f = 2T$ and $\tau_{pl} = 5T$ for nuclear spacing $R = 1.2$ (*1*) and 2.0 Å (*2*). Dark and light circles correspond to the longitudinal and transverse orientations of the molecular axis relative to the polarization vector of the electric field of the wave.

Special calculations proved that all these methods for determining the ionization probability lead to identical results; for the size of the central zone, it is convenient to take $r_0 = 5$ Å. Parameter r_0 corresponds to the region of localization of approximately ten lower electronic states of the molecule.

The main results of calculating the electron dynamics of the molecular H_2^+ ion are shown in Fig. 4. The figure gives the dependences of the ionization probability for H_2^+ on the radiation intensity, calculated for two values of the nuclear spacing as well as for longitudinal and transverse orientation of the molecular axis relative to the field polarization vector. First of all, we must note a sharp increase in the ionization probability (especially in the case of relatively weak fields) upon an increase in the nuclear spacing. This effect was discovered in [12] and was repeatedly discussed in the literature [13, 14]. On the other hand, for $R = R_0$, the ionization probability is found to be negligibly small in the region of $P \le 4 \times 10^{14}$ W/cm^2. It should also be noted that in the given case the ionization probability weakly depends on the orientation of the molecule relative to the direction of polarization of the electric field of the wave, especially for an equilibrium nuclear spacing. The corresponding dependences for specific values of the intensity are shown in Fig. 5. We believe that such a smooth dependence is associated with a high degree of

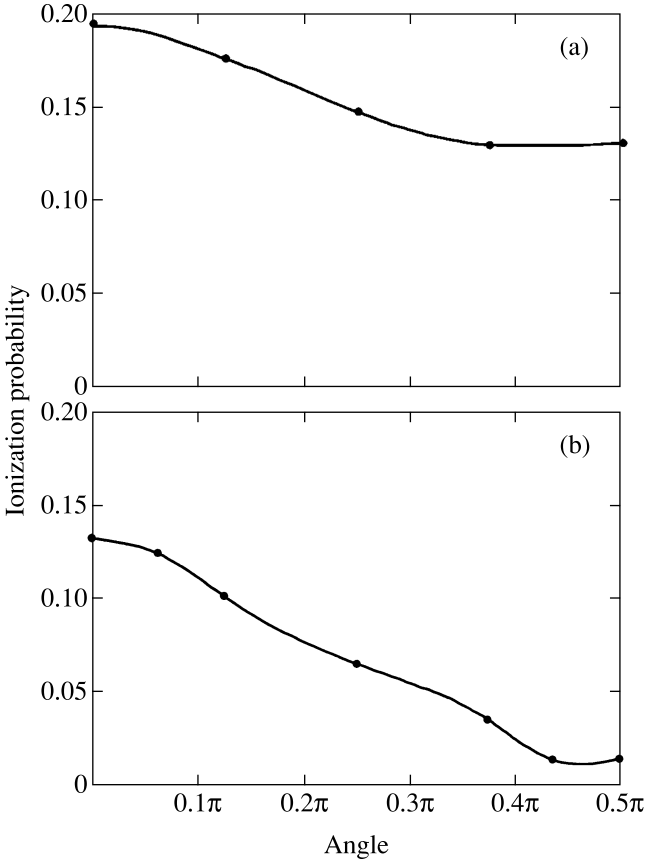

Fig. 5. Dependences of the ionization probability for H_2^+ on the orientation of the molecular axis relative to the polarization vector of the electric field of the wave for the following values of parameters: (a) $R = 1.2$ Å, $P = 7 \times 10^{14}$ W/cm^2; (b) $R = 2.0$ Å, $P = 2 \times 10^{14}$ W/cm^2.

spatial symmetry of the initial state (see Fig. 2) and with the absence of a preferred direction in the system for $R = R_0$. An increase in the nuclear spacing leads to symmetry breaking (see Fig. 3) and, hence, to a stronger angular dependence of the ionization probability (see Fig. 5). It should be noted that the angular dependence of ionization $W_i(\theta)$ of a molecular system was analyzed both analytically [15, 16] and numerically [17] in various limiting cases. The results of our computations are closest qualitatively to the data obtained in [17]. In the latter study, the angular dependence of the ionization probability for three-dimensional H_2^+ was studied in the tunneling mode by numerical simulation. In the approximation of molecular energy levels rearranged in the field, the ionization probabilities per unit time of H_2^+ were calculated for the two lower *u* and *g* states. It was found that, in contrast to model [15], the dependence of the ionization probability on the angle between the axis of the molecular ion and the polarization of radiation is weak (the ionization rate for the lower rearranged *g* state is higher for the molecular orientation along the electric field of the wave for nuclear spacings that do not exceed four Bohr radii). This result is in good qualitative agreement with the data obtained for $R \approx R_0 = 1.06$ Å and shown in Fig. 4 (curve *1*). It should be noted that a direct comparison of our results with the data on the ionization rates for rearranged states is not physically justified since we determined the ionization probability per pulse, while the population of various "dressed" states of the molecule was not investigated.

Let us now consider the results of calculating the probability of excitation for various electronic states of the molecule. It was found that the probability of population of all electronic states except the lower excited $|n = 2\rangle$ state is negligibly low in the entire investigated range of parameters and does not exceed 3×10^{-3}. Since the Keldysh parameter $\gamma \le 1$ under the conditions of our computations in the region of $P \ge 3 \times 10^{14}$ W/cm^2—i.e., we are talking about the tunneling mechanism of H_2^+ ionization—such a small probability of molecular excitation appears natural.

The results are of fundamental importance for constructing a model describing the motion of the nuclear subsystem of the molecule. Indeed, our calculations show that the processes of ionization of the molecule and the excitation of all its states except the lower excited state $|n = 2\rangle$ can be neglected in the range of intensity values $P \le 2 \times 10^{14}$ W/cm^2. This means that the dynamics of the nuclear subsystem can be considered within an approximation that takes into account only two lower energy levels of the molecule.

4. EVOLUTION OF THE NUCLEAR SUBSYSTEM OF H_2^+ IN A LASER FIELD

Within the two-term approximation, the total wave function of a molecular hydrogen ion can be written in form (1), the sum containing only two addends:

$$\begin{aligned}\Psi(\mathbf{r}, \mathbf{R}, t) &= \Phi_1(R, \theta, t)\varphi_1(x, y, R) \\ &+ \Phi_2(R, \theta, t)\varphi_1(x, y, R).\end{aligned} \tag{9}$$

Here, $\mathbf{r} = (x, y)$ and $\mathbf{R} = (R, \theta)$ is the set of coordinates of the electronic and nuclear subsystems, respectively, and subscripts "1" and "2" correspond to the ground and the first excited electronic states of the molecule. It was noted above that the action of an external electromagnetic field induces transitions between the energy levels, the transition probability being determined by the matrix element of dipole operator $\mathbf{r}_{12}$. Taking into account relation (9), we find that the system of equations describing the vibration–rotation dynamics of the molecule under the action of an external electromagnetic field has the form

$$\begin{aligned} i\hbar\frac{\partial\Phi_1(R,\theta,t)}{\partial t} &= [\hat{T}_N + V_{\text{eff}}^{(1)}(R)]\Phi_1(R,\theta,t) \\ &- e\varepsilon(t)\cos\omega t[x_{12}(R)\cos(\theta) + y_{12}(R)\sin\theta]\Phi_2(R,\theta,t), \\ i\hbar\frac{\partial\Phi_2(R,\theta,t)}{\partial t} &= [\hat{T}_N + V_{\text{eff}}^{(2)}(R)]\Phi_2(R,\theta,t) \\ &- e\varepsilon(t)\cos\omega t[x_{12}(R)\cos(\theta) + y_{12}(R)\sin\theta]\Phi_1(R,\theta,t). \end{aligned} \tag{10}$$

Here,

$$\hat{T}_N = -\frac{\hbar^2}{2\mu}\left\{\frac{1}{R}\frac{\partial}{\partial R}R\frac{\partial}{\partial R} + \frac{1}{R^2}\frac{\partial^2}{\partial\theta^2}\right\}$$

is the kinetic energy operator of nuclear motion (the first and second terms correspond to the energy of vibrational and rotational motion, respectively), μ is the reduced mass of the molecule, and $V_{\text{eff}}^{(1,2)}$ is defined by formula (4). The quantities ex_{12} and ey_{12} are the matrix elements of the projections of the dipole operator onto the molecular axis and on the direction perpendicular to the axis. It was noted above that the transverse component $y_{12}(R) \equiv 0$ on account of the symmetry of the wave functions of the two lower electronic states; the dependence of quantity x_{12} on the nuclear spacing is shown in Fig. 6. For the ground vibrational state, the transition probability is determined by the value of $x_{12}(R_0) \approx 0.65$ Å. In solving system of equations (10) numerically, we assumed that the molecule was on the lower energy level in the ground vibration–rotation state at the initial instant.

In order to analyze the results of numerical integration of system (10), we calculated the wave functions of stationary vibration–rotation nuclear states, satisfying the equation

$$[\hat{T}_N + V_{\text{eff}}^{(1)}(R)]\phi_{vm}(R,\theta) = E_{vm}\phi_{vm}(R,\theta). \tag{11}$$

Here, v and m are the vibrational and rotational quantum numbers. It should be noted that the nuclear function of the stationary state can be represented in the form

$$\phi_{vm}(R,\theta) = \chi_{vm}(R)\frac{\exp(im\theta)}{\sqrt{2\pi}} \tag{12}$$

and function $\chi_{vm}(R)$ of the vibrational motion can be determined from the solution of the eigenvalue problem

$$\begin{aligned} &-\frac{\hbar^2}{2\mu}\frac{1}{R}\frac{\partial}{\partial R}R\frac{\partial\chi_{vm}}{\partial R} \\ &+ \left\{V_{\text{eff}}^{(1)}(R) + \frac{\hbar^2 m^2}{2\mu R^2}\right\}\chi_{vm} = E_{vm}\chi_{vm}. \end{aligned} \tag{13}$$

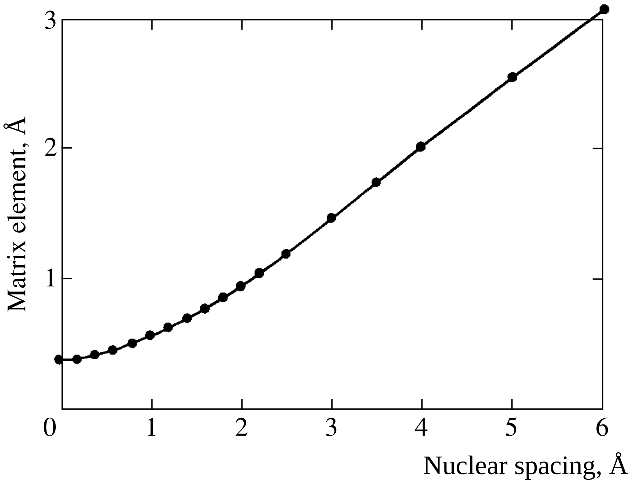

Fig. 6. Matrix element of the projection of the electron dipole moment of the transition onto the molecular axis x_{12} as a function of the nuclear spacing.

Table 2 contains some of the results of calculation of the vibration–rotation spectrum of the nuclear subsystem, derived from formula (13). It can be seen from the data represented in the table that, for small values of the vibrational and rotational quantum numbers, the obtained spectrum is successfully approximated by the expression

$$\begin{gathered} E_{vm} = \hbar\Omega(v + 1/2) + Bm^2, \\ v = 0, 1, 2, \ldots, \quad m = 0, \pm 1, \pm 2, \ldots. \end{gathered} \tag{14}$$

Here, $\hbar\Omega = 0.255$ eV is the vibrational quantum corresponding to the transition

$$|v = 0, m = 0\rangle \longrightarrow |v = 1, m = 0\rangle,$$

and $B = \hbar^2/2\mu R_0^2$ is the rotational constant calculated for the equilibrium nuclear spacing. It should be noted that, by virtue of the condition $B \ll \hbar\Omega$, which holds in

Table 2. Energies of vibration–rotation energy levels of the model H_2^+ ion, belonging to the ground electron energy level

v	m			
	0	1	2	8
0	−2.267	−2.264	−2.256	−2.087
1	−2.012	−2.010	−2.000	−1.843
2	−1.763	−1.760	−1.752	−1.601
3	−1.529	−1.526	−1.519	−1.3741
4	−1.309	−1.307	−1.299	−1.161
5	−1.103	−1.098	−1.094	−0.960
6	−0.911	−0.908	−0.902	−0.773

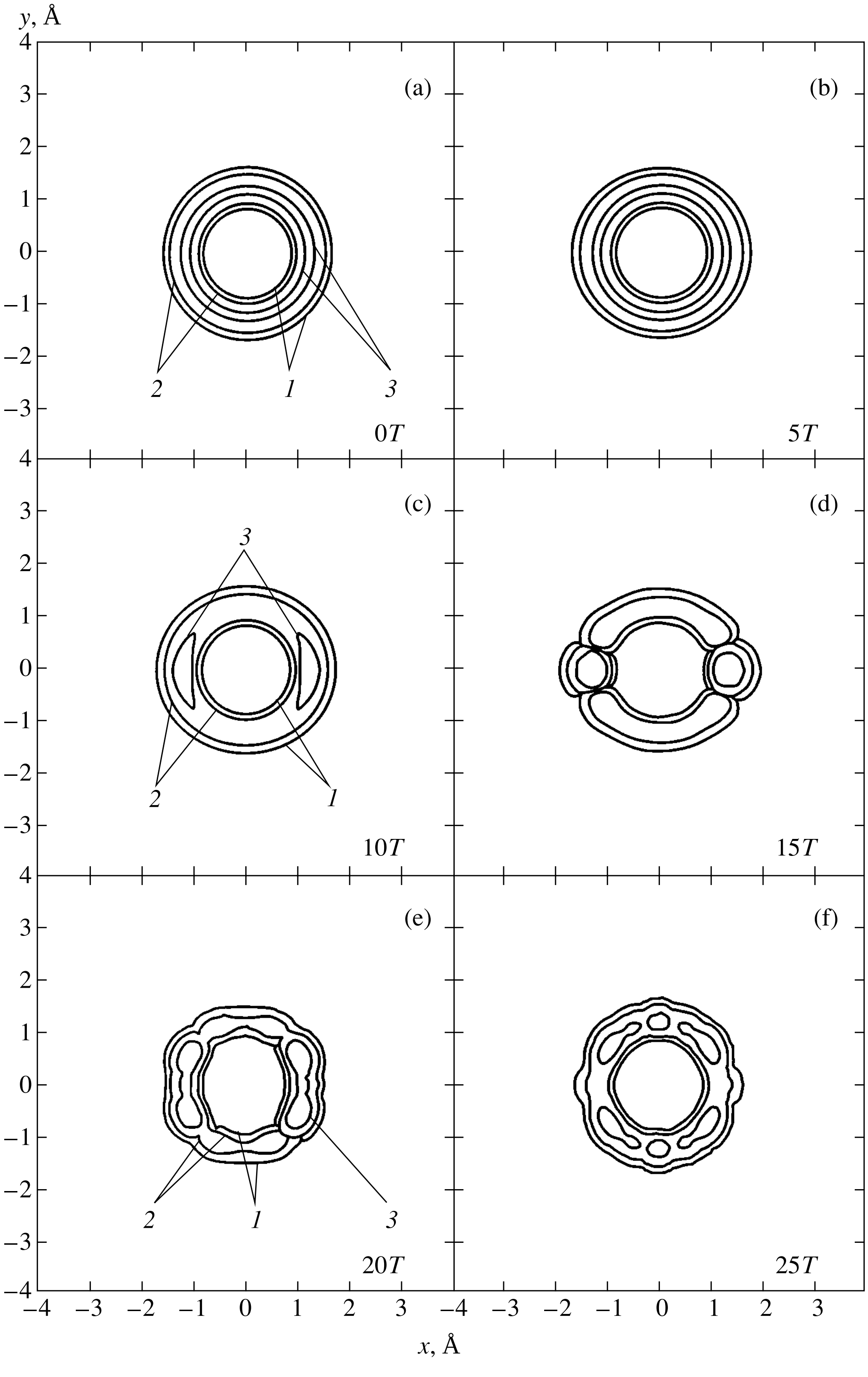

Fig. 7. Dynamics of the nuclear wave packet in the field of laser radiation with an intensity of $P = 10^{14}$ W/cm^2 and parameters of $\tau_f = 5T$ and $\tau_{pl} = 10T$. The isolines of probability density (in relative units): 1 (*1*), 0.1 (*2*), and 0.01 (*3*).

this case, the effect that rotational motion of the molecule has on the value of the vibrational quantum can be neglected.

The data contained in Table 3 show that the main parameters for the two-dimensional H_2^+ ion are close to the experimentally measured parameters of the real three-dimensional molecular ion.

Let us now consider the results of calculating the vibration–rotation dynamics of the nuclear motion, which were obtained using the two-term approximation. The main part of the computations were performed for the initial nuclear state $|v = 0, m = 0\rangle$. A typical example of the evolution of the nuclear probability density calculated for a pulse with $\tau_f = 5T$, $\tau_{pl} = 10T$, and $P = 10^{14}$ W/cm^2 is shown in Fig. 7. It can be seen that the initial state is completely delocalized in the angle and the nuclear packet is localized as a result of laser pulse action predominantly in the direction of the polarization vector of the electric field of the wave, which is changed at the rear front of the laser pulse by a complex rotational dynamics emerging as a result of the population of a large number of rotational states of the lower electronic state of the molecule. The structure of the nuclear wave packet represented in Fig. 7 indicates that the region of packet localization practically does not change along the radius. This means that the molecule remains predominantly in the ground vibrational state. Since the ground vibrational state is localized along the radial coordinate R in the region

$$\Delta R \approx \sqrt{\hbar/\mu\Omega} \approx 0.18\ \text{Å} \ll R_0,$$

the effect of the increase in the ionization rate with increasing R can be disregarded. This justifies once again the application of the two-term approximation under the given conditions. The degree of orientation of the molecule relative to the field direction ε can be conveniently characterized by the value of the squared cosine of angle θ averaged over the quantum state [11]. We calculated this value using the formula

$$\langle \cos^2\theta \rangle = \int \cos^2\theta \times \{|\Phi_1(R, \theta, t)|^2 + |\Phi_2(R, \theta, t)|^2\} R dR d\theta, \tag{15}$$

where Φ_1 and Φ_2 are the solutions of system (10). For the above parameters, the dynamics of $\langle\cos^2\theta\rangle$ represented in Fig. 8 demonstrates a considerable "alignment" of the system by the end of the "plateau" of the laser pulse. The subsequent "disorientation" of the state is associated, to a considerable extent, with a decrease in the radiation intensity at the rear front of the pulse. The ability of the laser field to confine molecules in the aligned state for a long time obviously follows from the data presented in Fig. 9. In this case, the plateau duration is $\tau_{pl} = 50T$; during this time, the molecule remains aligned predominantly along the direction of the wave field, although the extent of orientation changes with time. It should be emphasized that the degree of the electronic and vibrational excitations of the molecules is small. This makes it possible in principle to control the spatial orientation of the molecule without changing its electronic and vibrational states, which is quite important for modern laser femtochemistry.

An increase in the laser radiation intensity to $P \geq 2 \times 10^{14}$ W/cm^2 changes the pattern of the process significantly. In this case, a considerable population of the vibrational continuum of the lower energy level takes place, leading to disintegration of the molecule. The corresponding data on the probability of finding the molecule at the lower electron energy level (including the bound level) and not in the dissociated state depend-

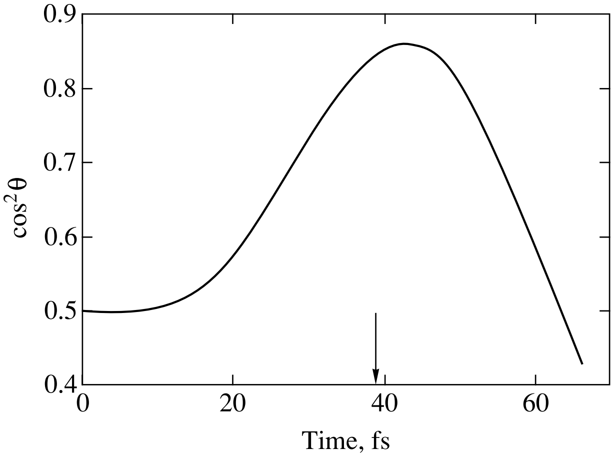

Fig. 8. Mean value of the squared cosine of the angle between the molecular axis and the polarization vector of the electric field of the wave as a function of time for a pulse with $\tau_f = 5T$, $\tau_{pl} = 10T$, and $P = 10^{14}$ W/cm^2. The arrow indicates the end of the plateau of the laser pulse.

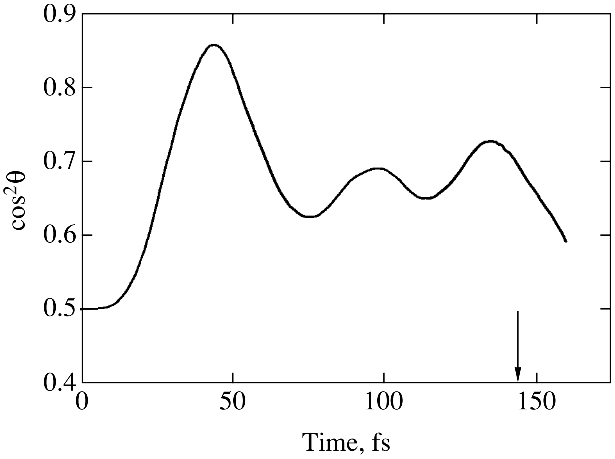

Fig. 9. The same as in Fig. 8 for $\tau_{pl} = 50T$.

Table 3. Comparison of the parameters of the model (two-dimensional) and real (three-dimensional) H_2^+ ions

	2*D*	3*D*
Equilibrium nuclear spacing, R_0, Å	1.2	1.06
Dissociation energy, eV	2.3	2.64
Vibrational quantum, $\hbar\Omega$, eV	0.25	0.28
Energy of transition between the lower electron energy levels for $R = R_0$, eV	10	11.5
Ionization potential, eV	29.8	29.9

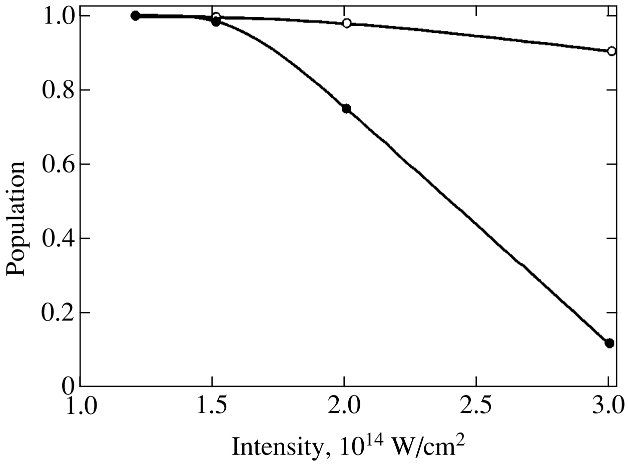

Fig. 10. Dependence of the probability of finding the system in the ground electronic state (light circles) and also in non-dissociated state (dark circles) after the end of a laser pulse on the radiation intensity for a pulse with $\tau_f = 5T$ and $\tau_{pl} = 10T$.

ing on the radiation intensity are given in Fig. 10. It can be seen that, for the given laser pulse parameters ($\tau_f = 5T$ and $\tau_{pl} = 10T$), the possibility of controlling the spatial orientation of the molecular axis without its dissociation is limited to intensities of $P \leq 1.5 \times 10^{14}$ W/cm^2.

5. CONCLUSIONS

Thus, we have studied the dynamics of the electronic and nuclear subsystems of the simplest molecule H_2^+ in a strong laser field. It is shown that, for an equilibrium value of the nuclear spacing, the ionization probability of the system is a smooth function of the angular orientation of the molecular axis relative to the direction of polarization of the electric field of the wave. In this case, even in the intense ionization regime ($W_i \geq 0.1$), the probabilities of excitation of all electronic states of the molecule (except the first excited state) are negligibly low. This makes it possible to analyze the vibration–rotation dynamics of the molecule using the two-term approximation in the range of fields for which the ionization probability is low. The application of this approximation has made it possible to determine the range of radiation parameters in which the effective action on the rotational states of the molecule not accompanied by its vibrational excitation or dissociation is possible. It should also be noted that an increase in the radiation intensity to values corresponding to considerable ionization of the molecule extremely complicates an analysis of the vibration–rotation dynamics in the framework of the adiabatic approximation since it is necessary to take into account a large number of electronic states of the molecule in the continuum. In such a situation, it is obviously more expedient to directly solve the time-dependent Schrödinger equation for the wave function of the molecule outside the scope of the Born–Oppenheimer approximation.

ACKNOWLEDGMENTS

The authors are grateful to C. Cornaggia and L. Roso for fruitful discussions.

This study was supported by the Russian Foundation for Basic Research (project no. 03-02-16234) and INTAS (grant no. 99-1495).

REFERENCES

1. V. A. Cho and R. B. Bernstein, J. Phys. Chem. **95**, 8129 (1991).
2. P. B. Corkum, C. E. Ellert, and M. Mehendale, Faraday Discuss. **113**, 47 (1999).
3. L. J. Frasinski, K. Codling, P. A. Hatherly, *et al.*, Phys. Rev. Lett. **58**, 2424 (1987).
4. K. Codling, L. J. Frasinski, and P. A. Hatherly, J. Phys. B **22**, L321 (1989).
5. Ch. Ellert and P. B. Corkum, Phys. Rev. A **59**, R3170 (1999).
6. C. Cornaggia, J. Lavancier, D. Normand, *et al.*, Phys. Rev. A **44**, 4499 (1991); L. Quaglia, M. Brewczyk, and C. Cornaggia, Phys. Rev. A **65**, 031404(R) (2002).
7. E. Charron, A. Giusti-Suzor, and F. H. Mies, Phys. Rev. Lett. **75**, 2815 (1995).
8. T. Seideman, J. Chem. Phys. **103**, 7887 (1995).
9. A. I. Andryushin and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **116**, 1551 (1999) [JETP **89**, 837 (1999)].
10. M. Tsubouchi, B. Whitaker, L. Wang, *et al.*, Phys. Rev. Lett. **86**, 4500 (2001).
11. A. M. Popov, O. V. Tikhonova, and E. A. Volkova, Laser Phys. **13**, 435 (2003).
12. T. Zuo and A. D. Bandrauk, Phys. Rev. A **52**, R2511 (1995); S. Chelkowski and A. D. Bandrauk, J. Phys. B **28**, L723 (1995).
13. T. Seideman, M. Yu. Ivanov, and P. B. Corkum, Phys. Rev. Lett. **75**, 2819 (1995).
14. A. M. Popov, O. V. Tikhonova, and E. A. Volkova, Laser Phys. **7**, 844 (1997).
15. N. B. Delone, V. P. Kraĭnov, and M. A. Sukharev, Tr. Inst. Obshch. Fiz., Ross. Akad. Nauk **57**, 32 (2000).
16. A. A. Kudrin and V. P. Krainov, Laser Phys. **13**, 1024 (2003).
17. M. Plummer and J. F. McCann, J. Phys. B **30**, L401 (1997).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 711–721.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 792–802.
Original Russian Text Copyright © 2003 by Arkhipkin, Myslivets, Timofeev.

ATOMS, SPECTRA, RADIATION

Stark-Chirped Rapid Adiabatic Passage: Propagation of Laser Pulses and Spacetime Evolution of Populations and of Two-Photon Coherence

V. G. Arkhipkin*, S. A. Myslivets, and I. V. Timofeev
Kirensky Institute of Physics, Russian Academy of Sciences, Akademgorodok, Krasnoyarsk, 660036 Russia
**e-mail: avg@iph.krasn.ru*

Received April 11, 2003

Abstract—On the basis of a vector model, the propagation of laser pulses under the conditions of a two-photon quasiresonance in the case of Stark-chirped rapid adiabatic passage through the resonance is studied with allowance for a diabatic character of the interaction. It is shown that the shape of a pulse propagating in a medium changes, the sweeping of its carrier frequency occurring concurrently. Special features of the spacetime evolution of the population difference in a two-photon transition and of the two-photon coherence during pulse propagation are analyzed. It is established that a complete population inversion and a maximum coherence may exist over a long length of the medium if the corresponding conditions are satisfied at the boundary. A new possibility for achieving a high coherence (close to a maximum value) is proposed. *© 2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

In recent years, new concepts based on atomic coherence and effects of quantum interference have given a strong impetus to the development of the resonance nonlinear optics of gaseous media. Various effects in these realms, such as induced transparency [1], the coherent trapping of populations [2, 3], and the adiabatic transfer of populations [4] (or stimulated Raman adiabatic passage [5]), have found widespread use. The above phenomena make it possible to control linear and nonlinear optical properties of matter. Concurrently, laws that govern the propagation of laser pulses under resonance conditions change significantly (there arise matched pulses [1], adiabatons [6], a decrease in the group velocity of a pulse by a factor of 10^7 to 10^8 [7], and some other effects [8–10]). New possibilities open up here, including that for a high-efficiency implementation of nonlinear-optics interaction even for very weak light fields [11, 12] and for the generation of subfemtosecond light pulses [13].

The new effect of Stark-chirped rapid adiabatic passage (referred to in the following as SCRAP) was recently demonstrated in [14]. It is closely related to a rapid adiabatic passage due to sweeping the pumping-laser frequency (see, for example, [15]) or due to the intrinsic Stark shift of levels in two-photon interaction [16]. In contrast to what occurs in the phenomena described in those studies, the passage through a resonance in the case of SCRAP is accomplished owing to the dynamic Stark shift caused by an additional strong laser pulse that is delayed in time with respect to the pumping-laser pulse. Specific conditions must be satisfied for the observation of SCRAP to be possible. The SCRAP phenomenon also differs from the coherent two-quantum interaction of ultrashort pulses of light under the conditions of self-induced transparency [17].

SCRAP makes it possible to coherently control both populations and coherence in atomic and molecular systems. It enables one to obtain the inversion of populations between high-lying levels and the ground-state level [14] or to prepare atoms in states that are characterized by a maximum two-photon coherence [18–20]. It was proposed that the latter be employed to ensure an efficient generation of coherently tuned radiation—in the vacuum-ultraviolet region inclusive—in the process of nonlinear-optics mixing of short pulses [19, 20]. In all probability, this effect may be used to control chemical reactions and to treat quantum information; also, it may find applications in atomic optics.

Only the time dynamics of population inversion between levels and of the maximum two-photon coherence in the SCRAP process was investigated in [14, 18] in terms of adiabatic states. In the present study, we examine the propagation of a pumping-laser pulse under the SCRAP conditions and analyze the spacetime dynamics of level population and of two-photon coherence with allowance for the propagation of this pulse. The analysis is performed on the basis of a vector model that is quite clear and which makes it possible to obtain analytic expressions both for populations and for the off-diagonal density-matrix element describing atomic coherence with allowance for the first nonadiabatic correction. Below, we show that this correction

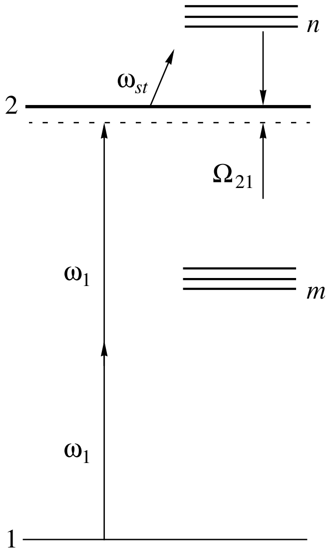

Fig. 1. Simplified diagram of energy levels for Stark-chirped rapid adiabatic passage.

must be taken into account in considering the propagation of a pulse in a medium. We also show that, under the SCRAP conditions, population inversion in a resonance transition and the induced maximum coherence may occur over a long distance in a medium.

This article is organized as follows. In Section 2, we discuss the SCRAP phenomenon on the basis of a vector model, disregarding pulse propagation. There, we present basic equations and their solutions, describe the pattern of the SCRAP effect in terms of the vector model, and discuss conditions necessary for its observation. In Section 3, we derive equations that describe the propagation of a pumping-laser pulse under the SCRAP conditions, give their solution, and discuss the spacetime dynamics of population inversion and atomic coherence with allowance for pulse propagation. In the Conclusions, we consider the possible applications of our results.

2. EFFECT OF RAPID ADIABATIC PASSAGE THROUGH A TWO-PHOTON RESONANCE DUE TO A LASER-INDUCED STARK SHIFT: VECTOR REPRESENTATION

A simplified energy-level diagram corresponding to the process under consideration is shown in Fig. 1. In order to observe the SCRAP effect, use is made of two laser pulses following each other with a specific time delay. One of them (triggering or pumping pulse) interacts with the *1–2* two-photon transition (see Fig. 1), its frequency being initially detuned from the two-photon resonance; the other, strong, off-resonance (Stark), pulse leads to a change in the two-photon frequency, inducing a time-dependent dynamic Stark shift. As the two-photon detuning changes from large negative to large positive values (or vice versa), the population of the ground-state (lower) level (*1*) can be transferred almost entirely, by the end of the triggering pulse, to the excited (upper) state (level *2*); that is, a complete population inversion may arise in the *1–2* transition, the lifetime of this inversion being determined by the relaxation time for state *2*. The pulse durations are assumed here to be shorter than all atomic relaxation times.

This process is usually explained in terms of dressed (adiabatic) states [14], but it can also be described in terms of the vector model proposed in [16]. Within this model not only can one visualize the time dynamics of the process, but also obtain analytic expressions for population inversion and for two-photon coherence.

2.1. Equations for Interacting Fields and for the Density Matrix in the Approximation of a Generalized Two-Level System

We will first consider the time evolution of the population inversion $\rho_2 – \rho_1$ in the *1–2* two-photon transition and of the off-diagonal density-matrix element ρ_{21}, which describes two-photon coherence under the SCRAP conditions. For ρ_{21}, we hereafter use the term “two-photon coherence” or merely “coherence.” The triggering pulse $\mathscr{E}_1(z, t)$ and the Stark pulse $\mathscr{E}_s(z, t)$ propagate along the z axis in an isotropic medium,

$$\mathscr{E}_{1,s}(z,t) = \frac{1}{2}E_{1,s}(z,t)\exp[-i(\omega_{1,s}t - k_{1,s}z)] + \text{c.c.}, \tag{1}$$

where $k_{1,s} = \omega_{1,s}/c$ is the wave number at the frequency $\omega_{1,s}$, c is the speed of light, and $E_{1,s}(z, t)$ stands for complex-valued slowly varying amplitudes (envelopes). Since the Stark pulse undergoes off-resonance interaction with the medium, we assume that it is preset in the process of propagation and that the amplitude is a real-valued quantity.

The propagation of the triggering pulse is described by the reduced wave equation

$$\frac{\partial E_1(z,t)}{\partial z} + \frac{1}{c}\frac{\partial E_1(z,t)}{\partial t} = -i2\pi k_1 P_1(z,t), \tag{2}$$

where $P_1(z, t)$ is the complex-valued amplitude that characterizes polarization induced by the triggering pulse. It can be calculated on the basis of the equation for the density matrix,

$$P_1(\omega_1) = N\sum_m (d_{1m}\rho_{m1} + d_{m2}\rho_{2m}), \tag{3}$$

where N is the density of atoms, d_{ij} stands for the transition electric dipole moments, and ρ_{ij} are the components of the density matrix.

In the following, we assume, for the frequency of the two-photon quasiresonance, that

$$\Omega_{21} = 2\omega_1 - \omega_{21} \ll \omega_{ij},$$

we also assume that the one-photon detuning satisfies the conditions

$$\Omega_{ji} = \omega_k - \omega_{ji} \gg |G_{ji}| \gg T_k^{-1}, \qquad (4)$$

where $G_{ji} = E_k d_{ji}/2\hbar$ is the Rabi frequency and T_k is the pulse duration. In the approximation specified by Eq. (4), one can pass from a multilevel system to a generalized two-level scheme (see, for example, [21]). Level *1* is the ground-state level; the group of levels labeled with *m* makes a leading contribution to the two-photon quantum transition and to the Stark shift due to the triggering field; and the group of levels labeled with *n* contributes to the Stark shift due to the Stark field.

If use is made of the rotating-wave approximation, the equations describing the density matrix for the generalized two-level system in the interaction representation have the form

$$\begin{aligned} \frac{\partial(\rho_2 - \rho_1)}{\partial t} &= -2qA_1^2 \operatorname{Im}\rho_{21}, \\ \frac{\partial\rho_{21}}{\partial t} - i(\Omega'_{21} - \Omega_s)\rho_{21} &= \frac{i}{2}qA_1^2(\rho_2 - \rho_1). \end{aligned} \qquad (5)$$

In Eq. (5), we introduced the following notation:

$$E_1 = A_1 \exp(-i\varphi_1),$$

where A_1 and φ_1 are, respectively, the real-valued amplitude and phase, which are functions of coordinates and time;

$$q = \frac{1}{2\hbar^2}\sum_m \frac{d_{2m}d_{m1}}{\Omega_{m1}}, \quad \Omega_{m1} = \omega_1 - \omega_{m1},$$

$$\Omega'_{21} = \Omega_{21} + \frac{2\partial\varphi_1}{\partial t} = 2\omega'_1 - \omega_{21}, \quad \Omega_{21} = 2\omega_1 - \omega_{21}$$

is the initial (static) detuning, with ω_{21} and $\omega'_1 = \omega_1 + \partial\varphi_1/\partial t$ being, respectively, the resonance transition frequency and the instantaneous frequency; and $\Omega_s = S_1 + S$ is the Stark shift of the two-photon resonance due to the triggering ($S_1 = a_1 A_1^2$) and the Stark ($S = a_s E_s^2$) pulse, with

$$\begin{aligned} a_1 &= \frac{1}{4\hbar^2}\sum_m \left(\frac{|d_{m1}|^2}{\omega_1 - \omega_{m1}} + \frac{|d_{2m}|^2}{\omega_1 - \omega_{2m}} \right), \\ a_s &= \frac{1}{4\hbar^2}\sum_n \frac{|d_{n1}|^2}{\omega_s - \omega_{n1}} \end{aligned} \qquad (6)$$

being parameters that describe the Stark shifts.

In the above approximation, the polarization given by Eq. (3) can be represented in the form

$$\begin{aligned} P_1(\omega_1) = {} & [\chi_1(\omega_1)\rho_1 - \chi_2(\omega_1)\rho_2] N A_1 \exp(i\varphi_1) \\ & - 2N\hbar q A_1 \rho_{21} \exp(i\varphi_1), \end{aligned} \qquad (7)$$

where $\chi_{1,2}(\omega_1)$ are linear susceptibilities,

$$\chi_i(\omega_1) = \frac{1}{2\hbar}\sum_m \frac{|d_{mi}|^2}{\omega_1 - \omega_{mi}}, \quad i = 1, 2. \qquad (8)$$

2.2. Vector Representation

It is convenient to rewrite Eq. (5) by using the vector model proposed in [15, 16] and by introducing the notation

$$r_1 = \rho_{21} + \rho_{12}, \quad r_2 = i(\rho_{21} - \rho_{12}), \quad r_3 = \rho_2 - \rho_1.$$

The quantities r_j ($j = 1, 2, 3$) can be treated as the components of the vector

$$\mathbf{r} = \mathbf{e}_1 r_1 + \mathbf{e}_2 r_2 + \mathbf{e}_3 r_3$$

($\mathbf{e}_{1, 2, 3}$ are unit vectors) in some vector space (vector **r** is sometimes referred to as a pseudospin). The pseudospin satisfies the equation of motion

$$\frac{\partial\mathbf{r}}{\partial t} = \boldsymbol{\gamma} \times \mathbf{r}, \qquad (9)$$

where $\boldsymbol{\gamma}$ is a vector whose components are given by

$$\gamma_1 = qA_1^2, \quad \gamma_2 = 0, \quad \gamma_3 = (\Omega'_{21} - \Omega_s).$$

These parameters have a clear physical meaning: γ_1 is the effective two-photon Rabi frequency and γ_3 is the instantaneous detuning away the two-photon resonance.

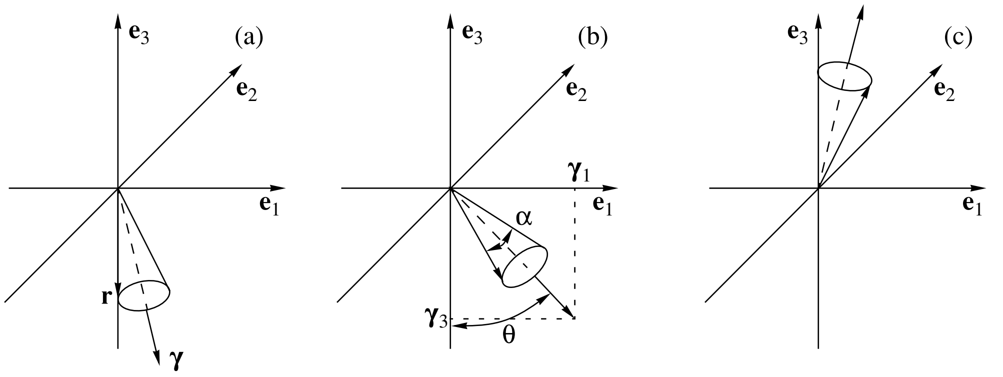

Fig. 2. Evolution of vector **r** during Stark-chirped rapid adiabatic passage: (a) the case where the two-photon frequency of the triggering pulse is less than the resonance frequency and where vector **r** precesses over the cone surface about vector $\boldsymbol{\gamma}$; (b) the case where, as the sweeping of the frequency of the passage through the resonance due to the Stark shift occurs, the angle θ between the vector $\boldsymbol{\gamma}$ and the negative direction of the e_3 axis increases, while vector **r** continues precessing about the vector $\boldsymbol{\gamma}$; and (c) the case where the doubled frequency of the triggering pulse is greater than the resonance transition frequency and where, upon the completion of motion, vector $\boldsymbol{\gamma}$ is oriented in a direction close to the positive direction of the e_3 axis.

The equations of motion for the components of vector **r** have the form

$$\frac{\partial r_1}{\partial t} = \gamma_3 r_2, \quad \frac{\partial r_2}{\partial t} = -\gamma_3 r_1 + \gamma_1 r_3, \quad \frac{\partial r_3}{\partial t} = -\gamma_1 r_2. \tag{10}$$

In this approximation, the solution to Eqs. (10) with the allowance for the first nonadiabatic correction has the form (see, for example, [15])

$$r_1 = \mp\frac{\gamma_1}{\gamma}, \quad r_3 = \mp\frac{\gamma_3}{\gamma}, \quad r_2 = -\frac{1}{\gamma_3}\frac{\partial r_1}{\partial t}, \tag{11}$$

where

$$\gamma = |\gamma| = \sqrt{\gamma_1^2 + \gamma_3^2}.$$

In expressions (11) and in those that follow, the upper (lower) signs are chosen for the case of $\gamma_3(z = 0, t \longrightarrow -\infty) < 0$ $[\gamma_3(z = 0, t \longrightarrow -\infty) > 0]$.

We note that, upon the inclusion of the nonadiabatic correction, the coherence

$$\rho_{21} = \frac{r_1 - ir_2}{2}$$

becomes complex-valued. As will be shown below, this must be taken into account in considering the propagation of a pulse in a medium.

It is convenient to introduce the parameter $\theta = \arctan(\gamma_1/\gamma_3)$, which is the angle between vector $\boldsymbol{\gamma}$ and the negative direction of the e_3 axis aligned with unit vector $\mathbf{e}_3$ (see Fig. 2). In this notation, we have $r_1 = \pm\sin\theta$ and $r_3 = \mp\cos\theta$.

The adiabaticity condition implies that the angular velocity $\partial\theta/\partial t$ of vector $\boldsymbol{\gamma}$ is much less than the frequency γ of precession of vector **r**; that is,

$$\left|\frac{\partial\theta}{\partial t}\right| = |\dot{\theta}| \ll \gamma = \sqrt{\gamma_1^2 + \gamma_3^2}. \tag{12}$$

By using expression (12), it can easily be shown that $r_2 = -\dot{\theta}/\gamma$. This means that the angle between **r** and $\boldsymbol{\gamma}$ is much smaller than the angle between $\boldsymbol{\gamma}$ and $\mathbf{e}_3$ (see Fig. 2); that is, $|r_2| \ll |r_1|$.

Adiabaticity condition (12) can be represented in the form

$$\left|\frac{\dot{\gamma}_1\gamma_3 - \dot{\gamma}_3\gamma_1}{(\gamma_1^2 + \gamma_3^2)^{3/2}}\right| \ll 1. \tag{13}$$

This condition is similar to that obtained in [5] in terms of adiabatic states.

Let us show that solution (11), together with condition (13), describes the SCRAP phenomenon well. We consider the case where the two-photon Rabi frequency $\gamma_1(t)$ and the Stark shift $S(t)$ of the frequency of the transition induced by the second pulse have Gaussian shape; that is,

$$\gamma_1(t) = \gamma_{10}\exp\left(-\frac{t^2}{T_1^2}\right), \quad S(t) = S_0\exp\left[-\frac{(t-\delta t)^2}{T_s^2}\right], \tag{14}$$

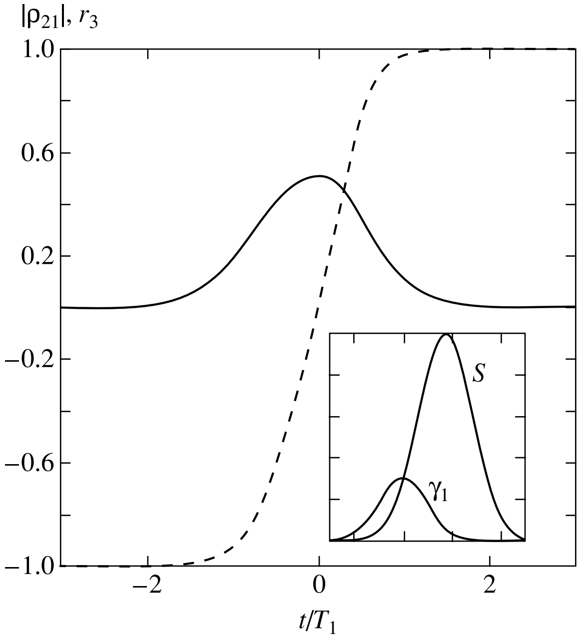

Fig. 3. Time dependences of (solid curve) the two-photon coherence $|\rho_{21}|$ and (dashed curve) the population difference $r_3 = \rho_2 - \rho_1$ for the pulse sequence shown in the inset (the triggering pulse precedes the Stark pulse). The following parameter values were used here: $\delta t/T_1 = 1.7$ for the delay between the pulses; $T_s/T_1 = 1.6$ for the duration of the Stark pulse; $S_0T_1 = 50$ for the maximum Stark shift; $\gamma_{10}T_1 = 15$ for the amplitude of the two-photon Rabi frequency; and $\Omega_{21}T_1 = -16$ for the initial detuning.

where T_1 and T_s are the durations of, respectively, the triggering and the Stark pulse and δt is the time delay between them. In the ensuing analysis, we disregard, for the sake of simplicity, the shift due to the triggering pulse. In order to achieve SCRAP, it is necessary to meet certain requirements. Fulfillment of the adiabaticity condition (13) can be ensured if the initial detuning Ω_{21}, the maximum Rabi frequency γ_{10}, and the delay time δt satisfy a specific relation. For passage through a two-photon resonance to occur, it is necessary, first of all, that the maximum Stark shift S_0 exceed the initial detuning Ω_{21} ($|S_0| > |\Omega_{21}|$) and that the two quantities in question have the same sign. It is obvious that, under these conditions, the resonance is swept twice, at the instants of time

$$t_{1,2} = \delta t \mp T_s\sqrt{\ln\frac{S_0}{\Omega_{21}}}.$$

In this case, the triggering pulse must be rather strong in order to ensure an adiabatic transition during the first passage through the resonance, but it must be at the same time rather weak in order to prevent an adiabatic transition during the second passage. Under the assumption that the triggering pulse attains a maximum at the instant of time $t_1 = 0$, these requirements lead to the conditions [14]

$$1 \ll \frac{\gamma_{10}^2 T_s^2}{\Omega_{21}\delta t} \ll \exp\frac{8\delta t^2}{T_1^2}. \qquad (15)$$

Relation (15) yields an upper and a lower limit on the detuning Ω_{21} and the maximum value γ_{10} of the Rabi frequency. From the formula for r_3 in (11), it follows that, for the population transfer to be maximal at the end of the triggering pulse, fulfillment of the condition $\gamma_1 \ll |\gamma_3|$ must be ensured, which is most easily achieved at the instant $t = \delta t$, when the Stark shift S is maximal. This condition can be recast in the form

$$\gamma_{10}\exp\left(-\frac{\delta t^2}{T_1^2}\right) \ll |\Omega_{21} - S_0|. \qquad (16)$$

At a specific instant of time when the population difference vanishes, $r_3 = 0$, the two-photon coherence attains the maximum value of $|\rho_{21}| = 1/2$. These conditions can be satisfied in a rather wide region.

For numerical illustrations, we have chosen the diagram that represents transitions in the Kr atom and which was used in experiments aimed at generating vacuum-ultraviolet radiation under the conditions of electromagnetically induced transparency [22]: there, a triggering pulse of a wavelength of $\lambda_1 \approx 212.55$ nm undergoes two-photon interaction with the $4p^{6\,1}S$–$4p^5 5p$ transition. Since relaxation times characteristic of the transitions being considered are about a few tens of nanoseconds, the duration of the laser pulses used must not exceed a few nanoseconds.

The population difference r_3 and the modulus of the atomic coherence, $|\rho_{21}| = |(r_1 - ir_2)/2|$, versus time are shown in Fig. 3, whence it can be seen that the population inversion and the maximum coherence in the two-photon transition are realized under the above conditions.

In terms of the vector model, the SCRAP process can be visualized as follows. Suppose that, at the initial instant of time ($t \longrightarrow -\infty$), all atoms are in the ground state *1* ($r_3 = -1$) and that the double frequency of the triggering pulse is tuned rather far away from the resonance and is, for definiteness, less than the resonance frequency of the *1–2* transition. This means that the vector $\boldsymbol{\gamma}$ is directed nearly along the negative direction of the e_3 axis and is virtually parallel to vector $\mathbf{r}$ (that is, $\theta \approx 0$). In this case, vector $\mathbf{r}$ moves over the surface of a cone (precesses about vector $\boldsymbol{\gamma}$) having a small apex angle proportional to r_2 (see Fig. 2). As a Stark pulse is applied, the sweeping of the frequency toward the resonance begins, which is accompanied by the growth of

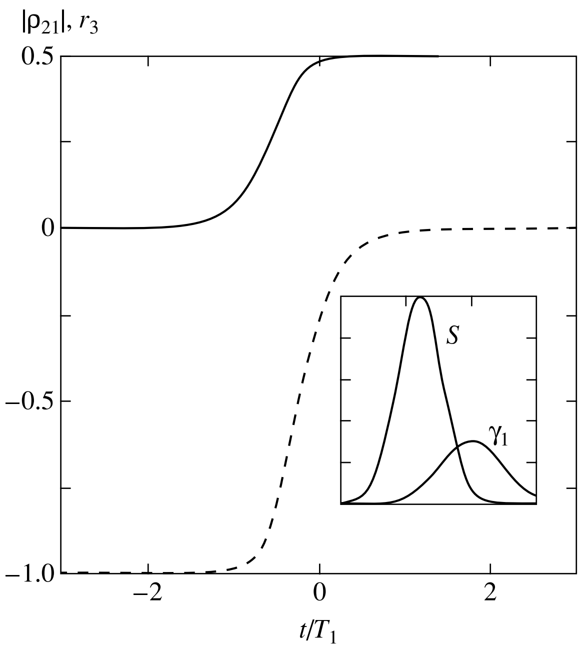

Fig. 4. Time dependences of (solid curve) the two-photon coherence $|\rho_{21}|$ and (dashed curve) the population difference $r_3 = \rho_2 - \rho_1$ for the case where the Stark pulse precedes the triggering pulse, as is shown in the inset. The following parameter values were used here: $\delta t/T_1 = -1.6$, $T_s/T_1 = 1$, $S_0T_1 = 50$, $\gamma_{10}T_1 = 15$, and $\Omega_{21} = 0$.

angle θ. If angle θ varies more slowly than the precession frequency γ [condition (13) is satisfied], vector $\mathbf{r}$ will follow the vector $\boldsymbol{\gamma}$ as the angle θ varies, its precession persisting. The amplitude of the Stark pulse must be such that, upon passage through the resonance, the frequency of the triggering pulse appears to be considerably higher than the resonance transition frequency (see Fig. 2). At the end of the pulse, vector $\mathbf{r}$ will form a small angle with the positive direction of the e_3 axis. Thus, the inversion of the level populations is realized owing to sweeping through the resonance. The inclusion of r_2 does not change the pattern qualitatively, but one can show that, in this case, the precession of vector $\mathbf{r}$ occurs about vector $\boldsymbol{\gamma} + \dot{\boldsymbol{\theta}}$, where $\boldsymbol{\gamma} = (0, 0, -\gamma_1)$ and $\dot{\boldsymbol{\theta}} = (0, -\dot{\theta}, 0)$ are the vectors written in the doubly rotating coordinate frame (see, for example, [15]).

By using this technique, one can prepare atoms in a coherent superposition of states *1* and *2* that is characterized by a maximum coherence [18]. This effect, known as the half-SCRAP effect [18], can also be described in terms of the vector model. In contrast to SCRAP, it arises under different conditions. From the formula for r_1 in (11), it follows that $r_1 \longrightarrow \mp 1$ if $\gamma_1 \gg |\Omega_{21} - S|$ (near the instant of time at which the triggering field takes a maximum value). At the initial instant of time, we have $r_3 = -1$; therefore, it is obvious that the inequality $\gamma_1 \ll |\Omega_{21} - S|$ must hold at this instant. These conditions are satisfied most readily at $\Omega_{21} = 0$. Moreover, they must be consistent with adiabaticity condition (13). Analysis shows that these conditions may be satisfied for the sequence of pulses shown in the inset to Fig. 4, in which case the Stark pulse is switched on earlier than the triggering pulse. The case of the opposite sequence of pulses—that is, the case in which the triggering pulse is switched on earlier than the Stark pulse [18]—is not considered in the present approximation.

Figure 4 shows the time dependences of r_3 and $|\rho_{21}| = |(r_1 - ir_2)/2|$ for the chosen sequence of pulses. It can be seen that, in contrast to what occurs in the case of SCRAP, there arises here a plateau of maximum coherence. As the Stark and the triggering field change, vector $\mathbf{r}$ initially oriented nearly along the negative direction of the e_3 axis then rotates through angle $\pi/2$, becoming aligned with the e_1 axis.

3. PROPAGATION OF TRIGGERING PULSE AND SPACE DYNAMICS OF THE POPULATION INVERSION AND OF THE COHERENCE

Let us now consider the propagation of a triggering pulse under the SCRAP conditions. We express the complex-valued polarization $P_1(\omega_1)$, which is defined by formula (7), in terms of quantities $r_{1,2,3}$ as

$$\begin{aligned} P_1(\omega_1) &= \frac{1}{2}\{[\chi_1(1 - r_3) - \chi_2(1 + r_3)]NA_1\exp(i\varphi_1)\} \\ &\quad - \hbar q(r_1 - ir_2)NA_1\exp(i\varphi_1) = (U + iV)\exp(i\varphi_1), \end{aligned} \tag{17}$$

where N is the density of atoms. It should be emphasized that the imaginary part V of the polarization owes its existence to the nonadiabatic correction r_2 and leads, despite its smallness, to a change in the shape of a pulse as it propagates in a medium.

Substituting expression (17) into Eq. (2), we obtain equations for the real-valued amplitude and phase; that is,

$$\frac{\partial A_1}{\partial z} + \frac{1}{c}\frac{\partial A_1}{\partial t} = \pm 2\pi k_1 V, \tag{18}$$

$$\frac{\partial \varphi_1}{\partial z} + \frac{1}{c}\frac{\partial \varphi_1}{\partial t} = \mp 2\pi k_1 \frac{U}{A_1}, \tag{19}$$

where

$$U = \frac{1}{2}\left[\chi_1(\omega_1)\left(1-\frac{\gamma_3}{\gamma}\right)-\chi_2(\omega_1)\left(1+\frac{\gamma_3}{\gamma}\right)\right]NA_1 - \frac{\hbar q^2 N A_1^3}{\gamma}, \tag{20}$$

$$V = -\frac{\hbar q^2 N A_1}{\gamma_3}\frac{\partial}{\partial t}\frac{A_1^2}{\gamma}.$$

Here, the upper (lower) signs correspond to a negative (positive) initial detuning away the two-photon resonance, $\Omega_{21} = 2\omega_1 - \omega_{21} < 0$).

If we discard the nonadiabatic correction ($V = 0$ corresponds to the ideal adiabatic limit), the right-hand side of Eq. (18) vanishes, in which case one can see that the pulse propagates in a medium without any distortion of its shape. As to the phase of the pulse, it will change anyway according to Eq. (19)—that is, sweeping of the instantaneous frequency occurs as the pulse propagates. Upon inclusion of the nonadiabatic correction, the two-photon coherence becomes complex-valued, with the result that the medium-induced macroscopic polarization develops an imaginary part proportional to the derivative $\partial(A_1^2/\gamma)/\partial t$. Its sign may be either positive (static detuning is negative, $\Omega_{21} < 0$) or negative; that is, the generated field can either enhance or suppress a pulse propagating in a medium. Below, we will show that this is determined by the sign of the derivative of the instantaneous detuning away from the two-photon resonance with allowance for the Stark shift.

Evaluating the derivative $\partial(A_1^2/\gamma)/\partial t$ and substituting the expression for V into Eq. (18), we find that the real-valued amplitude obeys the equation

$$\frac{\partial A_1}{\partial z} + \frac{1}{c}\frac{\partial A_1}{\partial t} = \pm\frac{2\pi\omega_1\hbar q^2 N}{c\gamma^3}\left(-A_1^3\frac{\partial\Delta}{\partial t} + 2\Delta A_1^2\frac{\partial A_1}{\partial t}\right), \tag{21}$$

where

$$\Delta = 2\omega_1' - \omega_{21} - \Omega_s^{(2)}$$

is the instantaneous detuning away from the two-photon resonance with allowance for the shift induced by the external Stark pulse.

It is convenient to pass from the equation for the phase to the equation for the instantaneous frequency $\omega' = \omega + \partial\varphi/\partial t$; that is,

$$\frac{\partial\omega_1'}{\partial z} + \frac{1}{c}\frac{\partial\omega_1'}{\partial t} = \mp\frac{2\pi\omega_1\hbar q^2 N}{c\gamma^3}\left(-A_1^3\frac{\partial\Delta}{\partial t} + 2\Delta A_1^2\frac{\partial A_1}{\partial t}\right)\frac{\Delta}{A_1}. \tag{22}$$

Equations (21) and (22) describe the propagation of a light pulse under the SCRAP conditions as long as adiabaticity condition (13) is satisfied. One can see that the right-hand sides of these equations differ only by the factor Δ/A_1. This indicates that the changes in the frequency and in the pulse shape occur simultaneously and that there is a relation between them. The first terms on the right-hand sides of the equations in question are due to the high-frequency Kerr effect, which leads both to a change in the shape of a pulse as it propagates in a medium and to phase self-modulation during the propagation process. The second terms lead to a change in the velocity of pulse propagation.

We note that, if there occurs a passage through the resonance ($\Delta = 0$) in scanning the frequency, the sign of the derivative does not change, as in the SCRAP case, since Δ is either an increasing (for a negative initial detuning, $\Omega_{21} < 0$) or a decreasing ($\Omega_{21} > 0$) function of time. The field generated by the imaginary part of the polarization is then in antiphase with the incident pulse, suppressing it. The situation is different if the scanned frequency does not reach the resonance. The derivative will be positive within one part of the pulse and negative within the remaining part, so that one part of the pulse will be suppressed, while the remaining part will be enhanced.

Going over to the coordinates $\xi = z$ and $\tau = t - z/c$ and taking into account Eq. (21), one can recast Eq. (22) in the form

$$\frac{\partial\omega_1'}{\partial\xi} = -\frac{\Delta(\tau,\xi)}{A_1(\tau,0)}\frac{\partial A_1}{\partial\xi}. \tag{23}$$

Integrating Eq. (23), we obtain

$$\frac{\Delta(\tau,\xi)}{\Delta(\tau,0)} = \frac{A_1^2(\tau,0)}{A_1^2(\tau,\xi)}. \tag{24}$$

This is a general relation that is valid in the adiabatic approximation. It reflects the interplay of the change in the instantaneous frequency of a pulse and the change in its envelope during the propagation of the pulse in a medium under the conditions of a two-photon resonance. Relation (24) also shows that the effects of the change in the pulse shape and of self-modulation are related to each other and cannot be considered separately.

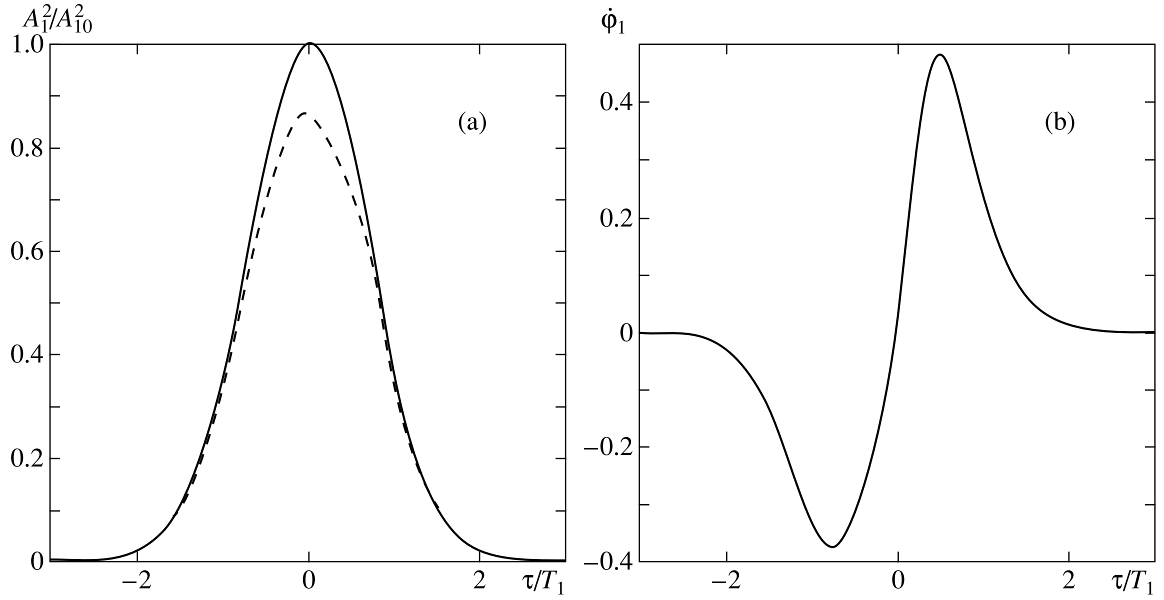

Fig. 5. Spacetime evolution of a pulse under the SCRAP conditions: (a) envelope of the pulse as a function of time at Z = (solid curve) 0 and (dashed curve) 10; (b) instantaneous frequency $\dot{\varphi}$ of the triggering pulse as a function of time at $Z = 10$ (it is normalized to $\alpha_0 = 8\pi k_1 N\chi_1(\omega_1)$). The parameters of the pulse were set to the values identical to those used for Fig. 3.

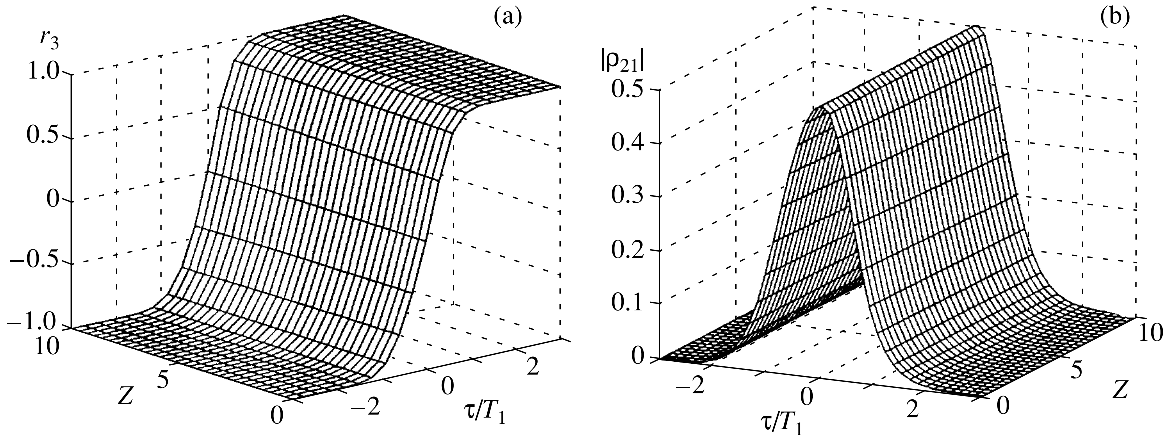

Fig. 6. Spacetime evolution of (a) the population difference $r_3 = \rho_2 - \rho_1$ and (b) the two-photon coherence $|\rho_{21}|$ under the SCRAP conditions. The pulse parameters were set to the values identical to those used for Fig. 3.

Let us rewrite Eq. (21) as

$$\frac{\partial A_1}{\partial z} + \frac{1}{v_p}\frac{\partial A_1}{\partial t} = \pm\frac{2\pi\omega_1\hbar q^2 N}{c\gamma^3}A_1^3\frac{\partial\Delta}{\partial t}, \qquad (25)$$

where

$$\frac{1}{v_p} = \frac{1}{c}\left(1 \pm \frac{4\pi\omega_1\hbar q^2 N}{\gamma^3}\Delta A_1^2\right). \qquad (26)$$

The results obtained by numerically solving Eq. (25) are illustrated in Fig. 5, where the time dynamics of the normalized intensity of the triggering pulse is shown at the input and output of the medium. At the input, the triggering and the Stark pulse have a Gaussian shape, their parameters and the time of delay between them being chosen in such a way as to ensure fulfillment of conditions (15) and (16), which are necessary for implementing SCRAP and a complete population transfer. One can see that the amplitude of a propagating pulse decreases smoothly, while its shape changes slowly. We can identify two effects that are responsible for the change in the pulse shape: (i) that which is due to a diabatic character of interaction [formula (26)] and (ii) that which is due to phase self-mod-

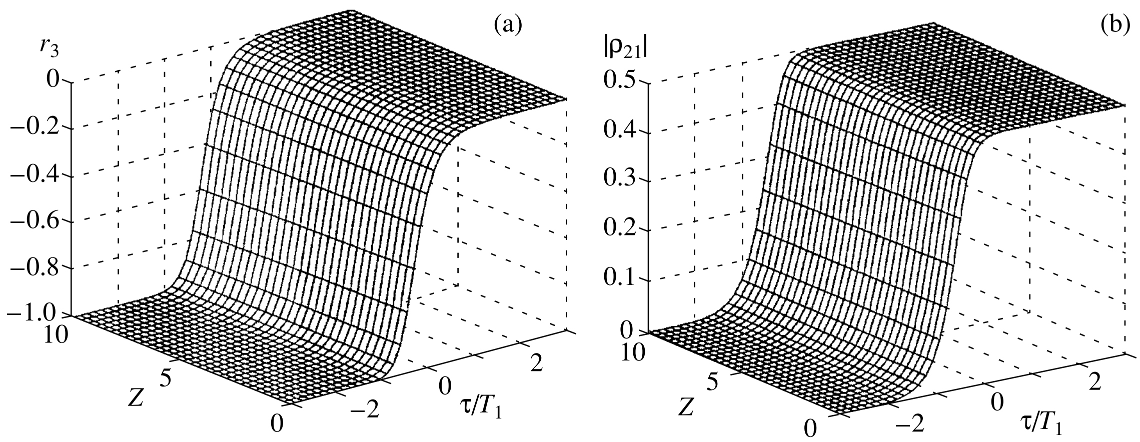

Fig. 7. Spacetime evolution of (a) the population difference $r_3 = \rho_2 - \rho_1$ and (b) the two-photon coherence $|\rho_{21}|$ under the half-SCRAP conditions. The pulse parameters were set to the values identical to those used for Fig. 4.

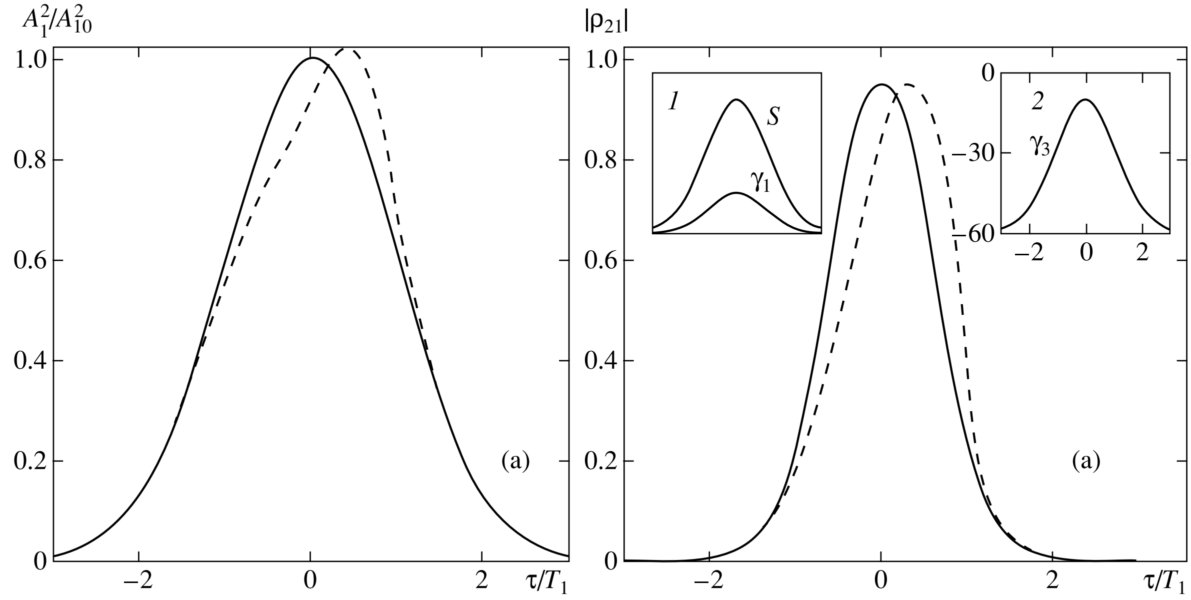

Fig. 8. Time evolution of the triggering pulse and the two-photon coherence $|\rho_{21}|$ in the case where the sweeping of the frequency does not lead to passage through the resonance: (a) envelope of the pulse at Z = (solid curve) 0 and (dashed curve) 70; (b) two-photon coherence $|\rho_{21}|$ (the notation for the curves is identical to that in Fig. 8a). The insets display (*1*) the timing of the triggering and the Stark pulse and (*2*) the instantaneous detuning γ_3 away the two-photon resonance. The following parameter values were used here: $\delta t = 0$, $\Omega_{21}T_1 = -60$, $T_s/T_1 = 1.6$, $S_0T_1 = 50$, and $\gamma_{10}T_1 = 15$.

ulation because of the high-frequency Kerr effect. The instantaneous frequency of the pulse varies with time as is shown in Fig. 5b. In order to calculate $\dot{\varphi}$, we used Eq. (24).

Figures 6 and 7 show the spacetime evolution of the population difference and of the two-photon coherence under the SCRAP and half-SCRAP conditions, respectively. We can see that the population inversion and the maximum coherence induced in the two-photon transition persist over a long length of the medium.

We would like to emphasize that, in order to achieve a high two-photon coherence (close to a maximum value), it is not necessary to pass through the resonance. As can be seen from the formula for r_1 in (11), fulfillment of condition $|\gamma_1| \gg |\gamma_3|$, along with adiabaticity condition (13), is sufficient for this. This condition means that the two-photon Rabi frequency must exceed the detuning away the resonance at the instant of time when the difference of the populations is close to zero, $\rho_2 - \rho_1 = r_3 \approx 0$. Figure 8a shows how the envelope of the triggering pulse that propagates in a medium varies in this case. A moderate increase in the pulse amplitude is associated with the redistribution of energy in the pulse due to the nonadiabatic correction, which leads to the weakening of the forward front of the pulse and to the strengthening of the backward front [see discussion after formula (22)], the total pulse energy (the area under the intensity curve) remaining virtually unchanged. Figure 8b shows that, under the quasiresonance conditions, a two-photon coherence close to a maximum value is induced over a long length of the medium. This can be used to implement an efficient generation of the third harmonic in the short-wavelength region of the spectrum.

4. CONCLUSIONS

On the basis of the vector model for a two-photon resonance, the SCRAP phenomenon, which can lead to a complete population transfer and a maximum coherence in the resonance transition, has been considered in the two-level approximation. This approach has enabled us to derive analytic expressions for populations and two-photon coherence with allowance for a diabatic character of the interaction, to analyze them over a broad region of parameter values, and to specify the adiabaticity conditions, as well as to develop a simple geometric interpretation of the effect. On this basis, we have investigated special features of the propagation of short laser pulses under various conditions. We have also proposed a new possibility for achieving a high coherence (close to a maximum value).

It has been shown that, because of a diabatic character of the interaction, the off-diagonal density-matrix element describing the two-photon coherence is complex-valued. Its imaginary part induces a second-order nonlinear Kerr polarization, which is responsible for the change in energy transfer between the propagating pulse and the medium, while the real part of the Kerr polarization leads to phase self-modulation. As the pulse propagates, these factors lead to a decrease in its amplitude and to a change in its shape (over small propagation lengths, the pulse shape remains virtually unchanged). In addition, the carrier frequency of the pulse changes with time, the region of linear sweeping existing over a large part of the pulse duration. We note that the change in the pulse frequency is related to the change in the pulse shape since they are both due to self-interaction via the high-frequency Kerr effect in a time-dependent laser field. It seems that the duration of such a pulse can be reduced by transmitting it through a dispersive delay line by using the compression effect [23].

We have also investigated the spacetime evolution of the population difference and of the two-photon coherence during pulse propagation. We have shown that the population inversion and the maximum coherence persist over a long medium length under specific conditions at the boundary. This method makes it possible to obtain a nearly complete inversion between the ground state and a high-lying excited level or to force a medium into a coherent state by using visible-range lasers that produce pulses of duration shorter than the relaxation time of the excited state. In view of this, the proposed method can find various applications in creating new sources of pulsed coherent radiation. By way of example, we indicate that, on the basis of the anti-Stokes Raman scattering, a medium inverted in this way can be used to generate, in the short-wavelength region of the spectrum, including the vacuum-ultraviolet region, short pulses tunable with respect to the wavelength. In all probability, cooperative anti-Stokes scattering can also be investigated by using this scheme. We would also like to note the possibility of generating the third harmonic in the short-wavelength region of the spectrum under the conditions of maximum coherence.

ACKNOWLEDGMENTS

This work was supported in part by the Russian Foundation for Basic Research (project no. 02-02-16325a), by the program Universities of Russia (grant no. UR.01.01.003), and by INTAS (grant no. 99-00019).

REFERENCES

1. S. E. Harris, Phys. Today **50**, 36 (1997).
2. B. D. Agap'ev, M. B. Gornyĭ, B. G. Matisov, and Yu. V. Rozhdestvenskiĭ, Usp. Fiz. Nauk **163** (9), 1 (1993) [Phys.–Usp. **36**, 763 (1993)].
3. E. Arimondo, in *Progress in Optics*, Ed. by E. Wolf (North-Holland, Amsterdam, 1996), Vol. 35, p. 257.
4. J. R. Kuklinsky, U. Gaubats, F. T. Hioe, *et al.*, Phys. Rev. A **40**, 6741 (1989).
5. N. V. Vitanov, M. Fleischhauer, and K. Bergmann, Adv. At. Mol. Opt. Phys. **46**, 55 (2001).
6. R. Grobe, F. T. Hioe, and J. H. Eberly, Phys. Rev. Lett. **73**, 3183 (1994).
7. A. B. Matsko, O. Kocharovskaya, Yu. Rostovtsev, *et al.*, Adv. At. Mol. Opt. Phys. **46**, 191 (2001).
8. M. Fleischhauer and A. S. Manka, Phys. Rev. A **54**, 794 (1996).
9. V. G. Arkhipkin and I. V. Timofeev, Phys. Rev. A **64**, 053811 (2001).
10. V. G. Arkhipkin and I. V. Timofeev, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 74 (2002) [JETP Lett. **76**, 66 (2002)].
11. S. E. Harris and L. V. Hau, Phys. Rev. Lett. **82**, 4611 (1999).
12. M. D. Lukin, P. H. Hemmer, and M. O. Scully, Adv. At. Mol. Opt. Phys. **42**, 347 (2000).
13. Fam Le Kien, Nguyen Hong Shon, and K. Hakuta, Phys. Rev. A **64**, 051803(R) (2001).
14. T. Rickes, L. P. Yatzenko, S. Steuerwald, *et al.*, J. Chem. Phys. **113**, 534 (2000).

15. R. Shumaker, in *Laser and Coherence Spectroscopy,* Ed. by J. I. Steinfeld (Plenum, New York, 1978; Mir, Moscow, 1982).
16. D. Grischkowsky, M. M. T. Loy, and P. F. Liao, Phys. Rev. A **12**, 2514 (1975).
17. I. A. Poluéktov, Yu. M. Popov, and V. S. Roĭtberg, Usp. Fiz. Nauk **114**, 97 (1974) [Sov. Phys. Usp. **17**, 673 (1974)].
18. L. P. Yatzenko, N. A. Vitanov, B. W. Shore, *et al.*, Opt. Commun. **204**, 413 (2002).
19. S. A. Myslivetz, A. K. Popov, T. Halfmann, *et al.*, Opt. Commun. **209**, 335 (2002); S. A. Myslivets, A. K. Popov, V. V. Kimberg, and T. F. George, in *Modern Topics in Chemical Physics*, Ed. by T. F. George, X. Sun, and G. P. Zhang (Res. Signpost, Trivandrum, India, 2002), p. 163.
20. E. A. Korsunsky, T. Halfmann, J. P. Marangos, *et al.*, quant-ph/0209161.
21. V. S. Butylkin, A. E. Kaplan, Yu. G. Khronopulo, and E. I. Yakubovich, *Resonance Interactions of Light with Substance* (Nauka, Moscow, 1977), p. 351.
22. C. Dorman, I. Kucukkara, and J. P. Marangos, Phys. Rev. A **61**, 013802 (2000).
23. S. A. Akhmanov, V. A. Vysloukh, Yu. G. Khronopulo, and A. S. Chirkin, *The Optics of Femtosecond Pulses* (Nauka, Moscow, 1988), p. 310.

Translated by A. Isaakyan

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 722–737.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 803–819.
Original Russian Text Copyright © 2003 by Sazonov.

ATOMS, SPECTRA, RADIATION

Resonant Transparency Effects in an Anisotropic Medium with a Permanent Dipole Moment

S. V. Sazonov
Kaliningrad State University, Kaliningrad, 236041 Russia
e-mail: nst@alg.kaliningrad.ru

Received March 14, 2003

Abstract—The propagation of a two-component laser pulse in an optically uniaxial medium is investigated under the conditions of the Zakharov–Benney resonance (viz., resonance of long and short waves). The short-wave ordinary component of the pulse, which is in resonance with the atomic subsystem, effectively generates a video pulse of the extraordinary wave (long-wave component). The latter dynamically detunes the ordinary pulse from the resonance and causes its phase modulation due to nonzero diagonal matrix elements of the dipole moment. An approximate operator approach is proposed for solving constitutive equations for the density matrix, which is equivalent to the asymptotic WKB method and makes it possible to reduce the analysis to solving a system of nonlinear wave equations for both components of the pulse. The possibility an extraordinary wave video pulse being generated with the help of a quasimonochromatic ordinary pulse with a longer wavelength. It is shown that, when the ordinary component dominates, the self-induced transparency mode is realized; in the opposite limit, the effect known as extraordinary transparency takes place. Solitary pulses corresponding to the latter case experience a decrease in the velocity of propagation, which is similar to that observed for self-induced transparency and practically do not change the population of quantum levels. Physical situations reducing the initial system of constituent and wave equations to familiar integrable models are analyzed. *© 2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The phenomenon of self-induced transparency (SIT) discovered in [1, 2] has been studied comprehensively both from the physical and the mathematical viewpoints [3, 4]. Quantum states in anisotropic media without a symmetry relative to spatial inversion do not possess a definite parity. As a result, the diagonal matrix elements of the dipole moment in such states differ from zero. Consequently, anisotropic molecules may have a permanent dipole moment when these molecules are in their energy eigenstates. The same applies to electronic states of anisotropic media such as crystals and quantum wells in semiconductors [5, 6]. The latter case should be treated separately since modern technologies make it possible to grow semiconductors containing quantum wells with preset properties [6, 7]. The SIT effect was studied in [5–9] for media possessing a permanent dipole moment in the framework of the scalar model, in which one electric field component for a pulse differs from zero. The approach proposed in [5, 6] does not require the application of the approximation of slowly varying amplitudes and phases (SVAP) in the constitutive and wave equations. In this case, stringent constraints are imposed on dipole moments and transition frequencies, which make it possible to reduce the systems of constitutive and wave equations to integrable models.

In an optically uniaxial medium, a light pulse bifurcates into ordinary and extraordinary components that can interact with quantum transitions. The presence of a permanent dipole moment at the nonlinear stage of such an interaction may play a decisive role. This study is devoted to analysis of the nonlinear stage of the resonance interaction of two-component pulses with optically uniaxial media.

The article is constructed as follows. In Section 2, the initial model is formulated and a system of equations of the Maxwell–Bloch type is derived for a pulse propagating at right angles to the optical axis. In this case, the ordinary component has a clearly manifested resonance frequency, while the extraordinary component is a video pulse. The next section is devoted to the derivation of a system of coupled nonlinear wave equations for the two components of the pulse on the basis of an approximate solution to the constitutive equations with the help of the operator version of the WKB asymptotic method, which is equivalent (see below) to disregarding the “dark” component of the dipole moment. In Section 4, analytic and numerical solutions to the system of wave equations in the form of solitary traveling formations including the ordinary envelope pulse and the extraordinary video pulse are considered and analyzed. It is shown that different quantitative relations between the amplitudes of the components correspond to different physical mechanisms of optical transparency of a resonant medium. Section 5 is devoted to analysis of situations leading to integrable models. The results of the study are summarized in Conclusions.

2. FORMULATION OF THE MODEL AND THE SYSTEM OF MAXWELL–BLOCH EQUATIONS FOR AN OPTICALLY UNIAXIAL MEDIUM

Let us suppose that a laser pulse propagates in a uniaxial medium along the x axis perpendicular to the optical z axis. Ordinary E_o and extraordinary E_e components of the electric field $\mathbf{E}$ of the pulse are parallel to the y and z axes, respectively.

We assume that anisotropy is natural and is created by the internal electric field. In the electric field, the degeneracy of electron energy levels in the modulus of projection M of the total angular momentum is removed [8]. As a result, the π ($\Delta M = 0$) and doubly degenerate σ transitions ($\Delta M = \pm 1$), which are allowed in accordance with the electric-dipole selection rules, take place in the electron subsystem.

In the absence of a pulse, the wave function of an optical electron in an axially symmetric field can be represented in the form

$$\psi_{\mu M} = R_{\mu M}(r, z)\exp(iM\varphi), \tag{1}$$

where r, φ, and z are the components of the cylindrical system of coordinates and μ is the set of quantum numbers corresponding to the cylindrical symmetry.

Using relation (1), we can find expressions for the Cartesian components of vector $\mathbf{d}_{\mu\nu}$ of the dipole moment of transition $\mu \longleftrightarrow \nu$:

$$d^x_{\mu\nu} = i\frac{d_{\mu\nu}}{\sqrt{2}}\Delta M_{\mu\nu}, \quad d^y_{\mu\nu} = \frac{d_{\mu\nu}}{\sqrt{2}}|\Delta M_{\mu\nu}|,$$

$$d^z_{\mu\nu} = D_{\mu\nu}(1 - |\Delta M_{\mu\nu}|),$$

where

$$\Delta M_{\mu\nu} = M_\mu - M_\nu = 0, \pm 1,$$

$$d_{\mu\nu} = -\sqrt{2}\pi e\int_0^\infty r^2 dr \int_{-\infty}^\infty R_\mu(r, z)R_\nu(r, z)dz,$$

$$D_{\mu\nu} = -2\pi e\int_0^\infty r dr \int_{-\infty}^\infty z R_\mu(r, z)R_\nu(r, z)dz.$$

Accordingly, taking into account the fact that the longitudinal component E_z for the normal propagation relative to the optical axis is equal to zero, we obtain the following expression for the matrix elements of Hamiltonian $\hat{V}$ of the electric-dipole interaction:

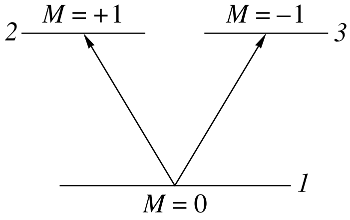

Fig. 1. Diagram of resonant quantum transitions in the case when the upper level is degenerate in the modulus of angular momentum component M.

$$V_{\mu\nu} = -\mathbf{d}_{\mu\nu}\cdot\mathbf{E} = -D_{\mu\nu}(1 - |\Delta M_{\mu\nu}|)E_e - \frac{d_{\mu\nu}}{\sqrt{2}}|\Delta M_{\mu\nu}|E_o. \tag{2}$$

The evolution of the state of an optical electron under the action of a laser pulse is described by the equations for the elements of density matrix $\hat{\rho}$,

$$\frac{\partial \rho_{\mu\nu}}{\partial t} = -i\left(\omega_{\mu\nu} - \frac{V_{\mu\mu} - V_{\nu\nu}}{\hbar}\right)\rho_{\mu\nu} - \frac{i}{\hbar}\sum_\lambda (V_{\mu\lambda}\rho_{\lambda\nu} - \rho_{\mu\lambda}V_{\lambda\nu}), \tag{3}$$

where $\hbar$ is the Planck constant and $\omega_{\mu\nu}$ is the frequency of the transition $\mu \longleftrightarrow \nu$.

It follows from formula (2) that only diagonal elements of $\hat{V}$ containing the extraordinary electric field component of the pulse differ from zero: $V_{\mu\mu} = -D_{\mu\mu}E_e$. Thus, the extraordinary component induces quantum π transitions and is responsible, due to the permanent dipole moment, for the dynamic frequency shift both for π and for σ transitions. The ordinary component excites only σ transitions.

In the subsequent analysis, we confine ourselves to a two-level system forming a σ transition of frequency ω_0 with a doubly degenerate upper level (Fig. 1). In this case, system (3), taking into account relation (2) under the assumption that the duration τ_p of a pulse is smaller than all relaxation times, assumes the form

$$\frac{\partial \rho_{21}}{\partial t} = -i\left(\omega_0 + \frac{D_{11} - D_{22}}{\hbar}E_e\right)\rho_{21} + i\frac{d_{21}}{\sqrt{2}\hbar}E_o(\rho_{11} - \rho_{22}) - i\frac{d_{31}}{\sqrt{2}\hbar}E_o\rho_{23}, \tag{4}$$

$$\frac{\partial \rho_{31}}{\partial t} = -i\left(\omega_0 + \frac{D_{11} - D_{33}}{\hbar}E_e\right)\rho_{31} + i\frac{d_{31}}{\sqrt{2}\hbar}E_o(\rho_{11} - \rho_{33}) - i\frac{d_{21}}{\sqrt{2}\hbar}E_o\rho_{32}, \tag{5}$$

$$\begin{aligned}\frac{\partial\rho_{32}}{\partial t} &= -i\frac{D_{22}-D_{33}}{\hbar}E_e\rho_{32}\\ &+ i\frac{E_o}{\sqrt{2}\hbar}(d_{31}\rho_{12}-d_{21}\rho_{31}),\end{aligned} \tag{6}$$

$$\frac{\partial\rho_{11}}{\partial t} = i\frac{E_o}{\sqrt{2}\hbar}[d_{21}(\rho_{21}-\rho_{12})+d_{31}(\rho_{31}-\rho_{13})], \tag{7}$$

$$\begin{aligned}\frac{\partial\rho_{22}}{\partial t} &= -i\frac{d_{21}E_o}{\sqrt{2}\hbar}(\rho_{21}-\rho_{12}),\\ \frac{\partial\rho_{33}}{\partial t} &= -i\frac{d_{31}E_o}{\sqrt{2}\hbar}(\rho_{31}-\rho_{13}).\end{aligned} \tag{8}$$

It follows hence that the ordinary component in this case induces a quantum transition, while the role of the extraordinary component is reduced to dynamic modulation of the frequency of this transition and, in accordance with formula (6), to dynamic removal of the degeneracy of the upper level for $D_{22} \neq D_{33}$.

The ordinary and extraordinary polarization components P_o and P_e created by the given medium are defined by the expressions

$$P_o = N\left[\frac{d_{21}}{\sqrt{2}}(\rho_{21}+\rho_{21}^*)+\frac{d_{31}}{\sqrt{2}}(\rho_{31}+\rho_{31}^*)\right], \tag{9}$$

$$P_e = N(D_{11}\rho_{11}+D_{22}\rho_{22}+D_{33}\rho_{33}), \tag{10}$$

where N is the atomic concentration.

We supplement relations (4)–(10) with the Maxwell equations for the electric field components of the pulse:

$$\frac{\partial^2 E_{o,e}}{\partial x^2}-\frac{n_{o,e}^2}{c^2}\frac{\partial^2 E_{o,e}}{\partial t^2} = \frac{4\pi}{c^2}\frac{\partial^2}{\partial t^2}(P_{o,e}-P_{o,e}^{(0)}). \tag{11}$$

Here, c is the velocity of light in vacuum; $n_o(n_e)$ is the ordinary (extraordinary) refractive index, created by the matrix of nonresonant dipole moments; and $P_o^{(0)}$ and $P_e^{(0)}$ are the values of the polarization components in zero external field, such that (see formulas (9) and (10)) $P_o^{(0)} = 0$, and

$$P_e^{(0)} = N(D_{11}W_1+D_{22}W_2+D_{33}W_3), \tag{12}$$

where W_j (j = 1, 2, 3) are the initial occupancies of the corresponding quantum levels.

It follows from formulas (10)–(12) that the extraordinary components of the polarizations and the field are determined by the dynamics of the occupancies of quantum levels (diagonal elements of $\hat{\rho}$). Consequently, in contrast to the ordinary component, the extraordinary component has no carrier frequency.

Taking into account the last remark, we will use in formulas (4)–(9) and (11) (in the equations for (E_o)) the SVAP approximation with the help of the representations

$$\begin{aligned}\rho_{21} &= R_{21}\exp[-i(\omega t-kx+\varphi)],\\ \rho_{31} &= R_{31}\exp[-i(\omega t-kx+\varphi)],\\ E_o &= 2\mathscr{E}_o\cos(\omega t-kx+\varphi),\end{aligned} \tag{13}$$

where ω and $k = \omega n_o/c$ are the carrier frequency and the wave number of the ordinary component, while amplitudes R_{21}, R_{31}, $\mathscr{E}_o$, and phase φ are slowly varying functions in the standard sense [9].

Substituting expressions (13) into formulas (4)–(8) and disregarding the terms oscillating at frequency 2ω, we obtain a system of equations coinciding in form with Eqs. (4)–(8) correct to the substitutions

$$\rho_{21} \longrightarrow R_{21}, \quad \rho_{31} \longrightarrow R_{31}, \quad E_o \longrightarrow \mathscr{E}_o,$$

$$\omega_0 \longrightarrow \Delta+\frac{\partial\varphi}{\partial t} = \omega_0-\omega+\frac{\partial\varphi}{\partial t},$$

this system can be written in the form

$$\frac{\partial\hat{R}}{\partial t} = i[\hat{\Omega},\hat{R}]. \tag{14}$$

Here, Hermitian matrices $\hat{R}$ and $\hat{\Omega}$ have the form

$$\hat{R} = \begin{pmatrix}\rho_{33} & \rho_{32} & R_{31}\\ \rho_{32}^* & \rho_{22} & R_{21}\\ R_{31}^* & R_{32}^* & \rho_{11}\end{pmatrix},$$

$$\hat{\Omega} = \begin{pmatrix}-\frac{1}{2}\left(\Delta+\frac{\partial\varphi}{\partial t}\right)+\Omega_{33} & 0 & \Omega_{31}\\ 0 & -\frac{1}{2}\left(\Delta+\frac{\partial\varphi}{\partial t}\right)+\Omega_{22} & \Omega_{21}\\ \Omega_{31} & \Omega_{21} & \frac{1}{2}\left(\Delta+\frac{\partial\varphi}{\partial t}\right)+\Omega_{11}\end{pmatrix}, \tag{15}$$

$$\Omega_{\mu\mu} = \frac{D_{\mu\mu}E_e}{\hbar}, \quad \mu = 1, 2, 3,$$

$$\Omega_{\mu 1} = \frac{d_{\mu 1}\mathscr{E}_o}{\sqrt{2}\hbar}, \quad \mu = 2, 3.$$

Substituting expressions (13) into formulas (9) and (11) and neglecting the second derivatives of the slowly varying amplitude and phase as well as the products of their first derivatives, we obtain

$$\frac{\partial \mathscr{E}_o}{\partial x} + \frac{n_o}{c}\frac{\partial \mathscr{E}_o}{\partial t} = i\frac{\pi N\omega}{\sqrt{2}cn_o}\sum_{\mu=2}^{3} d_{\mu 1}(R_{\mu 1} - \text{c.c.}), \tag{16}$$

$$\mathscr{E}_o\left(\frac{\partial \varphi}{\partial x} + \frac{n_o}{c}\frac{\partial \varphi}{\partial t}\right) = -\frac{\pi N\omega}{\sqrt{2}cn_o}\sum_{\mu=2}^{3} d_{\mu 1}R_{\mu 1} + \text{c.c.} \tag{17}$$

Let us suppose that the concentration of resonant σ transitions is so small that the following inequality holds:

$$\Lambda \equiv \frac{\pi d^2 N}{\hbar\omega} \ll 1.$$

In this case, Eq. (11) for E_e can be reduced in the order of the derivative with the help of the approximation of unidirectional propagation [10–12]. Disregarding the right-hand side of Eq. (11) in the zeroth approximation (in view of the smallness of N), we retain only the wave propagating to the right (to the bulk of the medium) in the solution to the obtained homogeneous equation: $E_e = E_e(t - n_e x/c)$. Then we take into account the right-hand side, deforming the given solution by introducing the dependence of E_e on the "slow" coordinate $\zeta = \Lambda x$ also: $E_e = E_e(\tau_e, \zeta)$, where $\tau_e = t - n_e x/c$ is the "local" time. Neglecting the terms on the order of Λ^2, we then obtain

$$\frac{\partial}{\partial t} = \frac{\partial}{\partial \tau_e}, \quad \frac{\partial}{\partial x} = -\frac{n_e}{c}\frac{\partial}{\partial \tau_e} + \Lambda\frac{\partial}{\partial \zeta},$$

$$\frac{\partial^2}{\partial x^2} \approx \frac{n_e^2}{c^2}\frac{\partial^2}{\partial \tau_e^2} - \frac{2\Lambda n_e}{c}\frac{\partial^2}{\partial \tau_e \partial \zeta}.$$

Using these expressions as well as formulas (10)–(12), integrating with respect to τ_e, and returning to the initial independent variables x and t, we obtain

$$\frac{\partial E_e}{\partial x} + \frac{n_e}{c}\frac{\partial E_e}{\partial t} = -\frac{2\pi N}{n_e c}\frac{\partial}{\partial t}\sum_{\mu=1}^{3} D_{\mu\mu}(\rho_{\mu\mu} - W_\mu). \tag{18}$$

It should be emphasized that, in contrast to Eq. (16), the quantity E_e in this equation is not the envelope, but the total extraordinary component of the field of the pulse (it was noted above that E_e has no carrier frequency).

Equations (14)–(18) form a self-consistent nonlinear system describing the propagation of a pulse whose ordinary component resonantly interacts with σ transitions in an optically uniaxial medium, while the extraordinary component causes modulation of the resonance frequency of the medium.

Following [6], we first assume that $d_{31} = d_{21} = d$ and $D_{33} = D_{22} \neq D_{11}$. In this case, the difference between transitions $1 \longleftrightarrow 2$ and $1 \longleftrightarrow 3$ disappears and we actually arrive at a two-level system. Indeed, from expressions (14) and (15), under the conditions specified above, we have

$$R_{21} = R_{31}, \quad \rho_{33} = \rho_{22}, \quad \rho_{32} - \rho_{22} = \text{const} = -W_2.$$

As a result, Eqs. (14)–(18) are transformed into a Maxwell–Bloch system for a medium with an optical symmetry:

$$\frac{\partial u}{\partial t} = -\left(\Delta + \frac{\partial \varphi}{\partial t} + \Omega_e\right)v, \tag{19}$$

$$\frac{\partial v}{\partial t} = \left(\Delta + \frac{\partial \varphi}{\partial t} + \Omega_e\right)u + \Omega_o w, \tag{20}$$

$$\frac{\partial w}{\partial t} = -\Omega_o v, \tag{21}$$

$$\frac{\partial \Omega_o}{\partial x} + \frac{n_o}{c}\frac{\partial \Omega_o}{\partial t} = \alpha v, \tag{22}$$

$$\Omega_o\left(\frac{\partial \varphi}{\partial x} + \frac{n_o}{c}\frac{\partial \varphi}{\partial t}\right) = -\alpha u, \tag{23}$$

$$\frac{\partial \Omega_e}{\partial x} + \frac{n_e}{c}\frac{\partial \Omega_e}{\partial t} = \beta\frac{\partial w}{\partial t}, \tag{24}$$

where

$$u = \frac{R_{21} + R_{21}^*}{\sqrt{2}}, \quad v = \frac{R_{21}^* - R_{21}}{i\sqrt{2}},$$

$$w = \frac{2\rho_{22} - \rho_{11} - W_2}{2}, \quad \Omega_o = \frac{2d\mathscr{E}_o}{\hbar}, \quad \Omega_e = \frac{DE_e}{\hbar},$$

$$\alpha = \frac{4\pi d^2 N\omega}{\hbar c n_o}, \quad \beta = \frac{2\pi D^2 N}{\hbar c n_o},$$

and $D = D_{11} - D_{22}$ (it is exactly the quantity D that will be referred to as the permanent dipole moment in the subsequent analysis).

System (19)–(24) generalizes the "scalar" model considered in [5] to the case when the field has two components (ordinary and extraordinary).

It should be emphasized once again that the ordinary component Ω_o of the Rabi frequency is determined by the envelope $\mathcal{E}_o$ of the ordinary component of the field of the pulse, while the extraordinary component Ω_e is determined by the component E_e of the field itself. Field equations (22) and (24) can also be reduced from the second to first derivatives for different reasons and with the help of different approximations (SVAP and unidirectional propagation, respectively). The meaning of the former approximation is that a large number of optical vibrations are accommodated under envelope $\mathcal{E}_o$, while the meaning of the latter approximation lies in the slow variation of the profile of E_e in a concomitant system of coordinates due to the low concentration of resonance centers (for this reason, the approximation of unidirectional propagation is sometimes referred to as the approximation of a slowly varying profile [13]). In this case, the extraordinary component, in contrast to the ordinary component, generally has no carrier frequency; i.e., this component is a video pulse or, according to the generally adopted terminology [14–17], an extremely short pulse. As can be seen from Eq. (24) and the expression for β, component E_e appears due to the presence of a permanent dipole moment for the uniaxial medium.

3. WAVE EQUATIONS
IN THE STRONG NONLINEARITY REGIME

In order to solve constituent equations (19)–(21) for $\Delta = 0$, we write this system of equations in matrix form,

$$\frac{\partial \boldsymbol{r}}{\partial t} = \hat{A}\boldsymbol{r}, \tag{25}$$

where $\boldsymbol{r} = (w, u, v)^T$ and

$$\hat{A} = \begin{pmatrix} 0 & 0 & -\Omega_o \\ 0 & 0 & -\tilde{\Omega}_e \\ \Omega_o & \tilde{\Omega}_e & 0 \end{pmatrix}. \tag{26}$$

Here,

$$\tilde{\Omega}_e = \Omega_e + \frac{\partial \varphi}{\partial t}.$$

The system written in this form coincides with the equations for the probability amplitudes of a three-level system, in which transitions that occur under the action of resonance fields form a λ diagram [18]. Following this analogy, we carry out the transformation of rotation through angle γ in the (w, u) plane around the v component of vector $\mathbf{r}$: $\tilde{\mathbf{r}} = \hat{L}\boldsymbol{r}$, $\hat{\tilde{\hat{A}}} = \hat{L}\hat{A}\hat{L}^{-1}$, where $\tilde{\mathbf{r}} = (\tilde{w}, \tilde{u}, v)$ and

$$\hat{L} = \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then Eq. (25) assumes the form

$$\frac{\partial \tilde{\mathbf{r}}}{\partial t} + \hat{L}\frac{\partial}{\partial t}(\hat{L}^{-1})\tilde{\mathbf{r}} = \hat{\tilde{\hat{A}}}\tilde{\mathbf{r}}.$$

Choosing angle γ so that $\tan\gamma = \tilde{\Omega}_e/\Omega_o$, we write the last equation in the form of the system

$$\begin{gathered} \frac{\partial \tilde{w}}{\partial t} = -\Omega v + \Omega_{-}\tilde{u}, \\ \frac{\partial \tilde{u}}{\partial t} = -\Omega_{-}\tilde{w}, \quad \frac{\partial \tilde{v}}{\partial t} = \Omega\tilde{w}, \end{gathered} \tag{27}$$

where

$$\Omega = \sqrt{\Omega_o^2 + \tilde{\Omega}_e^2},$$

$$\Omega_{-} = \frac{\partial\gamma}{\partial t} = \left(\frac{\Omega_o\partial\tilde{\Omega}_e}{\partial t} - \frac{\tilde{\Omega}_e\partial\Omega_0}{\partial t}\right)\frac{1}{\Omega^2}.$$

Analogously to the theory of "dressed" quantum states in a three-level system [18, 19], we will call v and

$$\tilde{u} = u\cos\gamma - w\sin\gamma = \frac{\Omega_o}{\Omega}u - \frac{\tilde{\Omega}_e}{\Omega}w, \tag{28}$$

"light" and "dark" components of the dipole moment and Ω and Ω_{-} as light and dark Rabi frequencies.

Obviously, quantity Ω_{-} characterizes the degree of mismatching in the dynamics of the ordinary component Ω_o and the effective extraordinary component $\tilde{\Omega}_e$ of the field of the pulse. We assume that both components follow each other almost adiabatically so that the condition [18, 19] $\Omega_{-} \ll \Omega$ holds. Assuming now that $\Omega_{-} = 0$ in system (27), we can easily obtain the solution to this system for the initial conditions $u(-\infty) = v(-\infty) = 0$ and $w(-\infty) = w_\infty$, which, for the initial variables, has the form

$$\begin{gathered} u = -2w_\infty\frac{\tilde{\Omega}_e\Omega_o}{\Omega^2}\sin\frac{\theta}{2}, \quad v = w_\infty\frac{\Omega_o}{\Omega}\sin\theta, \\ w = w_\infty\left[1 - 2\left(\frac{\Omega_o}{\Omega}\right)^2\sin^2\frac{\theta}{2}\right], \end{gathered} \tag{29}$$

where

$$\theta = \int_{-\infty}^{t} \Omega dt'.$$

It can be seen from system (27) that, under the above assumption ($\Omega_- \approx 0$), we have $\partial \tilde{u}/\partial t \approx 0$; i.e., the dark dipole moment remains practically unchanged, although an external action takes place.

Substituting expressions (29) into the right-hand sides of Eqs. (22)–(24), we obtain

$$\frac{\partial\Omega_o}{\partial x} + \frac{n_o}{c}\frac{\partial\Omega_o}{\partial t} = -b_o\frac{\Omega_o}{\Omega}\sin\theta, \quad (30)$$

$$\frac{\partial\varphi}{\partial x} + \frac{n_o}{c}\frac{\partial\varphi}{\partial t} = -2b_o\frac{\Omega_e + \partial\varphi/\partial t}{\Omega^2}\sin^2\frac{\theta}{2}, \quad (31)$$

$$\frac{\partial\Omega_e}{\partial x} + \frac{n_e}{c}\frac{\partial\Omega_e}{\partial t} = b_e\frac{\partial}{\partial t}\left[\left(\frac{\Omega_o}{\Omega}\right)^2\sin^2\frac{\theta}{2}\right]. \quad (32)$$

Here, $b_o = -\alpha w_\infty$ and $b_e = -2\beta w_\infty$.

Let us make the following important remark that will be used in the subsequent analysis of a more complex case. Solutions (29) obtained here using the “adiabaticity” criterion $\Omega_- \ll \Omega$ coincide exactly with analogous solutions to a system of form (25), obtained in [20] with the help of the WKB operator technique [21, 22]. For this reason, we can speak of the formal identity of these two approaches. The common property of the solutions obtained using the WKB asymptotic method is that the coefficients of periodic functions vary with time much slower than these functions [21, 22]. Under the adiabaticity condition, ratios of the form $\tilde{\Omega}_e/\Omega$ and, hence, the coefficients of the sines in expressions (29) are slowly varying functions of time. This circumstance is one more argument in favor of the above statement.

Let us now suppose that $d_{31} \neq d_{21}$. In this case, system (14), (15) can be solved approximately with the help of the above-mentioned operator version of the WKB method [20, 23]. We can write the solution to this system of equations in the form

$$\hat{R}(t) = \hat{U}\hat{R}(-\infty)\hat{U}^+, \quad (33)$$

where $\hat{U}$ is a unitary evolution operator.

Using formulas (15), we can easily establish that matrix $\hat{\Omega}$ does not commute with itself at different instants; i.e., $[\hat{\Omega}(t), \hat{\Omega}(t')] \neq 0$ for $t \neq t'$. If, however, the action of a pulse is quite short, the change in elements of matrix $\hat{\Omega}$ over this time interval Δt is insignificant and we can speak of approximate commutativity of $\Omega(t)$ and $\Omega(t')$. In this case, the evolutions operator has the form [20, 23]

$$\hat{U} = \lim_{\substack{\|\hat{\Omega}\| \to \infty \\ \Delta t \to 0}} \exp\left(i\int_{-\infty}^{t}\hat{\Omega}dt'\right), \quad (34)$$

where $\|\ldots\|$ is the norm of the operator.

In expression (34), we take into account the fact that the amplitude of a pulse increases upon a decrease in its duration; consequently, system of equations (14), (15) assumes the form of linear equations with large variable coefficients. It is this circumstance that allows us to speak of applicability of the WKB approach [21].

If all the eigenvalues

$$\hat{\theta} \equiv \int_{-\infty}^{t}\hat{\Omega}dt'$$

of the matrix are different, the exponential function in expression (34) can be evaluated with the help of the Sylvester formula [24]

$$\exp(i\hat{\theta}) = \sum_j\prod_{k\neq j}\frac{\hat{\theta} - \lambda_k\hat{I}}{\lambda_j - \lambda_k}\exp(i\lambda_j), \quad (35)$$

where $\hat{I}$ is a unit matrix and $\{\lambda_k\}$ is the set of eigenvalues $\hat{\theta}$. Since $\|\hat{\Omega}\| \longrightarrow \infty$, eigenvalues λ_k also tend to infinity. The indeterminacies in front of the exponentials can be determined using the L’Hospital rule, assuming that, in the limit $\Delta t \longrightarrow \infty$,

$$\hat{\theta} \approx \hat{\Omega}\Delta t, \quad \lambda_k \approx p_k\Delta t \approx \int_{-\infty}^{t} p_k dt,$$

where $\{p_k\}$ is the eigenvalue spectrum of operator $\hat{\Omega}$. Then formulas (33)–(35) lead to the following approximate expressions for the evolution operator in the limit corresponding to the condition of applicability for the WKB method [20]:

$$\hat{U} = \sum_j \exp\left(i\int_{-\infty}^{t} p_j dt'\right)\prod_{k\neq j}\frac{\hat{\Omega} - p_k\hat{I}}{p_j - p_k}. \quad (36)$$

The equation for the eigenvalues p of matrix $\hat{\Omega}$ has the form

$$\left(\Omega_{33}-\frac{1}{2}\frac{\partial\varphi}{\partial t}-p\right)$$
$$\times\left[\left(\Omega_{22}-\frac{1}{2}\frac{\partial\varphi}{\partial t}-p\right)\left(\Omega_{11}+\frac{1}{2}\frac{\partial\varphi}{\partial t}-p\right)-\Omega_{21}^2\right]$$
$$-\Omega_{31}^2\left(\Omega_{22}-\frac{1}{2}\frac{\partial\varphi}{\partial t}-p\right) = 0.$$

In the general case, this is a third-order equation. However, if we preserve one of the above assumptions ($D_{33} = D_{22}$), the roots of this equations can easily be found:

$$p_1 = \Omega_{22}-\frac{1}{2}\frac{\partial\varphi}{\partial t}, \quad p_{2,3} = \frac{\Omega_{11}+\Omega_{22}\pm\Omega}{2}. \tag{37}$$

Here,

$$\Omega = \sqrt{\left(\Omega_{11}-\Omega_{22}+\frac{\partial\varphi}{\partial t}\right)^2+4(\Omega_{21}^2+\Omega_{31}^2)}$$
$$= \sqrt{\tilde{\Omega}_e^2+\Omega_o^2},$$
$$\tilde{\Omega}_e = \frac{DE_e}{\hbar}+\frac{\partial\varphi}{\partial t}, \quad \Omega_o = \frac{2d\varepsilon_o}{\hbar},$$

and

$$d^2 = \frac{d_{21}^2+d_{31}^2}{2}.$$

It should be noted that the expression for Ω_o is transformed into that considered above for $d_{21} = d_{31} = d$, while the expression for $\tilde{\Omega}_e$ remains unchanged since the assumption $D_{33} = D_{22}$ is preserved.

We assume that, prior to the action of a pulse ($t = -\infty$), only the ground state is populated; i.e., only the element $\hat{R}(-\infty)$ differs from zero: $\rho_{11}(-\infty) = 1$. Then formulas (33), (36), and (37) lead to

$$\rho_{11} = 1-\left(\frac{\Omega_o}{\Omega}\right)^2\sin^2\frac{\theta}{2},$$
$$\rho_{\mu\mu} = 4\left(\frac{\Omega_{\mu 1}}{\Omega}\right)^2\sin^2\frac{\theta}{2}, \tag{38}$$
$$R_{\mu 1} = i\frac{\Omega_{\mu 1}}{\Omega}\sin\theta+2\frac{\tilde{\Omega}_e\Omega_{\mu 1}}{\Omega^2}\sin^2\frac{\theta}{2}, \quad \mu = 2, 3,$$

where θ is defined as in relations (29).

It can easily be seen that, for $d_{31} = d_{21} = d$, solutions (38) after the appropriate introduction of variables u, v, and w are transformed into the solutions (29) obtained by the other method. Thus, the conclusion concerning the equivalence of the two approaches is confirmed again. The smallness of the dark Rabi frequency Ω_- as compared to the light frequency Ω indicates a slow variation of ratios of the type $\Omega_{\mu 1}/\Omega$ and $\tilde{\Omega}_e/\Omega$ as compared to the variation of the quantities $\Omega_{\mu 1}$, $\tilde{\Omega}_e$, and Ω themselves. It is these ratios that form the coefficients of the trigonometric functions in expressions (38), and the slow rate of their variation is in accordance with the general properties of the solutions obtained with the help of the WKB method.

Substituting expressions (38) into the right-hand sides of Eqs. (16)–(18), we obtain a system of nonlinear wave equations of the form (30)–(32), where we now have

$$b_o = \frac{\pi N\omega(d_{21}^2+d_{31}^2)}{\hbar c n_o}, \quad b_e = \beta.$$

Thus, analysis of the propagation of a pulse is reduced to analysis of the system of nonlinear wave equations (30)–(32).

In the absence of a permanent dipole moment, we have $b_e = 0$. Then, in accordance with relations (32) and (31), these equalities are preserved in the medium itself if $\Omega_e = \partial\varphi/\partial t = 0$ at the entrance to the medium. In this case, $\Omega = \Omega_o$ and Eq. (30) can be reduced to the sine-Gordon equation

$$\frac{\partial^2\theta}{\partial x\partial\tau} = -b_o\sin\theta, \tag{39}$$

where $\tau = t - n_o x/c$, which describes the well-known SIT mode [3, 4] for the exact resonance of the field with the medium.

On the other hand, the presence of a permanent dipole moment inevitably leads to the formation of an extraordinary component that has no carrier frequency, as well as to the phase modulation of the ordinary component of the pulse (see Eqs. (32) and (31)), which ensures the dynamic detuning of the latter component from the resonance with the atomic subsystem. Thus, the nonlinear interaction between the short-wave ordinary and long-wave extraordinary components of an optical pulse occurs through the medium. Obviously, this interaction is most effective when $n_o = n_e$ (see Eqs. (30)–(32)). This equality corresponds to the resonance condition for long and short waves (the Zakharov–Benney resonance [25, 26]). If the high-frequency component of the wave field is not in resonance with the atomic system, the latter is excited insignificantly. The Zakharov–Benney resonance mode in this case is described by the Zakharov system exhibiting a weak (power) nonlinearity [25]. The unidirectional version of the Zakharov equations is the Yadjima–Oikawa integrable system [17, 27, 28]. However, in our case, the field generally interacts strongly with the medium due to the resonance of the ordinary component with the atomic subsystem. For this reason, the nonlinearity of system (30)–(32) is of a clearly manifested (not a power) nature.

4. SELF-INDUCED AND EXTRAORDINARY TRANSPARENCIES

We will seek the solution to system (30)–(32) in the form of solitary steady-state running waves, assuming that quantities Ω_o, Ω_e, and φ are functions of argument $t - x/V$, where V is the velocity of propagation of these waves. Taking into account the fact that Ω_o, Ω_e, and $\theta \longrightarrow 0$ for $t \longrightarrow -\infty$, we obtain from Eqs. (30)–(32)

$$\Omega_o' = g_o \frac{\Omega_o}{\Omega} \sin\theta,$$

$$\varphi' = 2g_o(\Omega_e + \varphi')\left(\frac{1}{\Omega}\sin\frac{\theta}{2}\right)^2, \tag{40}$$

$$\Omega_e = -g_e\left(\frac{\Omega_o}{\Omega}\right)^2 \sin^2\frac{\theta}{2},$$

where

$$g_{o,e} = b_{o,e}\left(\frac{1}{V} + \frac{n_{o,e}}{c}\right)^{-1}$$

and the prime on Ω_o and φ indicates the derivative with respect to $t - z/V$. We first assume that the role of the extraordinary component is relatively small: $\Omega_o^2 \gg \tilde{\Omega}_e^2$. This approximation corresponds to the generation of a video pulse of the extraordinary wave when the quasi-monochromatic ordinary component is supplied to the input of the medium (see Eq. (32)). In this case, we can put in Eqs. (40) $\Omega = \theta' \approx \Omega_o$; after this, the first equation assumes the form

$$\theta'' = g_o \sin\theta$$

and has a solution of the type of a solitary wave,

$$\Omega_o = \Omega_{om}\operatorname{sech}\frac{t - x/V}{\tau_p}, \tag{41}$$

where $\Omega_{om} = 2/\tau_p$. In this case, the relation between the pulse duration and velocity V is expressed as

$$\frac{1}{V} = \frac{n_o}{c} + b_o\tau_p^2. \tag{42}$$

Substituting this expression into two other equations, we obtain

$$\Omega_e = \varphi' = -\Omega_{em}\operatorname{sech}^2\frac{t - x/V}{\tau_p}, \tag{43}$$

where

$$\Omega_{em} = \frac{b_e}{(n_o - n_e)/c + b_o\tau_p^2}. \tag{44}$$

It can be seen from formulas (38) that, in this approximation,

$$\rho_{11} = 1 - \left[\left(\frac{2\Omega_{em}}{\Omega_{om}}\right)^2 + \cosh^2\frac{t - x/V}{\tau_p}\right]^{-1}. \tag{45}$$

Solution (41)–(45) corresponds to the SIT mode. Indeed, when a soliton-like pulse (41) propagates in the resonance atomic subsystem, almost complete inversion takes place (since $(2\Omega_{em}/\Omega_{om})^2 \ll 1$), followed by the return to the initial state. The extraordinary component (42) generated due to the presence of a permanent dipole moment detunes the atoms from resonance (see formulas (4) and (5)) with the ordinary component, affecting the latter component. In accordance with the Le Chatelier–Brown principle, the pulse responds to this action by phase modulation, as if tending to remain in resonance with the medium. It can be seen from formula (43) that the compensation in question in the given approximation is complete: the dynamic reduction of the frequency of the atomic transition is accompanied by an equivalent synchronous reduction of the carrier frequency of the ordinary component of the pulse. Thus, the medium remains in resonance with the field of the ordinary wave, which explains its nearly complete inversion at the region of the pulse maximum.

Using the explicit expression for b_o, we can write the denominator of relation (44) in the form

$$\frac{(n_o - n_e) + (\Lambda/n_o)(\omega\tau_p)^2}{c}.$$

On the one hand, $\Lambda \ll 1$, while on the other hand, $(\omega\tau_p)^2 \gg 1$. For this reason, both terms in the numerator may have comparable values. For media with negative birefringence ($n_o - n_e > 0$), the amplitude of the extraordinary component of the pulse increases upon a decrease in the pulse duration, and vice versa. In the case of positive birefringence ($n_o - n_e < 0$), however, two versions are possible: for $(\Lambda/n_o)(\omega\tau_p)^2 > n_e - n_o$, the situation is similar to that considered above, while for $(\Lambda/n_o)(\omega\tau_p)^2 < n_e - n_o$, first, the extraordinary component changes its polarity and, second, the value of the amplitude of Ω_e decreases with τ_p. If both refractive indices of the matrix are identical ($n_o = n_e$), which corresponds to the Zakharov–Benney resonance, we have

$\Omega_{em} = b_e/b_o\tau_p^2$. In this case, the generation of the extraordinary video pulse is the most effective. For $(\Lambda/n_o)(\omega\tau_p)^2 > n_e - n_o$, the denominator of relation (44) acquires a singularity. However, in this case, the initial assumption concerning the predominance of the ordinary component is violated. For this reason, in the case of positive birefringence, the solution under investigation may have a physical meaning only far away from this singularity. For $|n_o - n_e|/c \ll b_o\tau_p^2$, the ratio

$$\left|\frac{\Omega_e}{\Omega_o}\right|^2 \sim \frac{(D/d)^4}{(\omega\tau_p)^2},$$

and the initial condition $\tilde{\Omega}_e^2 \ll \Omega_o^2$ holds well for the quasimonochromatic ordinary component. If $|n_o - n_e|/c \gg b_o\tau_p^2$, the above-mentioned condition is satisfied even better since, in accordance with expression (41), the amplitude of the ordinary component always increases upon a decrease in its duration, while the quantity Ω_{em} tends to the limiting value $cb_e/|n_o - n_e|$.

Let us now consider the opposite case, $\tilde{\Omega}_e^2 \gg \Omega_o^2$. In this case, $\Omega = \theta' \approx \tilde{\Omega}_e$. Multiplying the first equation in system (40) by Ω_o under these conditions, we have

$$(\Omega_o^2)' = \frac{2}{\theta'}g_o\Omega_o^2\sin\theta. \tag{46}$$

Summing the second equation with the third, we obtain

$$\theta' = \frac{1}{\theta'^2}(2g_o\theta' - g_e\Omega_o^2)\sin^2\frac{\theta}{2}.$$

Expression Ω_o^2 from this relation and substituting the obtained expression into formula (46), we get

$$\begin{aligned}&\left[\theta'\left(2g_o - \frac{\theta'^2}{\sin^2(\theta/2)}\right)\right]' \\ &= 2g_o\left(2g_o - \frac{\theta'^2}{\sin^2(\theta/2)}\right)\sin\theta.\end{aligned} \tag{47}$$

Choosing the ansatz

$$\theta' = B\sin\frac{\theta}{2}, \tag{48}$$

where B is the constant to be determined, we can write expression (47) in the form

$$\theta'' = 2g_o\sin\theta.$$

Differentiating Eq. (48) and comparing the result with the last equation, we obtain $B^2 = 8g_o$. Integration of Eq. (48) gives

$$\tilde{\Omega}_e = \pm\frac{2}{\tau_p}\operatorname{sech}\frac{t - z/v}{\tau_p}. \tag{49}$$

Here, the relation between V and τ_p can be expressed via Eq. (42) to within the substitution $b_o \longrightarrow 2b_o$.

Using the solution and the second equation in system (40) in the approximation $\Omega \approx \tilde{\Omega}_e$, we obtain

$$\Omega_e = 3\varphi' = \frac{3\tilde{\Omega}_e}{4};$$

after this, from the last equation in system (40), we obtain in the same approximation

$$\Omega_o = \sqrt{-\frac{8g_o\Omega_e}{g_e}}$$

(it can be seen that the values of Ω_e and $\tilde{\Omega}_e$ are negative). This gives

$$\Omega_o = \Omega_{om}\operatorname{sech}^{1/2}\frac{t - x/V}{\tau_p}, \tag{50}$$

$$\Omega_e = -\Omega_{em}\operatorname{sech}\frac{t - x/V}{\tau_p}, \tag{51}$$

$$\varphi' = -\varphi'_m\operatorname{sech}\frac{t - x/V}{\tau_p}, \tag{52}$$

where $\Omega_{em} = 3/2\tau_p$, $\varphi'_m = 1/2\tau_p$, and

$$\Omega_{om} = 2\frac{d}{D}\sqrt{\frac{3\omega}{\tau_p}\left(1 + \frac{n_o - n_e}{2b_oc\tau_p^2}\right)}.$$

Substituting Eq. (50) into (51) and (38), we obtain

$$\rho_{11} = 1 - \left(\frac{3\Omega_{om}}{4\Omega_{em}}\right)^2\operatorname{sech}\frac{t - x/V}{\tau_p}. \tag{53}$$

The expression for the velocity for $n_e = n_o$ can be written in the form

$$\frac{1}{V} = \frac{n_o}{c} + 2b_o\tau_p^2 = \frac{n_o}{c}\left[1 + \Lambda\left(\frac{3\omega}{n_o\Omega_{em}}\right)^2\right]. \tag{54}$$

In accordance with this expression and the approximation $(\Omega_{om}/\Omega_{em})^2 \ll 1$ under which it was derived, the occupancies of the quantum levels of the medium practically do not change upon the passage through the medium of a bound state of pulses of the form (50), (51). It can be seen from formula (28) that, in the limit considered here, the dark component of the dipole moment is $u' \approx -w$. According to its meaning, $u' \approx \text{const}$ in the WKB approximation and, hence, $w \approx \text{const}$, which is in accordance with relation (53).

In accordance with formulas (19), (20), (51), and (52), the amplitude of effective detuning of the ordinary component from the resonance with the atomic subsystem is given by

$$|\varphi' + \Omega_e|_{t = z/V} = \frac{2\Omega_{em}}{3}$$

and is higher than the ordinary Rabi frequency Ω_o at least by a factor of several units. Thus, the quasimonochromatic ordinary component of the pulse is effectively detuned from the resonance with the medium by the extraordinary component having no carrier frequency. Consequently, in this case a transparency mode different from the SIT is realized. Proceeding from the above analysis, we can refer to the physical phenomena underlying the solutions (49)–(53) as extraordinary transparency (ET). We emphasize in this way the key role of the video pulse of the extraordinary wave in the given effect. It should be noted that the total area A of a pulse of the form (49)–(52) is equal to 2π as in the case of SIT. Indeed,

$$A = \int_{-\infty}^{\infty} \Omega dt \approx \int_{-\infty}^{\infty} |\tilde{\Omega}_e| dt = 2\pi$$

(see Eq. (49)). On the other hand, the area corresponding to a pulse of the ordinary wave is given by

$$A_o \equiv \int_{-\infty}^{\infty} \Omega_o dt \ll 2\pi.$$

At the same time, the velocity of propagation of the pulse decreases, in accordance with relation (54), and becomes of the same order of magnitude as the velocity of SIT solitons. This suggests a certain analogy to the effect of electromagnetically induced transparency [29, 30], in which a sharp increase in the blooming of the medium at the center of the resonance absorption line is accompanied by a considerable decrease in the group velocity of the signal, which is much stronger than in the case of SIT. This effect is realized in a system of three-level atoms with a λ diagram of allowed transition and is initiated by an intense resonant pumping on the $2 \longleftrightarrow 3$ transition. In this case, all atoms remain in the ground state. In our case, however, the trapping of atoms in the ground state is attained with the help of a video pulse of the extraordinary wave, which emerges due to the presence of a permanent dipole moment in an anisotropic medium and propagating together with a quasimonochromatic ordinary pulse with a retarded velocity (54). On the one hand, $\Lambda \ll 1$, and on the other hand, $(\omega/\Omega_{em})^2 \sim (\omega\tau_p)^2 \gg 1$; consequently, the value of the velocity strongly depends on the parameters of the medium and of the pulse and may change over a wide range.

The retardation mechanism in the case of ET can be explained most easily on the basis of the spectral analysis.

In view of the quasimonochromaticity condition $\omega\tau_p \gg 1$ for the ordinary wave, we have $|\varphi'| \sim 1/\tau_p \ll \omega$; i.e., the spectral broadening associated with the phase modulation is not very strong and the SVAP approximation remains in force.

Integrating Eq. (52), we obtain

$$\varphi = -\arctan\exp\left(\frac{t - x/V}{\tau_p}\right).$$

Then, in accordance with Eqs. (13) and (50), we can write the expression for E_o in the form

$$\frac{dE_o}{\hbar} = \frac{\Omega_{om}}{\sqrt{2}}\operatorname{sech}\frac{\tau}{\tau_p}\left[\exp\left(-\frac{\tau}{2\tau_p}\right)\cos(\omega t - kx) + \exp\left(\frac{\tau}{2\tau_p}\right)\sin(\omega t - kx)\right], \quad (55)$$

where $\tau = t - x/V$ and $k = n_o\omega/c$.

Applying the Fourier transformation to Eq. (55),

$$G_o(\nu) = \int_{-\infty}^{\infty} \frac{dE_o}{\hbar} e^{i\nu\tau} d\tau,$$

averaging over x, and taking into account the quasimonochromaticity condition $\omega\tau_p \gg 1$, we obtain

$$|G_o(\nu)|^2 = \left(\frac{\pi\Omega_{om}\tau_p}{2}\right)^2 (1 + e^{-\eta}\operatorname{sech}\eta)\operatorname{sech}\eta, \quad (56)$$

where $\eta = \pi(\nu - \omega)\tau_p$.

Analyzing expression (51), we obtain the following relation for the Fourier transform of the extraordinary component:

$$|G_e(\nu)|^2 = \left|\int_{-\infty}^{\infty} \Omega_e \exp(i\nu\tau) d\tau\right|^2 = \frac{9\pi^2}{4}\operatorname{sech}^2\frac{\pi\nu\tau_p}{2}. \quad (57)$$

Dependence (56) is represented in Fig. 2, which can be interpreted as follows. At the developed stage of ET, high-frequency photons of the ordinary wave transfer a part of their energy (according to the Zakharov–Benney mechanism) to the low-frequency extraordinary component, experiencing as a result a red shift; it can be seen from Fig. 2 that the spectrum of the ordinary com-

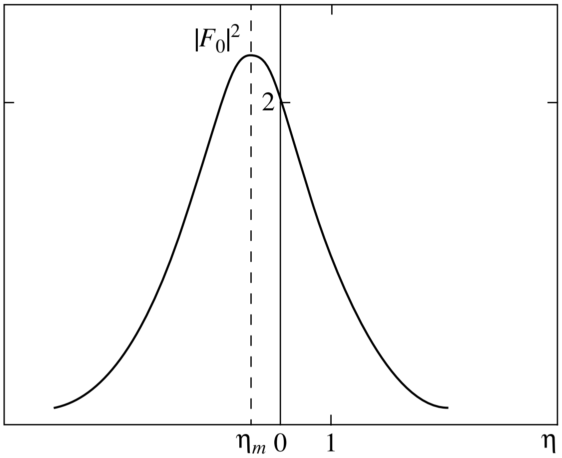

Fig. 2. Spectral energy distribution for the ordinary component of a pulse for extraordinary transparency; $|G_o|^2 = (\pi\Omega_{om}\tau_p/2)^2|F_o|^2$.

ponent is centered at a frequency of $\nu_m < \omega$, corresponding to $\eta = \eta_m < 0$. Analysis of formula (56) shows that $\eta_m = -0.38$. It follows hence that $\nu_m = \omega - 0.12/\tau_p$. In addition, the spectral width $\delta\omega_{ET}$ of the ET pulse (50) is larger than the spectral width $\delta\omega_{SIT}$ of the SIT soliton (41): $\delta\omega_{ET} = 0.74/\tau_p$, while $\delta\omega_{SIT} = 0.57/\tau_p$. Here, the spectral widths are determined by the values of frequencies at half the maximal values of $|G_o(\nu)|^2$.

Assuming that Ω_o, $\Omega \sim \exp(i\delta\omega t - i\delta kx)$ in Eq. (30) and linearizing this equation ($\sin\theta \approx \theta = -i\Omega/\delta\omega$), we arrive at the dispersion relation

$$k = \frac{n_o\delta\omega}{c} - \frac{b}{\delta\omega}.$$

Using this relation, we obtain the following expression for the group velocity:

$$\frac{1}{V} = \frac{d\delta k}{d\delta\omega} = \frac{n_o}{c} + \frac{b_o}{(\delta\omega)^2}.$$

Taking $\delta\omega \sim \delta\omega_{ET}$ as the spectral linewidth, we arrive at expression (54). Thus, a modification of the dispersion relation for the extraordinary wave in the vicinity of resonance leads to a decrease in the velocity of pulse propagation. It should be noted that, on the contrary, the generation of a video pulse of the extraordinary wave due to the Zakharov–Benney resonance away from the atomic–field resonance increases the velocity of the pulse propagation [17].

In [6], solitons which practically do not change the population of the quantum states are considered. However, in contrast to our case, the scalar (one-component) model is investigated in this publication. In [20], solitary acoustic pulses in a dilute paramagnet are analyzed, which do not change the population of quantum states either, but experience a decrease in velocity. The role of the longitudinal component of an acoustic pulse is similar to that of the extraordinary component of a laser signal [20]. On the other hand, the transverse component of an elastic pulse, as well as the ordinary component of an optical pulse, cause quantum transitions.

The amplitude of the extraordinary component of the pulses in question, as well as the velocity of propagation, increases monotonically upon a decrease in duration τ_p. The amplitude of the ordinary component in a medium with negative birefringence ($n_o > n_e$) exhibits the same property for $n_o = n_e$ also. In the case of positive birefringence ($n_o < n_e$), the value of Ω_{om} increases upon a decrease in τ_p down to

$$\tau_p = \tau_c \equiv \sqrt{\frac{3\hbar n_o(n_e - n_o)}{4\pi d^2 N\omega}},$$

where it attains its maximal value. As the pulse duration decreases further, the amplitude Ω_{om} decreases and vanishes at $\tau_p = \tau_c/\sqrt{3}$. The lower limit of τ_p corresponds to the maximal value of the amplitude of the extraordinary wave: $\Omega_{em} = 3\sqrt{3}/3\tau_c$. Let us find the condition under which the ET effect can be observed. For simplicity, we consider the case of Zakharov–Benney resonance: $n_o = n_e$. Then, using formulas (50)–(52), we can write the condition $\Omega_{om}^2 \ll \tilde{\Omega}_{em}^2$ in the form

$$1 \ll \omega\tau_p \ll \frac{1}{3}\left(\frac{D}{d}\right)^2. \qquad (58)$$

Let the left-hand side of this inequality correspond to the quasimonochromaticity condition for the ordinary component. In semiconducting GaAs/Al$_{0.14}$Ga$_{0.86}$ crystal with cutoff barriers of Al$_{03}$Ga$_{03}$As which have been synthesized recently, the value of $D/d \approx 7$ [6, 7], which is insufficient for satisfying both sides of inequality (58). On the other hand, the above-mentioned generation of a video pulse of the extraordinary wave with the help of an ordinary quasimonochromatic pulse, as well as the SIT mode for two-component pulses, can be successfully observed in crystals of this type. The condition $\Omega_o^2 \ll \tilde{\Omega}_e^2$, which is valid in the latter case, has the form $(\omega\tau_p)^2 \gg (D/d)^4$ for $n_e = n_o$ in accordance with formulas (41) and (44). For the above value of the ratio D/d, the latter conditions can easily be satisfied for a pulse of duration $\tau_p \sim 0.1$ ps, in which the ordinary component has a carrier frequency of $\omega \sim 10^{15}$ s^{-1}; for $d \sim 10^{-18}$ CGSE units, the intensity is given by

$$I_o \sim \frac{c}{4\pi}\left(\frac{\hbar}{d\tau_p}\right)^2 \sim 10^{10}\ \text{W/cm}^2.$$

In this case, the intensity of a video pulse of the extraordinary component has the form

$$I_e \sim \frac{c}{4\pi}\left(\frac{\hbar\Omega_e}{D}\right)^2 \sim I_o\left(\frac{D}{d}\right)^2\frac{1}{(\omega\tau_p)^2} \sim 10^6\ \text{W/cm}^2.$$

As regards the fulfillment of conditions (5) under which the ET effect can be observed, modern technologies make it possible to elevate the ratio D/d [6, 7]. For $D/d \approx 20$, we can speak of the fulfillment of condition (58) and, hence, of the possibility of observing ET under experimental conditions.

Let us now consider the solution to system (40) for an arbitrary ratio of the amplitudes of the ordinary and extraordinary components of a pulse. Summing the second and third equations, we obtain

$$\sin^2\frac{\theta}{2} = \frac{\tilde{\Omega}_e\Omega^2}{2g_o\tilde{\Omega}_e - g_e\Omega_o^2}. \tag{59}$$

Considering that

$$\Omega_o' = \frac{d\Omega_o}{d\theta}\theta' = \frac{\Omega d\Omega_o}{d\theta}$$

and integrating the first equation from system (40), we obtain

$$\int_0^{\Omega_o}\frac{\Omega^2 d\Omega_o}{\Omega_o} = g_o\int_0^{\theta}\sin\theta d\theta = 2g_o\sin^2\frac{\theta}{2}.$$

In the integration, we assumed natural conditions at infinity: $\Omega_o \longrightarrow 0$ as $\theta \longrightarrow 0$. Taking this into account, using Eq. (59), and differentiating with respect to Ω_o, we obtain

$$\frac{\Omega^2}{\Omega_o^2} = \frac{d}{d\Omega_o}\frac{2g_o\tilde{\Omega}_e\Omega^2}{2g_o\tilde{\Omega}_e - g_e\Omega_o^2}.$$

Introducing new dimensionless variables X and Y in accordance with relations $\Omega_o^2 = 4\sigma^2 X$ and $\tilde{\Omega}_e = -2\sigma Y$, where $\sigma = 2g_o/g_e$, we arrive at the following differential equation expressing the relation between the two components of the pulse in the traveling wave mode:

$$\frac{dY}{dX} = \frac{(X+Y^2)(Y+2X)^2 + 2XY^2(2Y-1)}{4X[X(X+Y^2) + Y^2(Y+2X)]}. \tag{60}$$

Using the first equation from system (40) and taking into account relation (59), as well as the new variables, we find that

$$\frac{dX}{d\xi} = \pm 2X\sqrt{\frac{2Y}{Y+2X}\left[1 - 2g\frac{Y(X+Y^2)}{Y+2X}\right]}, \tag{61}$$

where

$$\xi = \frac{t - x/V}{\tau_p}, \quad g = \frac{4g_o}{g_e^2},$$

and the minus and plus sign correspond to regions $X < X_{\max}$ and $X > X_{\max}$, respectively.

It should be noted that, under the Zakharov–Benney resonance condition, ($n_e = n_o$), we have

$$g = 4\left(\frac{d}{D}\right)^4(\omega\tau_p)^2.$$

Here, the pulse duration τ_p is introduced in accordance with relation (42).

Using relations (38), (40), and (59), we obtain

$$\begin{gathered}\Omega_o = 2\sigma\sqrt{X}, \quad \Omega_e = -4\sigma\frac{XY}{Y+2X}, \\ \varphi' = -2\sigma\frac{Y^2}{Y+2X}, \quad \rho_{11} = 1 - 2g\frac{XY}{Y+2X}.\end{gathered} \tag{62}$$

Thus, having solved system of equations (60) and (61), we can use expressions (62) to determine the physical parameters of the medium and of the pulse propagating in it.

The first of the cases considered here ($\tilde{\Omega}_e^2 \ll \Omega_o^2$) corresponds to the condition $Y^2 \ll X$, 1. In this case, Eq. (60) can be approximately written in the form

$$\frac{dY}{dX} = 1 + \frac{Y}{2X} - \frac{1}{2}\left(\frac{Y}{2X}\right)^2,$$

which has the solution of $Y = 2X$ (satisfying the condition $Y(0) = 0$) we are interested in. Using further formulas (61) and (62), we arrive at expressions (41)–(45) describing the SIT mode.

For $Y^2 \gg X$, 1 (or $\tilde{\Omega}_e^2 \gg \Omega_o^2$), expression (57) has the form

$$\frac{dY}{dX} = \frac{(2 + Y/X)^2 + 4Y/X}{4(3 + Y/X)};$$

it can readily be seen from formulas (58) and (59) that the solution $Y = 2X/3$ to this equation corresponds to the ET mode (relations (49)–(53)).

The intermediate case between SIT and ET cannot be analyzed analytically. In this case, Eq. (60) has to be solved numerically. The result of the corresponding analysis is shown is Fig. 3. Figure 4 shows the profiles of a running pulse and the variation of the population of the ground state for various ratios of amplitudes Ω_o and Ω_e. In the two limiting cases corresponding to SIT (a) and ET (c), good agreement is also observed between

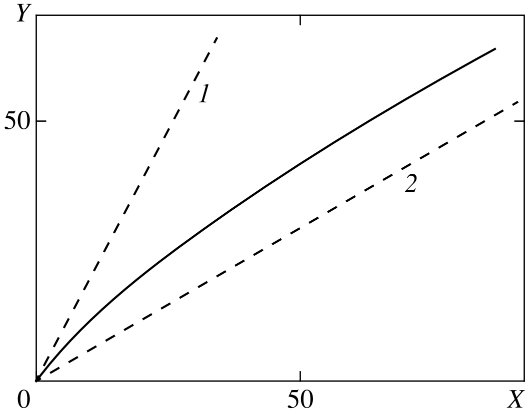

Fig. 3. Dependence $Y(X)$ plotted on the basis of the numerical solution of Eq. (60); the initial segment of straight line *1* (in the vicinity of $X = 0$) corresponds to the SIT mode, while straight line *2* for $X \gg 1$ corresponds to the ET mode.

the result of numerical calculations and the analytic solutions considered above. Figure 5 illustrates the structure of a two-component pulse taking into account phase modulation of the ordinary wave.

Thus, SIT is transformed into ET upon a smooth increase in the amplitude of the extraordinary pulse component. Phase modulation at the center of the ordinary component increases in this case.

A two component running pulse in the form of a bound state of the ordinary high-frequency component and a video pulse of the extraordinary component (which are mutually orthogonal) can be created by various methods before it is supplied to a resonant medium. One of the mechanisms was considered in Sections 3 of this article. It is connected with the generation of a video pulse of the extraordinary wave by a longer quasimonochromatic ordinary pulse due to nonzero diagonal matrix elements of the dipole moment. A natural question may arise concerning the stability of steady-state modes of pulse propagation. Let us suppose, for example, that both the high-frequency and the low-frequency components have two parts (ordinary and extraordinary). It was noted above that, for a pulse propagating at right angles to the optical axis, the ordinary part induces σ transitions and the extraordinary part, π transitions (see Eq. (2)).

The characteristic frequencies of resonant σ transitions are $\omega_\sigma \sim \omega \sim 10^{15}$ s^{-1}. On the other hand, the duration of a two-component pulse cannot be smaller than 10 fs, in accordance with the requirements of quasimonochromaticity of the high-frequency component. Consequently, the maximal spectral widths of the two components are on the order of 10^{14} s^{-1}. Since the spectrum of a video pulse has a peak at zero frequency (see Fig. 5), its spectrum does not contain Fourier components which are in resonance with σ transitions. For this reason, the low-frequency ordinary component interacts weakly with the medium and cannot noticeably affect the parameters of a two-component pulse experiencing retardation in the velocity of propagation due to a strong resonant coupling with σ transitions in the

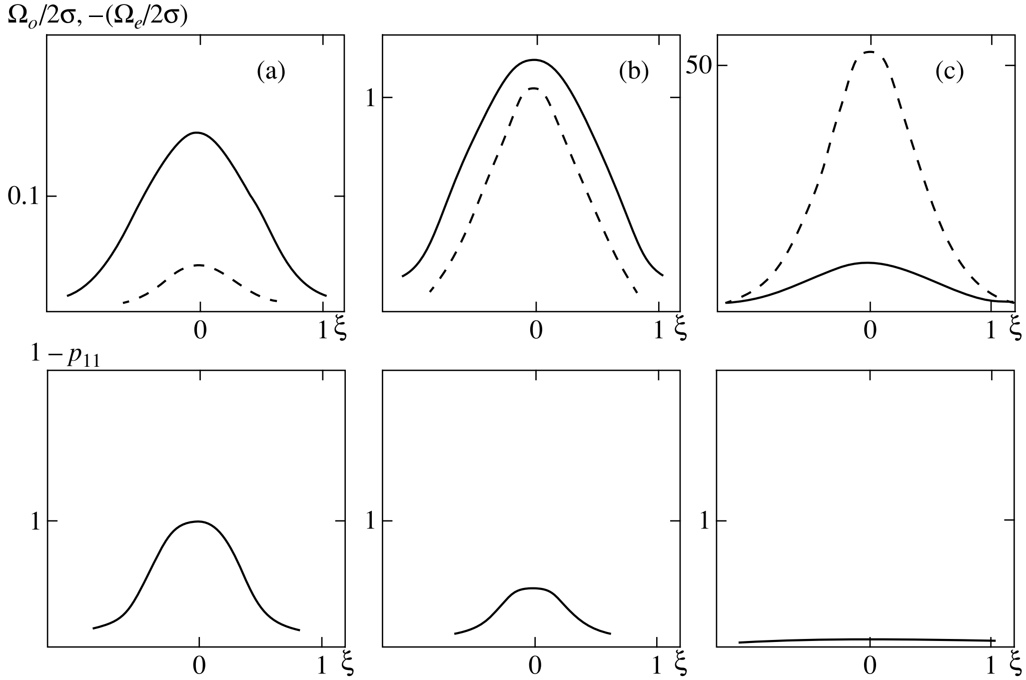

Fig. 4. Results of numerical solution of system (60)–(62): $g = 40.6$ (a), 0.38 (b), and 3.2×10^{-4} (c). Solid and dashed curves in the upper figures correspond to ordinary (Ω_o) and extraordinary ($-\Omega_e$) components, respectively.

medium. Similar considerations lead to the conclusion that the high-frequency extraordinary component cannot deform the above-mentioned steady-state running pulses either if its spectral width (which is equal in order of magnitude to the Rabi frequency) remains smaller than the frequency interval between the σ transitions considered by us here and the π transition nearest to it. The latter constraint appears as important since the extraordinary components can excite π transitions if their spectrum contains the corresponding resonance frequencies.

5. OTHER APPROACHES IN THE CASE OF EXACT ZAKHAROV–BENNEY RESONANCE

The case $n_e = n_o$ makes it possible to use a somewhat different approach to analysis of problems formulated in this article, which stems from [31, 32]. Multiplying Eq. (22) by $2\Omega_o$ and using Eqs. (21) and (24) after the integration, taking into account the conditions at infinity (Ω_o, $\Omega_e \longrightarrow 0$ for x, $t \longrightarrow \infty$), we obtain

$$\Omega_e = -\frac{\beta}{2\alpha}\Omega_o^2. \tag{63}$$

The subsequent analysis can be reduced to solving system (19)–(23) taking into account Eq. (63).

Phase modulation of the ordinary wave, which is defined by Eq. (23), in the SVAP approximation can be taken into account implicitly in the complex envelope $\mathcal{E}_o$ of field E_o, which is defined as

$$E_o = \mathcal{E}_o \exp[i(\omega t - kx)] + \text{c.c.},$$

as well as the envelopes for ρ_{21} and ρ_{31}.

In this case, for $d_{31} = d_{21} = d$ and $D_{33} = D_{22}$, we easily arrive at the system

$$\frac{\partial s}{\partial t} = i(\Delta + \Omega_e)s + i\Omega_o w, \tag{64}$$

$$\frac{\partial w}{\partial t} = \frac{i}{2}(\Omega_o^* s - \Omega_o s^*), \tag{65}$$

$$\frac{\partial \Omega_o}{\partial x} + \frac{n_o}{c}\frac{\partial \Omega_o}{\partial t} = -i\alpha s, \tag{66}$$

where $s = u + iv$, and the Rabi frequency $\Omega_o = 2d\mathcal{E}_o/\hbar$ is a complex quantity.

The equation for Ω_e coincides with Eq. (24), while expression (63) assumes the form

$$\Omega_e = -\frac{\beta}{2\alpha}|\Omega_o|^2.$$

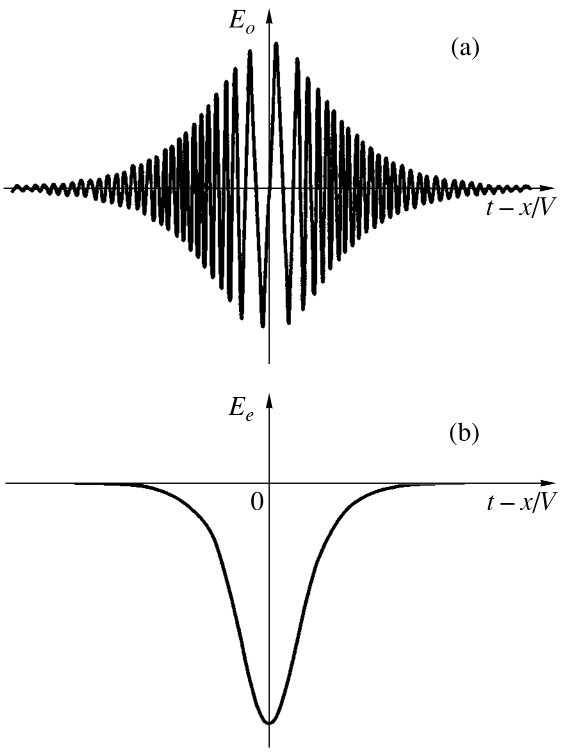

Fig. 5. Schematic diagram of the ordinary (a) and extraordinary (perpendicular to the ordinary) (b) components of a pulse, which propagate in the regime of the Zakharov–Benney resonance.

The substitution of this expression into Eq. (64) reduces system (64)–(66) to the integrable model analyzed in [32].

Along with expression (63), below we will use approximate system (30)–(32) derived from Eqs. (19)–(23) by eliminating material variables with the help of the WKB asymptotic method. It can easily be seen that, when relation (63) holds, Eqs. (31) and (32) for $n_e = n_o$ lead to the requirement $\partial\varphi/\partial t = \Omega_e$. Let us consider in greater detail the case of ET corresponding to the condition $\tilde{\Omega}_e \gg \Omega_o$. In this case,

$$\Omega \approx |\tilde{\Omega}_e| = \frac{\beta}{\alpha}\Omega_o^2.$$

Multiplying Eq. (30) by $(2\beta/\alpha)\Omega_0$ and passing to the local time τ, we obtain a sine-Gordon equation of the form (39) with the substitution $b_o \longrightarrow 2b_o$. In this case, the "area" is given by

$$\theta = \frac{\beta}{\alpha}\int_{-\infty}^{\tau}\Omega_o^2 dt' = \int_{-\infty}^{\tau}|\tilde{\Omega}_e|dt'.$$

It is interesting to note that we arrive at the same equation proceeding from formula (32) and taking into account (63) if we disregard the derivative of the "slow" variable $(\Omega_o/\Omega)^2$ on the right-hand side of Eq. (32) in accordance with the above-mentioned properties of the solutions obtained by using the WKB method. In this case, we have

$$\frac{\partial}{\partial t}\left[\left(\frac{\Omega_o}{\Omega}\right)^2\sin^2\frac{\theta}{2}\right] \approx \left(\frac{\Omega_o}{\Omega}\right)^2\frac{\partial}{\partial t}\left(\sin^2\frac{\theta}{2}\right) = \frac{\Omega_o^2}{2\Omega}\sin\theta$$

and relation (63) follows from Eqs. (30) and (32) automatically.

Using formula (63), we arrive again at Eq. (39) with the substitution $b_o \longrightarrow 2b_o$. Solving this equation, we obtain expressions of the type (49)–(52). The velocity is defined by relations (54). The difference between expressions (54) and (42) is in complete correspondence with the above-mentioned difference in the coefficients of the two sine-Gordon equations.

Thus, the approximations made above suggest that both modes, SIT and ET, are described by a sine-Gordon equation. Consequently, 2π pulses in the ET mode for $n_e = n_o$ with the correction, taking into account the approximations used, can be regarded as solitons as well as 2π pulses in the SIT mode. The difference between these two types of solitons is in the extent of their action on the medium: SIT solitons completely invert atoms with their subsequent return to the ground state, while ET solitons practically do not change the occupancies of the levels of the medium.

The general property of stability of solitons of the sine-Gordon equation makes it possible to speak of the stability of propagation of two-component pulses in the SIT and ET modes investigated here relative to small distortions of the profiles of both components.

On the other hand, the algorithm of the inverse scattering problem, which is developed for sine-Gordon equations, makes it possible to analyze the dynamics of formation of two-component pulses in both modes in the course of evolution of the input signal supplied to the medium.

Following [31], we can also propose an approach that does not employ the SVAP approximation, but assumes only a low density of the medium ($\Lambda \ll 1$). In this case, Eq. (11) for E_o can be reduced in the same way as was done for the extraordinary component. We use the system of constitutive equations (4)–(8) in the same way as for deriving Eqs. (64)–(66) for $d_{31} = d_{32} = d$ and $D_{33} = D_{22}$. This gives

$$\frac{\partial q}{\partial t} = i(\omega_0 + \Omega_e)q + i\tilde{\Omega}_o w, \tag{67}$$

$$\frac{\partial w}{\partial t} = -\tilde{\Omega}_o \mathrm{Im}\,q, \tag{68}$$

$$\frac{\partial\tilde{\Omega}_o}{\partial x} + \frac{n_o}{c}\frac{\partial\tilde{\Omega}_o}{\partial t} = -i\tilde{\alpha}\frac{\partial}{\partial t}\mathrm{Re}\,q, \tag{69}$$

where

$$\tilde{\alpha} = 2\omega\alpha, \quad \tilde{\Omega}_o = \frac{2dE_o}{\hbar}, \quad q = \sqrt{2}\rho_{21}.$$

The equation for Ω_e has the form of (24).

It was shown in [31] that, for $n_e = n_o$, system (67)–(69), (24) has an integral of motion connecting in our case the ordinary and extraordinary components with each other. Under the natural conditions at infinity, this integral can be written in the form

$$\tilde{\Omega}_o^2 = -\frac{\tilde{\alpha}}{2\beta}(\Omega_e^2 + 2\omega_o\Omega_e). \tag{70}$$

It follows hence that, in particular, $\Omega_e < 0$.

In the given case ($n_e = n_o$), system (67)–(69), (24) is integrable [31]. This system can be regarded as a generalization of the reduced system of Maxwell–Bloch equations [10, 11] to the case of an anisotropic medium.

6. CONCLUSIONS

The above analysis reveals a considerable influence of a video pulse of the extraordinary wave on the propagation of an ordinary envelope pulse in an optically uniaxial resonant medium. When a pulse propagates at right angles to the optical axis, the roles of the two components of the pulse are strictly differentiated: the ordinary wave excites resonant σ transitions, and the extraordinary wave dynamically shifts the frequency of these transitions. This results in the emergence of a new mechanism of resonant transparency, which can be described as follows. A pulse with a clearly manifested carrier frequency supplied to the input of a uniaxial resonant medium generates a video pulse of the extraordinary wave due to the permanent dipole moment; as a result, the medium is detuned from the resonance with the input pulse, and the latter pulse experiences phase modulation. This mechanism was referred to as extraordinary transparency to emphasize the decisive role in its formation of the extraordinary component generated in the medium. The occupancies of the quantum levels of the medium remain practically unchanged in this case, but the velocity of propagation of the pulse decreases in accordance with Eq. (54) and does not depend on the value of the permanent dipole moment. It should be noted that solitons that do not change the population of the quantum levels were considered in [6] in the framework of the scalar (one-component) model for the field, while acoustic two-component soliton-like pulses, which also preserve the occupancies, were investigated in [20]. In the framework of the two-component model considered in [6], both field components induce σ transitions in contrast to our case.

In our opinion, the two-component Maxwell–Bloch system of equations (19)–(24) is interesting itself from the point of view of its analytical structure.

The operator version of the WKB asymptotic method applied by us for solving the constitutive equations has made it possible, in contrast to previous studies, to analyze a medium with a degenerate upper level for $d_{31} \neq d_{21}$. In this case, the self-consistent system describing the dynamics of the field and the medium cannot be reduced to system (19)–(24), but, after the elimination of material variables ($d_{31} = d_{21}$ and $d_{31} \neq d_{21}$), the approximate wave equations have the form (30)–(32) in both cases.

Being integrable, the given system was obtained under a less stringent constraint imposed on the parameters of the medium as compared to those leading to integrable models.

The Maxwell equation for the extraordinary component is reduced to an equation with the first derivatives owing to the assumption of the low density of resonance centers. For our purposes, this assumption is not necessary. If we retain the corresponding equation in the initial (nonreduced) version,

$$\frac{\partial^2 \Omega_e}{\partial x^2} - \frac{n_e^2}{c^2}\frac{\partial^2 \Omega_e}{\partial t^2} = \frac{4\pi D^2 N}{\hbar c^2}\frac{\partial^2 w}{\partial t^2},$$

the solutions considered here have the form of running coupled formations of the envelope pulse of the ordinary wave and an extraordinary video pulse, and they preserve their form to within the substitution

$$g_e \longrightarrow \frac{c b_e}{2 n_o}\left(\frac{1}{V^2} - \frac{n_e^2}{c^2}\right)^{-1}.$$

However, in order to obtain more complex solutions using, for example, numerical methods, the reduced version of the equation for Ω_e in the form of (24) is obviously preferable.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 02-02-17710a).

REFERENCES

1. S. L. McCall and E. L. Hahn, Phys. Rev. Lett. **18**, 908 (1967).
2. S. L. McCall and E. L. Hahn, Phys. Rev. **183**, 457 (1969).
3. G. L. Lamb, Rev. Mod. Phys. **43**, 99 (1971).
4. A. I. Maimistov and A. M. Basharov, *Nonlinear Optical Waves* (Kluwer Academic, Dortrecht, 1999).
5. M. Agrotis, N. M. Ercolani, S. A. Glasgow, and J. V. Moloney, Physica D (Amsterdam) **138**, 134 (2000).
6. A. A. Zabolotskiĭ, Zh. Éksp. Teor. Fiz. **121**, 1012 (2002) [JETP **94**, 869 (2002)].
7. S. Kǒcinac, Z. Ikonić, and V. Milanovič, Opt. Commun. **140**, 89 (1997).
8. I. I. Sobel’man, *Atomic Spectra and Radiative Transitions* (Nauka, Moscow, 1977; Springer, Berlin, 1979).
9. L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (Wiley, New York, 1975; Mir, Moscow, 1978).
10. J. D. Gibbon, P. J. Caudrey, J. K. Eilbeck, and R. K. Bullough, J. Phys. A: Math. Gen. **6**, 1237 (1973).
11. P. J. Caudrey, J. K. Eilbeck, J. D. Gibbon, and R. K. Bullough, Lett. Nuovo Cimento **8**, 775 (1973).
12. S. V. Sazonov and E. V. Trifonov, J. Phys. B: At. Mol. Opt. Phys. **27**, L7 (1994).
13. M. B. Vinogradova, O. V. Rudenko, and A. P. Sukhorukov, *The Theory of Waves*, 2nd ed. (Nauka, Moscow, 1990).
14. S. A. Kozlov and S. V. Sazonov, Zh. Éksp. Teor. Fiz. **111**, 404 (1997) [JETP **84**, 221 (1997)].
15. A. I. Maĭmistov, Kvantovaya Élektron. (Moscow) **30**, 287 (2000).
16. A. M. Zheltikov, Usp. Fiz. Nauk **172**, 743 (2002) [Phys. Usp. **45**, 687 (2002)].
17. S. V. Sazonov and A. F. Sobolevskiĭ, Pis’ma Zh. Éksp. Teor. Fiz. **75**, 746 (2002) [JETP Lett. **75**, 621 (2002)].
18. J. H. Eberly, Quantum Semiclassic. Opt. **7**, 373 (1995).
19. M. Fleischhauer, Phys. Rev. A **54**, 794 (1996).
20. S. V. Voronkov and S. V. Sazonov, Zh. Éksp. Teor. Fiz. **120**, 269 (2001) [JETP **93**, 236 (2001)].
21. N. N. Moiseev, *Asymptotic Methods of Nonlinear Mechanics* (Nauka, Moscow, 1981).
22. A. H. Nayfeh, *Introduction to Perturbation Techniques* (Wiley, New York, 1981; Mir, Moscow, 1984).
23. A. Yu. Parkhomenko and S. V. Sazonov, Kvantovaya Élektron. (Moscow) **27**, 139 (1999).
24. V. A. Yakubovich and V. M. Starzhinskiĭ, *Parametric Resonance in Linear Systems* (Nauka, Moscow, 1987).
25. V. E. Zakharov, Zh. Éksp. Teor. Fiz. **62**, 1745 (1972) [Sov. Phys. JETP **35**, 908 (1972)].
26. R. K. Dodd, J. C. Eilbeck, J. Gibbon, and H. C. Morris, *Solitons and Nonlinear Wave Equations* (Academic, New York, 1982; Mir, Moscow, 1988).
27. N. Yadjima and M. Oikawa, Prog. Theor. Phys. **56**, 1719 (1976).
28. S. V. Sazonov, Zh. Éksp. Teor. Fiz. **118**, 20 (2000) [JETP **91**, 16 (2000)].
29. K. J. Boller, A. Imamoglu, and S. E. Harris, Phys. Rev. Lett. **66**, 2593 (1991).
30. O. Kocharovskaya, Phys. Rep. **219**, 175 (1992).
31. A. A. Zabolotskiĭ, Pis’ma Zh. Éksp. Teor. Fiz. **76**, 709 (2002) [JETP Lett. **76**, 607 (2002)].
32. A. A. Zabolotskiĭ, Phys. Lett. A **124**, 500 (1987).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 738–744.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 820–828.

The Accommodation of the Translational and Rotational Energy of a Gas in a Knudsen Flow Past a Thin Wire

A. K. Rebrov*, A. A. Morozov, M. Yu. Plotnikov,
N. I. Timoshenko, and A. V. Shishkin
Institute of Thermal Physics, Siberian Division, Russian Academy of Sciences,
pr. Akademika Lavrent'eva 1, Novosibirsk, 630090 Russia
**e-mail: rebrov@itp.nsc.ru*
Received March 12, 2003

Abstract—The paper presents the results obtained in determining the accommodation coefficients for the translational and rotational energy of gas molecules in a Knudsen flow past a thin wire. The method used was based on numerically solving the complete heat balance equation for a wire probe. The accommodation coefficients were determined for H_2, N_2, CH_4, and CO_2 on a gilded tungsten surface. For hydrogen with a quenched rotational energy, a negative accommodation coefficient of rotational energy was obtained due to the conversion of the rotational energy of incident molecules into the translational energy of reflected molecules. *© 2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The concept of accommodation (completeness of energy exchange between gas molecules and a surface) is important for studying nonequilibrium interactions between molecules and surfaces. Energy accommodation coefficients were introduced at the beginning of the past century, at the early stage of elucidating the features of processes in rarefied gases and establishing the concept of temperature jumps at the surface [1]. The classical equation for the energy accommodation coefficient is

$$\alpha = \frac{E_i - E_r}{E_i - E_w}, \qquad (1)$$

where E_i and E_r are the energies of incident and reflected molecules and E_w is the mean energy of reflected molecules corresponding to the temperature of the surface. The temperature jump value for a gas as a boundary condition for it near the surface in problems of mechanics of continua can easily be written via the accommodation coefficient. Over several decades, researchers have concentrated on translational energy accommodation. In the second half of the last century, scientists became interested also in the accommodation of internal energy. This interest was stimulated by the development of space technologies and the use of vacuum and vacuum–plasma processes as the material base for microelectronics and atomic technology, which initiated studies of strongly nonequilibrium processes in gases with internal degrees of freedom (rotational, vibrational, and electronic). This trend persists and has currently become even more pronounced [2–7].

The results obtained and the potential and prospects for studying interactions between molecules and surfaces cannot satisfy practical demands for creating a database sufficient for modeling the interaction of gas molecules with technical surfaces typical of various applications. This places emphasis on the development of methods for studying energy exchange on surfaces under the conditions identical or close to the particular requirements of physical experiments or technologies.

The use of the accommodation coefficient in form (1) presupposes equal energy accommodation on the surface for translational and internal energy. This assumption is not well grounded, because it is known that relaxation is faster for translational than for internal degrees of freedom [8]. In addition, this approach ignores possible energy redistribution over various degrees of freedom and their modes. A more exact description of energy exchange between molecules and surfaces should therefore use additional coefficients, which determine the fraction of energy transferred to the other degrees of freedom [9]. In several studies, four coefficients were introduced to describe not only energy exchange with a surface but also translational energy conversion to internal energy and vice versa [10]. The possibility of such processes can roughly be taken into account by the introduction of partial accommodation coefficients α_{tr} and α_{rot} for the translational and internal energy. These coefficients can be determined using the translational and internal energy, respectively, in (1).

Several different methods were used to experimentally determine the partial accommodation coefficients, including studies of ultrasonic surface oscillations in a rarefied gas [11], energy balance calculations from measured molecular velocities and level populations of incident and reflected molecules [12–14], recovery temperature measurements on a flat plate in a Knudsen flow [15], and studies of heat flow between a wire and a band in a rarefied gas [16]. More recently, the correctness of the last method was, however, questioned [17]. Theoretically, the partial accommodation coefficients were studied using quantum-mechanical calculations [18]. Both measurements and theoretical calculations show that the accommodation coefficient for translational energy can be substantially larger than the internal energy accommodation coefficient.

The method of a heated wire in a gas-dynamic Knudsen flow [19–23] occupies a special place in studies of energy exchange on surfaces and the determination of the partial accommodation coefficients. The distinguishing feature of this method is the possibility of studying energy exchange under the conditions of a supersonic flow around a body, which are characteristic of the dynamics of rarefied gases. Under these conditions, the potential and internal energies of incident molecules are substantially different. In addition, this method can be used in wide ranges of surface temperatures and gas flow energies.

Theoretical foundations for calculating energy accommodation coefficients in studies of heat transfer from a cylinder in a Knudsen flow were, in essence, developed in [24]. Thin wires are extensively used to measure total energy accommodation coefficients in a gas at rest at large Knudsen numbers [3, 5]. The possibility of determining total energy accommodation coefficients by probing a rarefied gas flow with a wire anemometer was discussed in [25], where simplified wire heat balance calculations were used in combination with other tools. It was shown in [19–23] that the idea of a complete analysis of heat balance for a cylinder in a Knudsen flow could be implemented in determining the partial accommodation coefficients for translational and internal energy. In these works, the accommodation coefficients were found using an analytic solution to a simplified heat balance equation suggested in [26]. Such simplifications, however, introduce substantial uncertainties into the analysis and are therefore inapplicable in the entire wire temperature range.

In this work, the translational and rotational energy accommodation coefficients on a surface were determined by numerically solving the heat balance equation taking into account the influence of the electric and thermophysical wire characteristics. This work continues studies initiated in [27]. Special attention is given to determining accommodation coefficients under the conditions of substantially different translational and rotational energies.

2. THE METHOD FOR DETERMINING ACCOMMODATION COEFFICIENTS

The heat balance for a wire of unit length heated by an electric current and placed into a gas is determined by the equation

$$\begin{aligned} cg\pi r_0^2\frac{\partial T_w}{\partial t} &= \frac{I^2\rho}{\pi r_0^2} + \lambda\pi r_0^2\frac{d^2T_w}{dx^2} \\ &+ 2\pi r_0 h(T_r - T_w) - 2\pi r_0\varepsilon\sigma(T_w^4 - T_\infty^4). \end{aligned} \qquad (2)$$

Here, x is the distance measured along the wire, T_w is the temperature of the wire, r_0 is the radius of the wire, c is the specific heat capacity, g is the density, I is the electric current strength, ρ is the specific electric resistance, λ is the heat conductivity coefficient, ε is the degree of blackness, σ is the Stefan–Boltzmann constant, T_∞ is the temperature of the surrounding surfaces, T_r is the recovery temperature of the wire, and h is the heat-transfer coefficient for the given flow parameters. It was assumed in the calculations that ρ, ε, and λ linearly depended on the temperature.

The boundary condition for (2) is the T_s temperature at the junctions between the wire and the holders. Correct determination of this temperature is one of the principal difficulties involved in solving the problem under consideration. Usually, holders are massive compared with the wire. If the temperature of the holder T_b at a large distance from the junction is known, the T_s temperature can be determined under the assumption that it depends on the heat flow from the wire to the holder and the heat resistance of the holder. Consider a cylindrical holder of radius r_b. An approximate estimate of T_s can be obtained under the assumption that heat transfer occurs as in a plane layer at small distances from the junction ($r \le r_0$) and as in a spherical layer at large distances ($r \ge r_0$). This gives

$$T_s \approx \left\{T_b + \frac{q}{2\pi\lambda_b}\left(\frac{1}{r_0} - \frac{1}{r_b}\right)\right\} + \frac{q}{\pi\lambda_b r_0},$$

where q is the heat flow from the wire to the holder (determined in the calculations using the temperature gradient at wire ends) and λ_b is the heat conductivity coefficient of the holder. If $r_b \gg r_0$, we obtain

$$T_s \approx T_b + \frac{3}{2}\frac{q}{\pi\lambda_b r_0}.$$

To solve (2), we must know the heat transfer coefficient h and the recovery temperature T_r. The kinetic theory of gases allows these values to be determined on convex surfaces in Knudsen flows [9]. We will use the equations for h and T_r for a cylinder situated normally with respect to the flow of a gas with internal degrees of

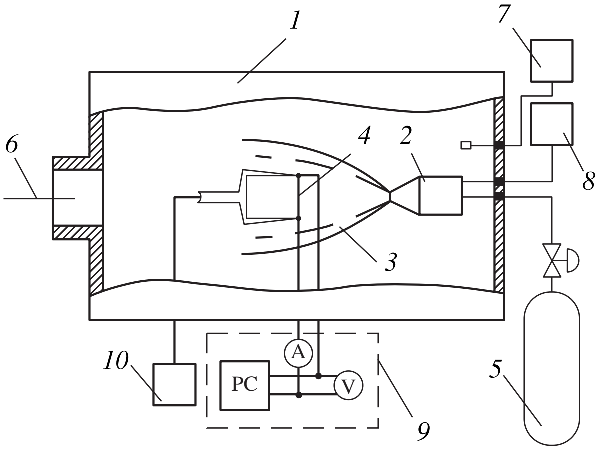

Fig. 1. Schematic diagram of the experimental setup: (*1*) vacuum chamber; (*2*) premix chamber; (*3*) gas jet; (*4*) wire probe; (*5*) gas feed; (*6*) vacuum stage; (*7*, *8*) pressure gauges; (*9*) measuring electronics; (*10*) power supply unit.

freedom [21] (these equations are derived in the Appendix), namely,

$$h = \frac{1}{4\sqrt{\pi}} nk\xi \exp\left(-\frac{s^2}{2}\right)\frac{U}{S}[I_0 + S^2(I_0 + I_1)], \qquad (3)$$

$$T_r = \frac{T_{\text{tr}}}{\xi}$$

$$\times\left\{\alpha_{\text{tr}}\left(2S^2 + 5 - \frac{1}{1 + S^2(1 + I_1/I_0)}\right) + \alpha_{\text{rot}} j \frac{T_{\text{rot}}}{T_{\text{tr}}}\right\}. \qquad (4)$$

Here, n, U, T_{tr}, and T_{rot} are the number density of particles, the velocity, and the translational and rotational gas flow temperatures, respectively;

$$S = \frac{U}{(2kT_{\text{tr}}/m)^{1/2}}$$

is the velocity ratio; m is the molecular weight; k is the Boltzmann constant;

$$\xi = 4\alpha_{\text{tr}} + j\alpha_{\text{rot}}, \quad j = \frac{5 - 3\gamma}{\gamma - 1}$$

is the number of internal degrees of freedom; γ is the adiabatic index; and I_0 and I_1 are the zeroth- and first-order modified Bessel functions in the $S^2/2$ argument, respectively.

Note that, in a gas at rest ($U = 0$), (4) reduces to

$$T_r = T_{\text{tr}}$$

and the heat transfer coefficient in (3) then only depends on the complex $\xi = 4\alpha_{\text{tr}} + j\alpha_{\text{rot}}$. For this reason, the translational and internal energy accommodation coefficients can only be separated for a directed gas flow.

All accommodation coefficient calculations were performed by minimizing the difference between the calculated and experimental probe electric resistances. The probe was a thin wire heated by an electric current. First, the resistance of the wire R'_k was measured at various electric current values I_k ($k = 1, \ldots, N$). Next, (2) was numerically solved for all possible pairs of accommodation coefficients (α_{tr} and α_{rot}) and all current strengths I_k to determine the distribution of temperature $T_w(x) = T_w(x, \alpha_{\text{tr}}, \alpha_{\text{rot}}, I_k)$ along the wire. The temperature distribution obtained was used to calculate the integral electric resistance of the wire by the equation

$$R(\alpha_{\text{tr}}, \alpha_{\text{rot}}, I_k) = \frac{1}{\pi r_0^2}\int_0^L \rho(T_w(x, \alpha_{\text{tr}}, \alpha_{\text{rot}}, I_k))dx,$$

where L is the wire length. Next, the discrepancy between the calculated and experimental electric resistance values was found,

$$\Delta_k = R(\alpha_{\text{tr}}, \alpha_{\text{rot}}, I_k) - R'_k,$$

and the functional

$$F(\alpha_{\text{tr}}, \alpha_{\text{rot}}) = \sum_1^N \Delta_k^2.$$

was minimized. The α_{tr} and α_{rot} values corresponding to the minimum of $F(\alpha_{\text{tr}}, \alpha_{\text{rot}})$ were taken to be the sought accommodation coefficients. This approach takes into account the special features of temperature distribution along the wire length and decreases the influence of errors inevitable in measurements.

3. EXPERIMENTAL

Measurements were performed as schematically shown in Fig. 1, on a vacuum gas dynamic unit in which a rarefied gas flow was created by free expansion from a sonic nozzle $d_* = 6$ mm in diameter at a braking pressure of $P_0 = 1000$–2100 Pa and a gas temperature of $T_0 = 290$–300 K. The background gas pressure surrounding the flow in the vacuum chamber was 2.7 Pa. A wire probe was mounted along the flow axis at distances $D = (2$–$7)d_*$ from the nozzle. Under these conditions, the flow near the wire was free from the influence of the background gas [28]. At the distances under consideration, flow parameters along the flow axis corresponded to the expansion of a nonviscous gas [29].

The probe was a gilded tungsten wire 8.3 μm in diameter, which ensured that the flow was a Knudsen flow under the conditions specified above. The diameter of the wire was measured using an electron microscope. The wire was welded to stainless steel holders. The holders were cylinders 0.6 mm in diameter reduced to cones, and their diameter at the ends was 0.2 mm. The typical wire length L was 3–10 mm. The surface of the wire was periodically controlled during experiments. The electric current and voltage on wires were measured by instruments of accuracy class 0.1. The range of currents was selected depending on the probe length from the requirement that the highest wire temperature determined in calculations should not exceed 700 K.

Experiments in a free flow were preceded by the determination of the temperature dependence of the electric and thermophysical characteristics of the wire, including specific resistance ρ, heat conductivity coefficient λ, and degree of blackness ε. It was necessary to create a stable high-temperature contact between the conductor and copper clamps in these experiments. In addition, we had to ensure the possibility of accurate wire length measurements. For this purpose, wire ends were lead–tin-plated and, under a microscope, introduced into nickel capillaries, where they were calked. The length of the wire from the edge of one capillary to the edge of the other was measured at room temperature under a microscope accurate to 0.02 mm. The nickel capillaries were clamped into massive copper clamps in experiments on determining the thermophysical characteristics. The $\rho = \rho(T_w)$ dependence was determined from measurements in a thermostat in air at atmospheric pressure.

The degree of wire surface blackness as a function of temperature, $\varepsilon(T_w)$, was determined by measuring the electric resistance of a long wire (L = 100 mm) at a low gas pressure ($P = 10^{-2}$ Pa). Under these conditions, the fraction of heat transferred to the holders and gas was small compared with thermal radiation. The degree of blackness $\varepsilon(T_w)$ was adjusted by solving (2) and comparing the calculated $R = R(\varepsilon, I)$ and experimental $R' = R'(I)$ dependences. The degree of blackness determined this way was further used to refine the heat conductivity coefficient of the wire. For this purpose, we measured the electric resistance of a wire L = 10 mm long at a low gas pressure. Under these conditions, the heat flow from the wire to the holders predominated. As with the degree of blackness, the heat conductivity coefficient λ was adjusted by comparing the calculated $R = R(\lambda, I)$ dependence with the results of measurements. Next, the degree of blackness was determined more accurately taking into account the obtained heat conductivity.

The calibration procedures described above gave the following characteristics of the wire under examination:

$$\varepsilon = 0.05[1 + 0.00335(T_w - 273)],$$

$$\rho = 5.65 \times 10^{-8}[1 + 0.0041(T_w - 273)]\ \Omega\ \text{m},$$

$$\lambda = \begin{cases} 145 - 0.11(T_w - 273)\ \text{W/(m K)}, & T_w \le 523\ \text{K} \\ 117 - 0.06(T_w - 523)\ \text{W/(m K)}, & T_w \ge 523\ \text{K}. \end{cases}$$

4. RESULTS AND DISCUSSION

The method was subjected to approbation with monoatomic Ar and He gases [27]. In experiments with polyatomic gases (H_2, N_2, CH_4, and CO_2), we used the procedure for enumerating pairs of accommodation coefficients (α_{tr} and α_{rot}) described above. An example of the projection of the differences between the experimental and calculated electric resistance values,

$$\Delta = \frac{1}{N}\sum_1^N \Delta_k^2,$$

onto the (α_{tr}, α_{rot}) plane for a flow of methane is shown in Fig. 2.

The accommodation coefficients averaged over many experimental results obtained with wire probes of various lengths, at various gas pressures, and at various distances from the nozzle are listed in Table 1. The data reported by different research teams for accommodation coefficients are usually difficult to compare, because it is difficult to find results obtained under identical conditions with respect to many key parameters, namely, the energy of incident molecules, the interaction angle, the initial structure of the surface, the degree of adsorption, and the surface temperature. This problem has not been completely defined for surfaces of technological interest. Table 2 contains some literature data on accommodation coefficients obtained at similar interaction energies. A comparison of Tables 1

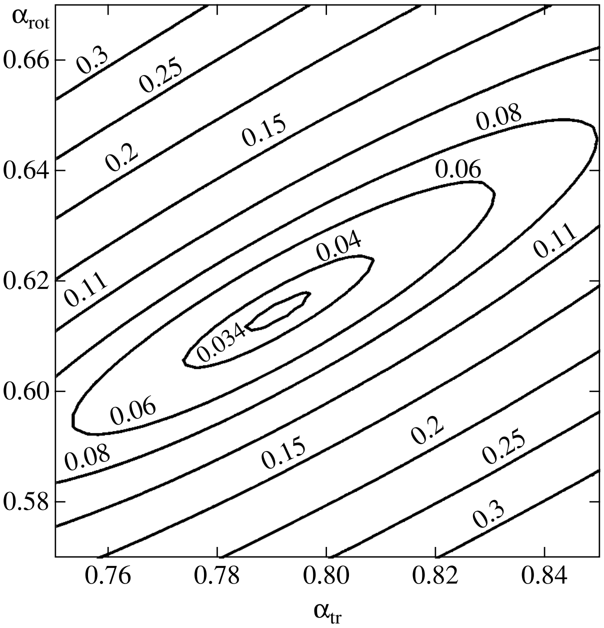

Fig. 2. Differences Δ between the experimental and calculated electric resistances of the wire probe in a flow of CH_4 obtained in determining accommodation coefficients ($P_0 \approx$ 2100 Pa, L = 9.4 mm, S = 4.16, α_{tr} = 0.79, and α_{rot} = 0.61).

Table 1. Measured accommodation coefficient values

Gas	α_{tr}	α_{rot}
N_2	0.67 ± 0.04	0.56 ± 0.08
CH_4	0.78 ± 0.10	0.73 ± 0.16
CO_2	0.90 ± 0.23	0.79 ± 0.15
H_2	0.44 ± 0.04	-0.15 ± 0.05

and 2 shows that our data qualitatively agree with those reported earlier.

The experimental data on N_2, CO_2, and CH_4 presented above lead us to conclude that, for molecules with low energies (corresponding to room braking temperature), the translational energy accommodation coefficient is higher than the internal energy accommodation coefficient. This substantiates the results of theoretical [18] and experimental [11, 13–16, 20, 22] works (some of them are given in Table 2).

The determination of internal energy accommodation coefficients by the suggested method is of interest. This method gives integral information about effective rotational energy conversion in collisions with surfaces when the translational and internal energies of a gas flow are substantially different. The available data on such accommodation coefficients are exceedingly scarce. Experimental units in which molecular beams are used are known to be the main tool for experimentally studying gas–surface interactions at the level of information about translational and internal energy distribution functions. In numerous experimental works performed with the use of molecular beams, accommodation coefficients are usually not given, and, often, the reported data are insufficient for calculating them. In addition, the most interesting molecular-beam experimental data were obtained for thoroughly cleansed, oriented crystalline surfaces specially prepared in conformity to the requirements of surface physics. Analysis of the works containing sufficient data for calculating accommodation coefficients shows that these coefficients can fall outside the range of values usually believed to be characteristic of them. For instance, the rotational energy accommodation coefficient becomes higher than one at a very high translational energy of incident molecules [30, 31]. This effect can correspond to energy transfer between translational and rotational degrees of freedom caused by inelastic interactions of molecules with surfaces.

Our experiments with hydrogen also challenge the traditional views on internal energy accommodation coefficients. The special feature of the adiabatic expansion of a hydrogen jet under our experimental conditions ($d_* = 6$ mm, $P_0 = 1300$–1400 Pa, $T_0 = 290$–300 K, and $D = 4d_*$) is quenching of rotational energy close to the nozzle exit. According to experimental and theoretical studies [32], the rotational temperature after quenching can be estimated at $T_{rot} \approx 220$ K. Our numerical processing of the experimental data gave a negative rotational energy accommodation coefficient (see Table 1), which was reproduced in repeat experiments. A simple analysis of (1) allows us to explain this result. At a wire temperature of $T_w > 300$ K and a rotational temperature of $T_{rot} \approx 220$ K of the incident gas, we obtain $E_{i,\,rot} < E_{w,\,rot}$. It then follows from the negative α_{rot} accommodation coefficient that the rotational energy of reflected molecules $E_{r,\,rot}$ should be smaller than the rotational energy of incident molecules $E_{i,\,rot}$ (although the wire is hot with respect to the gas and should transfer energy to gas molecules). Following the generally accepted approach [9, 10], let us write the rotational energy of reflected molecules in the form

$$E_{r,\,rot} = E_{i,\,rot} + \Delta E_s - \Delta E_{r \to t} + \Delta E_{t \to r},$$

where ΔE_s is the energy transferred from the surface, $\Delta E_{r \to t}$ is the energy of rotational degrees of freedom converted to translational energy, and $\Delta E_{t \to r}$ is the energy of translational degrees of freedom converted to

Table 2. Literature data on accommodation coefficients

Gas	Surface	T_w, K	α_{tr}	α_{rot}	Refs.
N_2	Au covered by the gas	293	0.89	0.78	[16]
N_2	Au	350	0.93	0.79	[22]
N_2	Au annealed in a vacuum	400	0.18	0.12	[13]
N_2	Pt	297	0.49	0.05	[11]
CH_4	Pt	373	0.81*	–	[3]
CO_2	W covered by CO_2	330	0.99*	–	[3]
NO	Pt cleansed by an ion beam	600	1	0.62	[14]
Air	W	350	0.97	0.84	[20]
H_2	Pt	338	0.337**	0.034**	[18]

Note: * Determined for the total energy.
** Calculated quantum-mechanically.

rotational energy. It follows from the condition $E_{r,\,\mathrm{rot}} < E_{i,\,\mathrm{rot}}$ that $\Delta E_{r\to t} > \Delta E_s + \Delta E_{t\to r}$. The negative rotational energy accommodation coefficient therefore arises because of the conversion of the rotational energy quenched at low energy levels into the translational energy of reflected molecules in collisions of molecules with the surface. The energy flux caused by this conversion substantially exceeds the energy flux from the surface to the gas. This explanation is indirectly substantiated by the data obtained in [33–35]. The authors of [33] studied the interaction of molecular hydrogen with the (001) MgO surface. They observed scattering that corresponded to rotational energy deactivation at certain low levels when the gas inelastically collided with the surface. According to [34], partial rotational energy conversion into the kinetic energy of scattered molecules was observed in the interaction of NO molecules with the surface of graphite. A microscopically grounded model of the interaction of a polyatomic molecule with the surface that takes into account energy exchange between translational and rotational degrees of freedom is described in [35].

5. CONCLUSIONS

The experimental data on accommodation coefficients and interactions of separate molecules with surfaces obtained in molecular-beam experiments do not meet the practical demands for describing processes in which technical surfaces participate. It is precisely for this reason that the development of the suggested technique for studying translational and rotational energy conversion in gas interactions with surfaces under the conditions that correspond to real technologies is of fundamental interest. The suggested method is a fairly simple and effective tool for obtaining data useful both for understanding energy exchange between a gas and a surface and for modeling gas–surface interactions in real technological applications.

Analysis of the literature and our experimental data performed above gives a good reason for revising views on accommodation coefficients taking into account internal energy. Changes in the translational and internal energy of molecules can be estimated, for instance, in terms of “conversion coefficients,” which should not necessarily lie in the interval 0–1.

ACKNOWLEDGMENTS

This work was financially supported by INTAS (grant no. 99-0749) and grant no. 910.2003.1 from the President of Russian Federation in support of leading scientific schools.

APPENDIX

The final form of Eqs. (1) and (2) can only be found in [21]. We therefore think it expedient to give their derivation. For Knudsen flow past a convex surface, the heat flow per unit surface is written as [9]

$$q = \frac{n(kT_{\mathrm{tr}})^{3/2}}{\sqrt{2\pi m}}\left\{\left[\alpha_{\mathrm{tr}}\left(S^2 + \frac{5}{2} - 2\frac{T_w}{T_{\mathrm{tr}}}\right) + \frac{1}{2}\alpha_{\mathrm{rot}} j\left(\frac{T_{\mathrm{rot}}}{T_{\mathrm{tr}}} - \frac{T_w}{T_{\mathrm{tr}}}\right)\right]\Psi(S_n) - \frac{1}{2}\alpha_{\mathrm{tr}}\exp(-S_n^2)\right\}, \tag{5}$$

$$\Psi(x) = \exp(-x^2) + x\sqrt{\pi}(1 + \mathrm{erf}(x)),$$

where n is the number density of particles; T_{tr} and T_{rot} are the translational and rotational temperatures of the gas flow, respectively; T_w is the wire temperature; $S = U/(2kT_{\mathrm{tr}}/m)^{1/2}$ is the velocity ratio; U is the velocity; m is the molecular weight; $S_n = S\sin\beta$; and β is the angle between the flow direction and the surface. The equation for the heat flow to a wire of length L and radius r oriented normally to the flow direction is obtained by integrating (5) over the surface of the cylinder,

$$Q = rL\int_{-\pi}^{\pi} q\,d\beta = 2rL\frac{n(kT_{\mathrm{tr}})^{3/2}}{\sqrt{2\pi m}}\left(AJ_1 - \frac{1}{2}\alpha_{\mathrm{tr}}J_2\right),$$

where

$$A = \alpha_{\mathrm{tr}}\left(S^2 + \frac{5}{2} - 2\frac{T_w}{T_{\mathrm{tr}}}\right) + \frac{1}{2}\alpha_{\mathrm{rot}} j\left(\frac{T_{\mathrm{rot}}}{T_{\mathrm{tr}}} - \frac{T_w}{T_{\mathrm{tr}}}\right),$$

$$J_1 = \int_{-\pi/2}^{\pi/2}\Psi(S_n)\,d\beta, \quad J_2 = \int_{-\pi/2}^{\pi/2}\exp(-S_n^2)\,d\beta.$$

Using the representation of the modified Bessel functions in the form [36]

$$I_n(x) = \frac{1}{\pi}\int_0^{\pi}\exp(x\cos\theta)\cos(n\theta)\,d\theta,$$

we can write the J_1 and J_2 integrals as

$$J_1 = \pi\exp\left(-\frac{S^2}{2}\right)\{I_0 + S^2[I_0 + I_1]\}$$

$$J_2 = \pi\exp\left(-\frac{S^2}{2}\right)I_0$$

(here, I_0 and I_1 are the modified Bessel functions zeroth- and first-order in the $S^2/2$ argument). The equation for Q then becomes

$$Q = 2\pi L\frac{n(kT_{\mathrm{tr}})}{\sqrt{2\pi m}}\exp\left(-\frac{S^2}{2}\right)$$

$$\times\left(A[I_0 + S^2(I_0 + I_1)] - \frac{1}{2}\alpha_{\mathrm{tr}}I_0\right).$$

Transforming the term in parentheses gives the eventual equation for the energy flux from a gas to a wire, namely,

$$Q = 2\pi L h(T_r - T_w),$$

$$T_r = \frac{T_{\mathrm{tr}}}{4\alpha_{\mathrm{tr}} + j\alpha_{\mathrm{rot}}}$$

$$\times\left\{\alpha_{\mathrm{tr}}\left(2S^2 + 5 - \frac{1}{1 + S^2(1 + I_1/I_0)}\right) + \alpha_{\mathrm{rot}} j\frac{T_{\mathrm{rot}}}{T_{\mathrm{tr}}}\right\},$$

$$h = \frac{1}{4\sqrt{\pi}} nk(4\alpha_{\mathrm{tr}} + j\alpha_{\mathrm{rot}})$$

$$\times\exp\left(-\frac{S^2}{2}\right)\frac{U}{S}[I_0 + S^2(I_0 + I_1)].$$

REFERENCES

1. M. Knudsen, Ann. Phys. (Leipzig) **34**, 593 (1911).
2. F. O. Goodman and H. Y. Wachman, *Dynamics of Gas–Surface Scattering* (Academic, New York, 1976; Mir, Moscow, 1980).
3. O. A. Kolenchits, *Thermal Accommodation of Gas–Solid Systems* (Nauka i Tekhnika, Minsk, 1977).
4. R. G. Barantsev, *Interaction of Rarefied Gases with Streamline Surfaces* (Nauka, Moscow, 1975).
5. S. F. Borisov, N. F. Balakhonov, and V. A. Gubanov, *Interaction of Gases with Solid-State Surfaces* (Nauka, Moscow, 1988).
6. C. Hurlbut, in *Rarefied Gas Dynamics: Proceedings of the 20th International Symposium*, Ed. by C. Shen (Peking Univ. Press, Beijing, China, 1997), p. 355.
7. A. Yu. Pankov, V. D. Borman, and S. Yu. Krylov, Zh. Éksp. Teor. Fiz. **113**, 1350 (1998) [JETP **86**, 737 (1998)].
8. G. N. Patterson, *Molecular Stream of Gases* (Fizmatgiz, Moscow, 1960), p. 187.
9. M. N. Kogan, *Rarefied Gas Dynamics* (Nauka, Moscow, 1967), pp. 88 and 351.
10. I. N. Larina and V. A. Rykov, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gaza, No. 5, 141 (1986).
11. G. M. Rosenblatt, R. S. Lemons, and C. W. Draper, J. Chem. Phys. **67**, 1099 (1977).
12. D. J. Marsden, in *Rarefied Gas Dynamics: Proceedings of the 4th International Symposium*, Ed. by J. H. de Leeuw (Academic, New York, 1966), Vol. 2, p. 566.
13. V. Ramesh and D. J. Marsden, Vacuum **24**, 291 (1974).
14. M. Asscher, W. L. Guthrie, T.-H. Lin, *et al.*, J. Chem. Phys. **78**, 6992 (1983).
15. H. Legge, in *Rarefied Gas Dynamics: Proceedings of the 14th International Symposium*, Ed. by H. Oguchi (Univ. of Tokyo Press, Tokyo, 1984), Vol. 1, p. 271.
16. K. Schafer and K. H. Riggert, Z. Elektrochem. **57**, 751 (1953).
17. B. Petek and I. Kuscer, in *Rarefied Gas Dynamics: Proceedings of the 11th International Symposium*, Ed. by R. Campargue (Commissiariat à l'Energie Atomique, Paris, 1979), Vol. 2, p. 1399.
18. P. Feuer and C. Osburn, in *Rarefied Gas Dynamics: Proceedings of the 6th International Symposium*, Ed. by L. Trilling and H. Y. Wachman (Academic, New York, 1969), Vol. 2, p. 1095.
19. V. M. Sankovich and N. A. Samsonov, Tr. Tsentr. Aérolog. Observ., No. 139, 76 (1979).
20. V. Sankovich, in *Rarefied Gas Dynamics: Proceedings of the 15th International Symposium*, Ed. by V. Boffi and C. Cercignani (Teubner, Stuttgart, 1986), Vol. 1, p. 627.
21. A. V. Bulgakov, in *Molecular Physics of Nonequilibrium Systems*, Ed. by E. V. Baklanov (ITF, Novosibirsk, 1984), p. 20.
22. A. V. Bulgakov and V. G. Prikhod'ko, Prikl. Mekh. Tekh. Fiz., No. 5, 126 (1986).
23. A. V. Bulgakov, V. G. Prikhodko, A. K. Rebrov, *et al.*, in *Rarefied Gas Dynamics: Physical Phenomena, Proceedings of the 16th International Symposium on Progress in Astronautics and Aeronautics*, Ed. by E. P. Muntz, D. Weaver, and D. Campbell (AIAA, Washington, DC, 1989), Vol. 117, p. 92.
24. H.-S. Tsien, J. Aeronaut. Sci. **15**, 573 (1948).
25. L. Gottesdiner, J. Phys. E: Sci. Instrum. **13**, 908 (1980).
26. R. G. Lord, J. Phys. E: Sci. Instrum. **7**, 56 (1974).
27. A. K. Rebrov, A. A. Morozov, M. Yu. Plotnikov, *et al.*, Rev. Sci. Instrum. **74**, 1103 (2003).
28. A. K. Rebrov, J. Vac. Sci. Technol. A **19**, 1679 (2001).
29. V. A. Zhokhov and A. A. Khomutskiĭ, *Atlas of Supersound Flows of Free-Expanding Ideal Gas Flowing out an Axis-Symmetric Jet* (Tsentr. Aerogidrodin. Inst., Moscow, 1970), Tr. TsAGI, No. 1224.
30. R. J. Lahaye, W. E. Stolte, S. Holloway, *et al.*, J. Chem. Phys. **104**, 8301 (1996).
31. S. Novopashin, C. Dankert, K. Lehmkoster, *et al.*, in *Rarefied Gas Dynamics: Proceedings of the 20th International Symposium*, Ed. by C. Shen (Peking Univ. Press, Beijing, China, 1997), p. 428.
32. H. P. Godfried, I. F. Silvera, and J. Van Straaten, in *Rarefied Gas Dynamics: Proceedings of the 12th International Symposium*, Ed. by S. S. Fisher (AIAA, New York, 1981), Vol. 74, Part 2, p. 772.
33. L. Rathbun and G. Ehrlich, in *Rarefied Gas Dynamics: Proceedings of the 10th International Symposium*, Ed. by J. L. Potter (AIAA, New York, 1977), Vol. 51, Part 1, p. 555.
34. J. Hager, D. Glatzer, H. Kuze, *et al.*, Surf. Sci. **374**, 181 (1997).
35. A. Yu. Pankov, S. Yu. Krylov, E. J. van Duijn, *et al.*, J. Chem. Phys. **112**, 8680 (2000).
36. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 745–753.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 829–839.

PLASMA, GASES

Current-Induced Implosion of a Multiwire Array as a Radial Plasma Rainstorm

V. V. Alexandrov[a], E. V. Grabovsky[a], G. G. Zukakishvili[a], M. V. Zurin[a], N. N. Komarov[b], I. V. Krasovsky[a], K. N. Mitrofanov[a], S. L. Nedoseev[a],*, G. M. Oleinik[a], I. Yu. Porofeev[a], A. A. Samokhin[a], P. V. Sasorov[c], V. P. Smirnov[a], M. V. Fedulov[a], I. N. Frolov[a], and A. A. Chernov[a]

[a] *Troitsk Institute of Innovation and Thermonuclear Investigations, Troitsk, Moscow oblast, 142092 Russia*
[b] *All-Russia Research Institute of Electromechanics, Moscow, Russia*
[c] *State Scientific Center, Institute of Theoretical and Experimental Physics, Moscow, 117218 Russia*
**e-mail: nedos@triniti.ru*

Received April 4, 2003

Abstract—We present our experimental results of the X-ray radiography of fast radiating *Z*-pinches based on cylindrical multiwire tungsten arrays. The experiments were carried out at the Angara-5-1 facility at an electrical power of up to 4 TW with a discharge current of up to 4 MA rising at a rate on the order of 5×10^{13} A s^{-1}. The linear mass of single and composite arrays reached 500 μg cm^{-1}, the initial radius was 4–10 mm, and the wire diameter was 5–8 μm. We have experimentally shown that for the current-induced implosion of multiwire tungsten arrays, significant azimuthal and axial plasma inhomogeneities result from discharge cold start and prolonged plasma production, which determine the subsequent course of the implosion. The *Z*-pinch structure also remains spatially inhomogeneous at the time of intense X-ray radiation. The generated inhomogeneous plasma collapses toward the array axis in the form of numerous radially elongated plasmoids with relatively small diameters. The stream of plasmoids is called a radial plasma rainstorm. As the plasmoids contract toward the array axis, they decrease in radial size and merge into isolated plasma current filaments, which are elongated mainly along the discharge axis. We critically discuss the models of a radiating *Z*-pinch in plasma composed of matter with a large atomic number that disregard the cold-start and prolonged plasma-production effects. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The Angara-5-1 facility is used to study multiterawatt-power *Z*-pinches in various plasma-producing media as well as to carry out experiments with multiwire arrays. In these experiments, we discovered and investigated the phenomenon of prolonged plasma production [1–3]. It manifests itself in fast *Z*-pinches with cold start at a high rate of rise of the discharge current ($dI/dt \sim 0.5–1.0 \times 10^4$ A s^{-1}). The following features are characteristic of fast *Z*-pinches with prolonged plasma production.

(i) At cold start, i.e., when a voltage pulse from the main electric power generator breaks down an initially unionized plasma-producing material, the plasma production is spatially inhomogeneous. This is characteristic of both homogeneous (gas) and heterogeneous (solid-state low-density foam, multiwire arrays) plasma-producing media, although the breakdown and plasma-production mechanisms differ for these media. In a gaseous plasma-producing medium, the spatial inhomogeneity manifests itself in the form of azimuthal current filamentation followed by axial stratification of the plasma formed [1]. For multiwire arrays, the azimuthal inhomogeneity is specified by the array pattern itself: the separation between the wires (200–1000 μm) is always much larger than the wire diameter (5–10 μm).

(ii) As the discharge current rises, the magnetic plasma acceleration in the plasma-production zone begins before the complete transformation of the entire plasma-producing material into accelerated plasma. The created rare plasma saturated with current is carried away with it from the plasma-production zone. In the specific case of a discharge through a multiwire array, the prolonged plasma production can be explained by the insufficient flux of heat and radiation from the rare plasma through which the discharge current flows on the dense thin plasma-producing filaments or microscopic drops of material of the wires that exploded with virtually no current. As a result, the plasma production is prolonged in time. So, the ionization of the plasma-producing material, the plasma transport from the pinch periphery to the axis, and the imploded radiating state of the pinch take place in mutually overlapping time intervals.

These factors significantly affect not only the initial discharge stage but also the subsequent stages—the implosion and the imploded state of the radiating

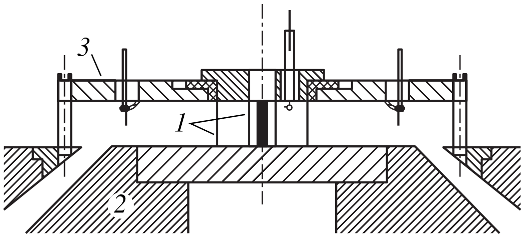

Fig. 1. Arrangement of multiwire arrays between the current-conducting electrodes with magnetic self-insulation: (*1*) multiwire array; (*2*) cathode; (*3*) anode.

Z-pinch. In our experiments [3], we found that the tungsten plasma from the wires in the plasma-production zone never merges into a single azimuthally inhomogeneous shell. In this paper, we present our experimental results showing that the azimuthal and axial plasma density inhomogeneities that arise at the stage of prolonged plasma production in a multiwire array give rise to a radial stream of numerous plasmoids. When the *Z*-pinch implodes, these plasmoids converge to the axis of symmetry of the array and create an inhomogeneous three-dimensional structure of the radiating *Z*-pinch.

This paper has the following structure. In Section 2, we briefly describe the experimental parameters, which were identical for [1, 3] and for this paper. In Section 3, we briefly present the characteristics of the implosion dynamics of *Z*-pinches, including multiwire arrays, which we studied previously. In Section 4, we present new experimental results of our study of the implosion of a multiwire array. In Section 5, we suggest a model for the current-induced implosion of a multiwire array based on experimental data.

2. BASIC PARAMETERS OF THE EXPERIMENTS WITH MULTIWIRE ARRAYS

Our experiments [1, 3] with multiwire cylindrical arrays at the Angara-5-1 facility have the following basic parameters: a discharge current of 2.5–4 MA, a current rise time of 100 ns, a discharge power of up to 4 TW, a multiwire array diameter of 6–20 mm, a height of 10–15 mm, a tungsten wire diameter of 6–10 μm, and a linear mass of single wire array from 120 to 500 μg cm^{-1}. We investigated both single arrays and two- and three-cascade coaxial arrays. Solid cylinders of low-density agar–agar foam saturated with tungsten micropowder were used as the third, inner cascade. The total linear mass of the three-cascade compositions reached 1 mg cm^{-1}. The arrangement of the arrays between current-conducting electrodes with magnetic self-insulation is schematically shown in Fig. 1.

For diagnostics, we used the following techniques in our experiments: measurements of the derivative of the total discharge current, magnetic probe measurements, laser shadow probing and interferometry, measurements of the total intensity of soft X-ray radiation, optical and X-ray image intensifiers in frame-by-frame imaging mode and an optical image intensifier in slit scan mode, X-ray probing of the dense component of the plasma-producing medium using an auxiliary *x*-pinch [3], and a time-integrated X-ray pinhole camera of high spatial resolution.

3. MAIN EXPERIMENTAL RESULTS FROM [1–3]

Without repeating the numerous experimental data from [1–3], we present their main results referring to the techniques used to obtain them.

3.1. The Plasma-Production Stage

During the plasma production, azimuthally separated plasma streams are formed in the *rz* plane from each wire (laser probing, an X-ray image intensifier, slit scans). They are carried off by the Ampere force in the radial direction without merging together, so a plasma configuration in the form of a star with the number of rays equal to the number of wires is formed in projection orthogonal to the *z* axis. The plasma density in the streams (~10^{18} cm^{-3}) is low compared to the density of the plasma-producing material of the exploded wires. An axial inhomogeneity of the electron density of the generated plasma with a characteristic scale of approximately 250 μm manifests itself near the wires (laser probing, an X-ray image intensifier). The axial inhomogeneity of the dense component of the plasma-producing array wires has scale smaller by an order of magnitude (probing by *x*-pinch radiation). The outer boundary of the rare plasma near the wires is not displaced from the starting radius of the multiwire array until almost complete evaporation of the plasma-producing medium (slit scans, laser probing, an X-ray image intensifier).

3.2. The Implosion Stage

The plasma production for different wires terminates at different times with a spread of 10–20 ns. In this case, the termination of the plasma production along one wire is axially nonsimultaneous, and the termination of the plasma production for the array as a whole is azimuthally nonsimultaneous (X-ray probing). In zones where the plasma production terminated, the outer plasma boundaries begin to implode (slit scans, laser shadow probing). The nonsimultaneous termination of the plasma production along the wires is responsible for the initial axial nonuniformity in the motion of the outer plasma sheet boundaries (laser shadow probing and interferometry, an X-ray image intensifier).

4. A RADIOGRAPHIC STUDY OF THE IMPLOSION STRUCTURE OF THE MULTIWIRE ARRAY

In this section, we present the results of our radiographic study of the implosion of multiwire arrays that were mainly obtained by using time-integrated pinhole cameras with a high spatial resolution. Holes with diameters from 10 to 60 μm were used in the pinhole camera; the hole profile was in the shape of a crater with a vertex angle of 30°–50°. The holes were punched by a laser pulse in 40-μm-thick tantalum foil. Since the pinhole camera was fairly close to the object under study and was exposed to intense X-ray and microparticle fluxes, the foil with a hole was changed after each experiment. Aluminum filters that cut off the soft radiation component were placed behind the 30- and 60-μm holes. The distance from the pinch axis to the pinhole was 10 cm, and the distance from the recorder to the pinhole was 5.3 cm. The X-ray image was recorded on photographic plates.

Let us first consider the X-ray photograph for the *Z*-pinch of the multiwire array in soft X-ray radiation (photons with energies > 150 eV). It is shown in Fig. 2. Since the image of the paraxial zone is overexposed, no structural features are seen in this region. The peripheral zone has a normal exposure, and structural features of the tungsten plasma radiation are clearly seen. The two-dimensional image shown in Fig. 2 is a time-integrated image of the evolution of the three-dimensional distribution of X-ray sources when they are generated and move toward the array axis during the current-induced implosion of the tungsten plasma and its radiation in the imploded state. First, we can say that the three-dimensional time-integrated image of the imploding plasma in soft X-ray radiation is in the form of a cylindrical brush composed of separate needles oriented mostly radially. To correctly formulate the experimental results that follow from the X-ray photograph in Fig. 2, we will take into account the following considerations. Various factors can be responsible for the soft X-ray radiation of the peripheral plasma in our experiments.

First, these may include the intrinsic radiation of the current-carrying plasma that emerges during its motion into the paraxial zone—the region of maximum radiation of the pinch as a whole. These may also include the intrinsic radiation of the plasma that has already filled the space inside the array and that arises from the subsequent MHD turbulent breakthrough of the current from the periphery into the paraxial zone.

Second, these factors may include the flash of radiation from the peripheral plasma that emerges immediately during maximum plasma implosion and radiation in the paraxial zone. It can result from the influence of the multiterawatt photon flux from the plasma that imploded at the pinch axis on the peripheral plasma. The photoionization of the peripheral plasma by this flux can cause an increase in its radiation characteristics. The temperature rise in the paraxial zone at this time may also cause an increase in the temperature and radiation intensity of the peripheral plasma due to thermal conductivity. Finally, the flash of radiation from the peripheral plasma at maximum *Z*-pinch implosion may result from a jump in voltage between the electrodes, which arises from a sharp increase in the impedance of the current-carrying plasma imploding at a high rate.

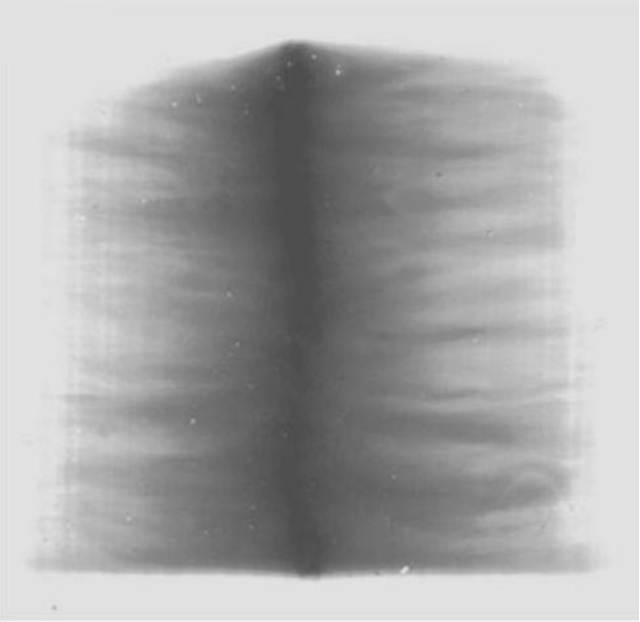

Fig. 2. Time-integrated image for the pinch of the multiwire array at energies above 150 eV. The initial array diameter and height are 12 and 10 mm, respectively.

If the first group of factors is correct, then the image in Fig. 2 can be left both by short, but long-radiating plasmoids moving toward the axis and by plasma jets with a length on the order of the initial array radius radiating in a relatively short time interval. When carefully examining the negatives, we can see that the initial points of the plasma jet trajectories are tied to the wires and do not go beyond the outer array radius. Consequently, these are the streams of imploding plasma rather than the expanding plasma after the *Z*-pinch implosion. The imploding plasma of the multiwire array largely consists of plasmoids with axial and azimuthal sizes of less than 1 mm, i.e., small compared to the array radius and length. The radial length of these plasmoids cannot be determined from our time-integrated X-ray photograph. However, it follows from the X-ray photographs taken with an exposure of about 3 ns that immediately before the main radiation pulse, the length of some plasmoids is comparable to the initial array radius (Fig. 3).

If the second group of factors is correct, then we have an X-ray photograph of the pinch taken during its maximum implosion and radiation (6–15 ns). We see from this photograph that by no means was all of the plasma collected in the paraxial zone by this time.

At present, we cannot choose the only factor from those listed above, although the second group of factors seems most plausible. However, in any case, the main conclusion that follows from the X-ray photograph

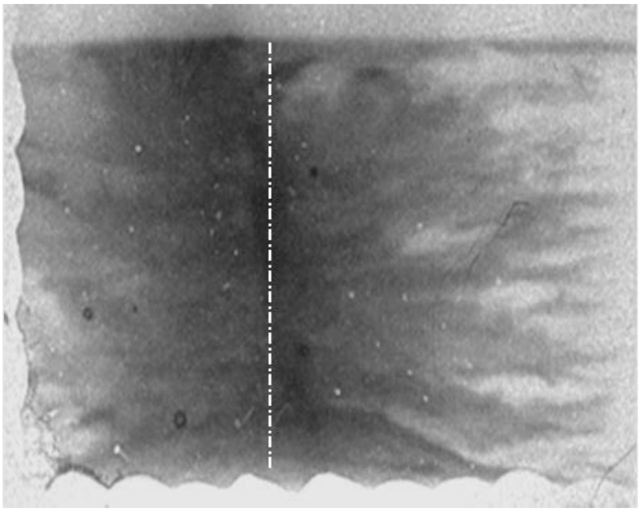

Fig. 3. Image of part of the imploding multiwire array at photon energies above 150 eV 138 ns after the onset of current; the exposure is 3 ns. The linear array mass is 285 μg cm^{-1}, the array consists of 30 tungsten wires 8 μm in diameter; the initial array radius and height are 10 and 10 mm, respectively. The dark regions indicate the plasma, and the white line indicates the location of the axis of symmetry of the array.

shown in Fig. 2 is beyond question: the spatial structure of the imploding plasma is axially and azimuthally so inhomogeneous that any one-dimensional or even two-dimensional imploding plasma shell is out of the question. The initial spatial inhomogeneity of the accelerated plasma and the nonsimultaneity of its generation at the starting radius are so large that it would be improper to use the formalism of Rayleigh–Taylor instability to describe the implosion dynamics of the pinch plasma.

Let us now turn to the images of the plasma that arises from the current-induced implosion of multiwire arrays at photon energies above 1 keV. First, note that when using a composition of two coaxial multiwire arrays 12 and 6 mm in diameter and 15 mm in length, we can obtain its imploded radiating state of very high geometrical quality. The X-ray photographs for this case are shown in Fig. 4. The diameter of the pinch image at energies above 1.5 keV does not exceed 400 μm, which corresponds to a degree of implosion of 30 when measured from the initial diameter of the outer array. The total discharge current in this experiment was approximately 3 MA, the duration of the X-ray pulse at half maximum was approximately 6 ns, and the active power in the discharge circuit was about 4 TW. The corresponding time dependences of the discharge current and the X-ray intensity are shown in Fig. 5.

In our experiments, the high-quality implosion of the composite pinch shown in Figs. 4a–4c and 5 was achieved by carefully choosing the initial parameters of the multiwire arrays. The implosion quality of single arrays and composite systems for other parameters was not so high. The time-integrated X-ray photographs of

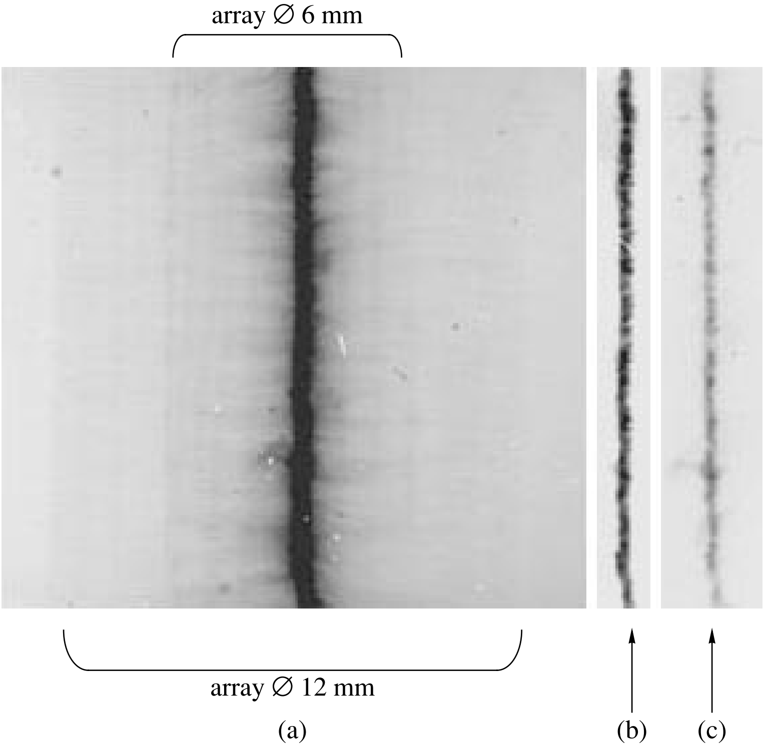

Fig. 4. Time-integrated pinhole image of a 15-mm-long composite pinch. The photon energy is above (a) 150 eV, (b) 1 keV, and (c) 1.5 keV; the pinch diameter in Fig. 4c is approximately equal to 400 μm.

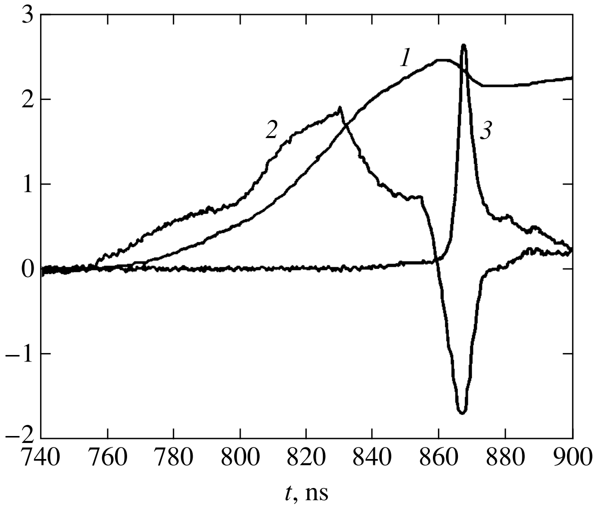

Fig. 5. Time dependences of (*1*) the discharge current *I*, MA, (*2*) the derivative of the current *dI*/*dt*, arb. units, and (*3*) the intensity of soft X-ray radiation, arb. units, for the X-ray photograph in Fig. 4.

the implosion of single arrays show that azimuthally and axially inhomogeneous structures of radiating plasma with hot spots and a number of radial outgrowths are formed in the paraxial zone of the multiwire array. These structures are clearly seen in the time-integrated X-ray photographs (see Figs. 6, 8, and 9) taken in the photon energy range 150–2000 eV. For example, the intensely radiating outgrowths in Fig. 6 are not connected with the main body of the pinch at all. In general, the pinch structure is most inhomogeneous in the near-cathode region. In cases of highly developed inhomogeneities, the X-ray pulse is broadened and contains additional maxima (see Fig. 7). We may assume that the structural elements of the pinch seen in the pinhole images radiate at different times, increasing the total duration of the X-ray pulse. The presence of radially oriented radiating pinch elements can also be explained by the fact that the discharge current flows through channels some of which have an appreciable radial component not only near the electrodes (cathode) but also in the bulk of the imploding plasma.

The high spatial resolution of our pinhole cameras allows us to examine the fine structure of the imploded radiating pinch plasma even without time resolution. Figures 8 and 9 show the fine structure of the radiating pinch regions with a resolution higher than 50 μm over the observed object. Axial inhomogeneity of the radiating plasma is more characteristic of the pinch shown in Fig. 8, while azimuthal inhomogeneity is clearly seen in Fig. 9. The set of our experimental results reveals not a tubular structure of the pinch but two separate current radiating channels interacting through their own magnetic fields.

5. DISCUSSION OF EXPERIMENTAL RESULTS

5.1. The Z-Pinch of a Multiwire Array as a Radial Plasma Rainstorm

Let us consider the schemes shown in Figs. 10 and 11 as a working model for the fast implosion of a multiwire array based on our experimental results and the results from [1–3]. During plasma production, a plane plasma stream propagates from each wire, as from the initial outer boundary in the (*rz*) plane, toward the axis of the multiwire array. The number of streams formed is equal to the number of wires (Fig. 10a). Let us con-

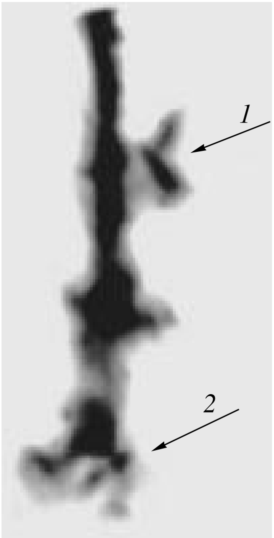

Fig. 6. The time-integrated X-ray photograph of part of a pinch about 5 mm in length with separated plasmoids for photons with *h*ν > 1 keV: (*1*) plasma outgrowths; (*2*) near-cathode pinch region.

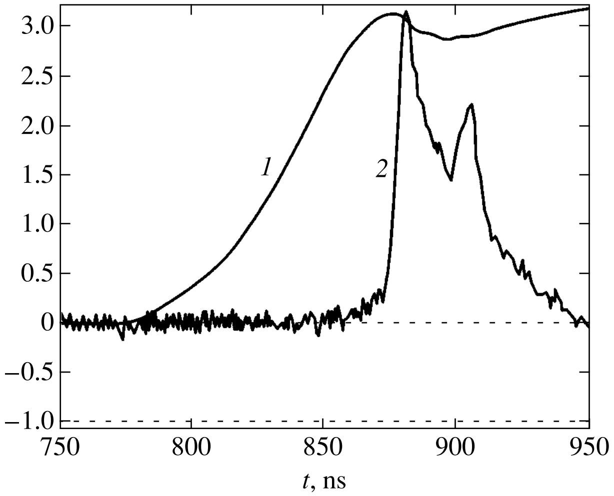

Fig. 7. The current (1, MA) and the intensity of soft X-ray radiation (2, arb, units) (*h*ν ~ 200 eV) corresponding to the pinhole image in Fig. 6.

Fig. 8. Azimuthally axial inhomogeneities of the radiation of a pinch region; the pinch diameter is approximately 0.8 mm.

sider the behavior of an arbitrarily chosen stream (Figs. 10b–10d). Throughout the plasma production, the outer boundary of the stream is rigidly tied to the plasma-generating zone—the exploded wire that shifts only slightly. Having reached the axis, the inner boundaries of the plasma streams from the wires produce a precursor pinch. The plasma production along an individual wire terminates nonsimultaneously. In the places where it terminated, the outer boundary of the plane plasma stream begins to radially shift toward the axis. Since the termination of the plasma production along the wire is not simultaneous, the outer boundary of the plasma stream in the (*rz*) plane is spatially inhomogeneous. The $[\mathbf{j} \times \mathbf{H}]$ force increases this inhomogeneity as the plane plasma stream is radially displaced and collapses.

The axially inhomogeneous radial collapse of plasma streams from the wires (or even their breakage) gives rise to isolated pinches with outgrowths directed radially, from the axis outward, toward the plasma-producing wires in the paraxial zone. The combination of neighboring plasma streams into isolated pinches and the radiation from them can take place earlier than the arrival of the pinches at the array axis. In addition, some of the isolated pinches can also merge with neighboring pinches before they reach the array axis. During the interaction of separate streams, an azimuthal velocity component can locally arise in some of the plasmoids. Plasmoids with a nonzero azimuthal velocity will not reach the axis of symmetry of the array. As a result, a radiating pinch structure arises (its model is shown in Fig. 11).

The recorded azimuthal and axial inhomogeneities of the imploding plasma of a multiwire array result from the break of the azimuthal magnetic flux through

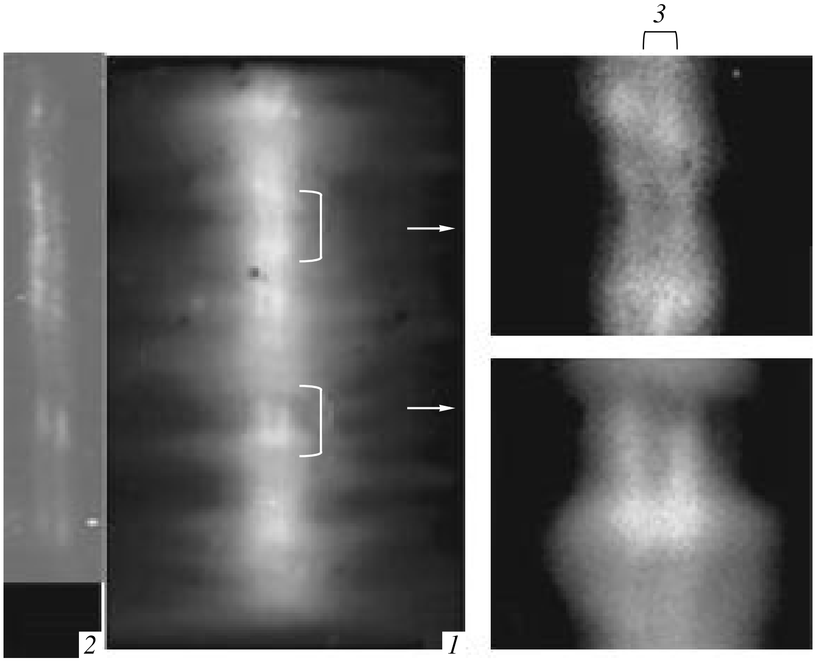

Fig. 9. The fine radially azimuthal pinch structure: (*1*) for photons with $h\nu > 150$ eV; (*2*) for photons with $h\nu > 1$ keV; (*3*) the diameter of the radiating filament, 200 μm.

the multiwire array in its parts where the plasma production terminated because of the evaporation of the wire material. For the plasma stream from each wire, the process of plasma production and its termination is random. Thus, the inhomogeneity is predetermined, on the one hand, by the well-defined initial structure of the array and, on the other hand, by the statistical nature of the plasma production.

The structure of the radially elongated plasmoids moving into the paraxial zone of the array clearly shows up against the background of a more homogeneous plasma with a lower density [4]. Below, we call this stream a radial plasma rainstorm.

5.2. Existing Models of a Z-Pinch

Giving this name to our phenomenological model for the implosion of a multiwire array, we seek to emphasize the distinctive features of its implosion dynamics. These features clearly show up in an actual experiment but are disregarded in the standard models of a *Z*-pinch. The "snowplow" and plasma-liner models of pinches as well as their modifications for numerical simulations [5] belong to these models. These models assume the presence of a more or less homogeneous cylindrical plasma shell before the onset of pinch implosion followed by the growth of two-dimensional MHD instabilities during the implosion. A "heuristic model" for the *Z*-pinch of a multiwire array and its modifications for numerical simulations, which define the initial phase and dynamics of the current-induced implosion of a multiwire array with an allowance made for the azimuthal and axial inhomogeneities of the initial plasma, have appeared in recent years [6–8]. A pinch model based on toroidal buoyant toroidal magnetic flux tubes forming on the periphery and converging toward the pinch axis was suggested to describe the energetics of the imploded state of a fast *Z*-pinch [9, 10]. This one-dimensional model assumes MHD turbulent mixing of the magnetic flux that penetrates into an imploded pinch with a velocity close to the Alfven velocity. The power of the magnetic energy transfer in this flux per unit pitch length is estimated in [9, 10] to be

$$W = \chi \frac{H_\varphi^2}{8\pi} 2\pi r V_A, \qquad (1)$$

where H_φ [Oe] is the azimuthal magnetic field strength of the current I [A] flowing through a pinch of radius r [cm], and V_A [cm s^{-1}] is the Alfven velocity. Assuming that

$$H_\varphi = 0.2\frac{I}{r}, \quad V_A \sim \frac{I}{10M^{1/2}},$$

where M [g cm^{-1}] is the linear pinch plasma mass, we obtain

$$H_\varphi = 10^{-10}\frac{\chi I^3}{rM^{1/2}} \text{ [W/cm]}. \qquad (2)$$

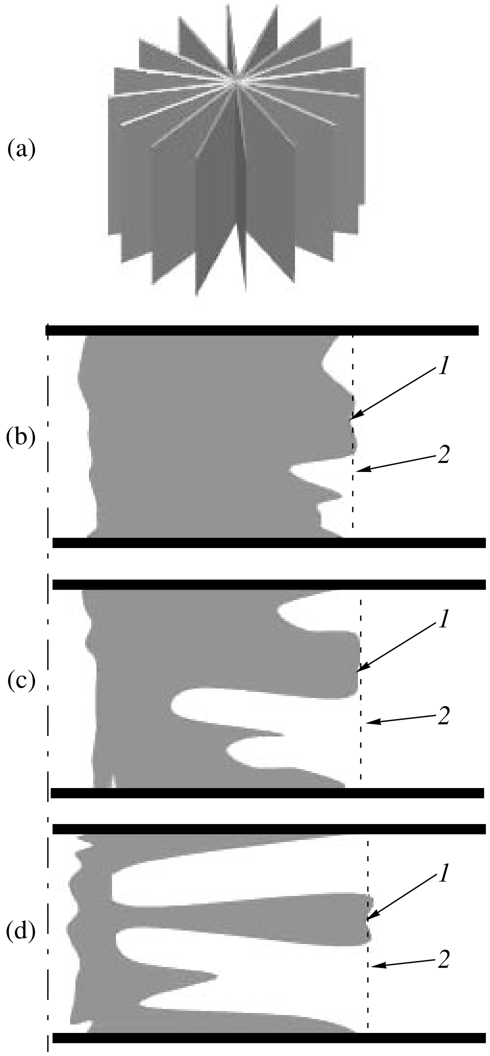

Fig. 10. A model representation of the imploding plasma of a multiwire array: (a) the orientation of the planes in which the plasma moves from the wires toward the array axis; (b), (c), and (d) the development of an instantaneous *r–z* plasma structure in one of the planes. The dark regions indicate the plasma; (*1*) the wire region where the plasma production is still taking place; (*2*) the region where the plasma production has already terminated.

Here, χ is a numerical dimensionless factor on the order of unity, whose value can actually be determined only experimentally. In [9, 10], it was calculated by assuming that part of the imploding pinch plasma is filled with toroidal magnetic flux tubes from the periphery converging toward the pinch axis. Assuming that the tube volume is small compared to the plasma volume, the authors of [8, 9] consider this model to be one-dimensional, which simplifies calculations.

Our experiments show that none of the above models corresponds to the actual dynamics of the fast *Z*-pinch of a multiwire tungsten array in the required measure. The model described in [9, 10] seems to better

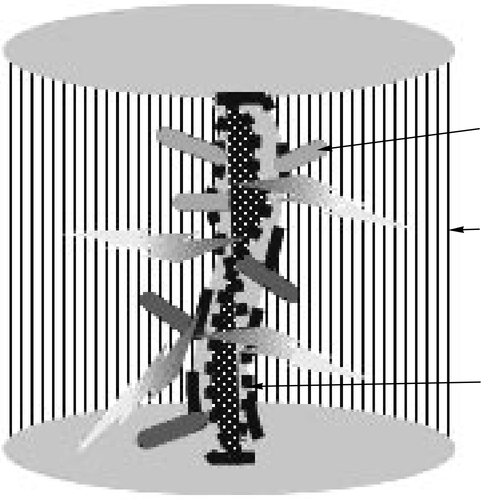

Fig. 11. A model representation of the structure of the radiating pinch plasma: (*1*) plasma outgrowths; (*2*) exploded wires; (*3*) hot spots.

agree with experiment, because it takes into account the MHD turbulent pattern of the imploded pinch. However, for the axially azimuthal inhomogeneity of the imploding plasma in an experiment, the concept of toroidal magnetic flux tubes [9, 10] is inapplicable, because the torus is an azimuthally homogeneous geometrical figure. In the model of a plasma rainstorm, the plasmoids transforming during plasma production and pinch implosion rather than the toroidal magnetic flux tubes against a homogeneous plasma background are the inhomogeneities. For this reason, to take into account the inhomogeneity structure in the plasma rainstorm, the above parameter χ should be calculated from different assumptions. In addition, the strong dependence of the energy transfer power W on the current I and radius r of the imploded plasma ($W \propto I^3/rM^{1/2}$) in the model presented in [9, 10] requires detailed knowledge of the radial current distribution in the pinch. The necessary theoretical modification of formula (2) is beyond the scope of our experimental study.

The plasma rainstorm shown in Figs. 2 and 4a can serve as an indicator for the implosion quality of a Z-pinch. The rainstorm occupies almost all of the space inside the array if not all of the plasma was collected in the paraxial zone by the time of intense pinch radiation; i.e., the plasma production is excessively prolonged, and the implosion quality is unsatisfactory (Fig. 2). In contrast, for the composite pinch with the best radiation parameters (Fig. 4), the rainstorm occupies only part of the space inside the inner array. In this case, the prolonged plasma production in the inner array delays the ultimate implosion of the entire composition, so the outer peripheral plasma has time to implode to a smaller radius.

5.3. Azimuthal Inhomogeneity of the Imploded State of a Z-Pinch

Let us now discuss the filaments—the azimuthal inhomogeneity of the imploded state of a Z-pinch. During the spatially inhomogeneous prolonged plasma production, the plasmoids are radially elongated clumps of dense plasma with relatively small diameters. As the plasmoids contract toward the array axis, they decrease in radial size and merge into isolated filaments—the plasma structures elongated mainly along the discharge axis.

The filamentary structures in Z-pinches are well known [11–16]. The necessary condition for their maintenance is an excess of the total discharge current I over the current J that corresponds to the Bennett equilibrium condition for the pinch as a whole:

$$I^2 > J^2 = 3.2 \times 10^{-22}(T_i + ZT_e)N_i. \tag{3}$$

Here, I and J are measured in MA, $N_i = M/m$ [cm^{-1}] is the linear ion density, m [g] is the ion mass, T_e and T_i [eV] are the electron and ion plasma temperatures, and Z is the effective ion charge.

The maximum possible number of current-carrying plasma filaments f at a given current I is determined by the condition for the maintenance of Bennett equilibrium for each filament:

$$f = \frac{I^2 m}{3.2 \times 10^{-22}(T_i + ZT_e)M}. \tag{4}$$

Naturally, the Bennett equilibrium of each filament does not hinder their group contraction.

Under typical experimental conditions at the Angara-5-1 facility for a total discharge current I = 2.5 MA and a linear tungsten plasma mass $M = 5 \times 10^{-4}$ g cm^{-1}, two current filaments will be maintained, for example, at an electron temperature of T_e = 200 eV, $Z \sim (T_e)^{0.5} \sim 14$, and an ion temperature of T_i = 3000 eV. We observed two filaments in an imploded pinch in a number of experiments. A larger number of filaments could be maintained in the pinch when the discharge current and the temperature and linear mass of the imploding plasma increased, because all of the plasma parameters in the denominator of formula (4) are small. However, in this period of process development, the plasma streams have such a pronounced axial inhomogeneity and radial orientation that it makes no sense to speak about the current filaments elongated along the discharge axis.

The number of filaments at the final stage of pinch implosion depends on many factors, including the initial azimuthal current distribution at discharge cold start and the magnetic-field diffusion time, which determines the rate of inhomogeneity smearing or merging of neighboring filaments.

We repeat in connection with the first of these factors that for the current-induced implosion of multiwire arrays, the initial conditions for the formation of structures are specified by the strong azimuthal inhomogeneity of the plasma-producing medium. Numerous current filaments are formed in discharges with a rapidly rising current through an azimuthally homogeneous medium (gas), which also merge into several current filaments during the implosion [1]. In this context, there is no qualitative difference between the implosion of a tungsten multiwire array and the implosion of a gas liner made of material with a large atomic number, for example, xenon. Quantitatively, the degree of implosion of multiwire arrays, especially composite ones, is higher than the degree of implosion of a xenon liner, because the initial plasma-production structure in the arrays is more regular than the current filaments randomly scattered with respect to azimuth in xenon.

We will estimate the smearing time τ of the plasmoids and filaments due to plasma diffusion in a magnetic field from the condition

$$\tau \sim \frac{4\pi\sigma\delta^2}{c^2\beta}, \quad (5)$$

where δ is the characteristic diameter of the structure, σ is the plasma electrical conductivity, c is the speed of light, and β is the ratio of the gas-kinetic plasma pressure to the magnetic plasma pressure. Assuming that $\beta \sim 1$, $\delta \sim 1$ mm, and $\sigma \sim 10^{14}$ s^{-1} for a state close to the imploded, equilibrium state, we obtain $\tau \sim 10^{-8}$ s. As we see, the estimated diffusion smearing time of filaments in the phase of an imploded pinch is on the order of the pinch pulse duration. At the stage of discharge development (~100 ns) under prolonged plasma production conditions (~70 ns), $\beta \ll 1$, so the diffusion smearing time scale for the forming inhomogeneities is much longer than 10^{-8} s. Consequently, the inhomogeneities will not be smeared at this time either. Thus, the estimated rates of diffusion smearing of the inhomogeneities both during the formation of a plasma rainstorm and in the imploded pinch state are consistent with the experimental evidence of their existence.

6. CONCLUSIONS

Our experimental results show that for the current-induced implosion of multiwire tungsten arrays, significant azimuthal and axial plasma inhomogeneities result from discharge cold start and prolonged plasma production, which determine the subsequent implosion dynamics of the generated plasma. The $[\mathbf{j} \times \mathbf{H}]$ forces cause the generated plasma to collapse toward the paraxial zone in the form of numerous radially elongated plasmoids with relatively small diameters. As the plasmoids contract toward the array axis, they decrease in radial size and merge into isolated plasma current filaments elongated mainly along the discharge axis. The plasma of the forming Z-pinch also remains spatially inhomogeneous at the time of intense X-ray radiation.

We call this stream of plasmoids a radial plasma rainstorm. Since the original inhomogeneities of the imploding plasma resulting from cold start and prolonged plasma production have a crucial effect on the final parameters of the multiterawatt radiating Z-pinch, we believe that the models of a superterawatt radiating Z-pinch in plasma of matter with a large atomic number that disregard the effects of cold start and prolonged plasma production do not correspond to reality.

ACKNOWLEDGMENTS

We thank G.S. Volkov for fruitful discussions and S.F. Medovshchikov for the technological maintenance of the experiments.

This study was supported by the Russian Foundation for Basic Research, project no. 02-02-17200.

REFERENCES

1. A. V. Branitskiĭ, V. V. Aleksandrov, E. V. Grabovskiĭ, *et al.*, Fiz. Plazmy **25**, 1060 (1999) [Plasma Phys. Rep. **25**, 976 (1999)].
2. V. V. Aleksandrov, A. V. Branitskiĭ, G. S. Volkov, *et al.*, Fiz. Plazmy **27**, 99 (2001) [Plasma Phys. Rep. **27**, 89 (2001)].
3. V. V. Alexandrov, I. N. Frolov, M. V. Fedulov, *et al.*, IEEE Trans. Plasma Sci. **30**, 559 (2002).
4. I. Frolov, E. Grabovsky, K. Mitrofanov, *et al.*, in *Proceedings of Euroconference on Advanced Diagnostics for Magnetic and Inertial Fusion, Varenna, Italy, 2001*, p. 415.
5. D. L. Peterson *et al.*, AIP Conf. Proc. **409**, 201 (1997).
6. M. G. Haines, IEEE Trans. Plasma Sci. **30**, 588 (2002).
7. J. P. Chittenden, S. V. Lebedev, A. R. Bell, *et al.*, Phys. Rev. Lett. **83**, 100 (1999).
8. M. H. Frese, S. D. Frese, S. E. Rosental, *et al.*, IEEE Trans. Plasma Sci. **30**, 593 (2002).
9. R. H. Lovberg, R. A. Riley, and J. S. Shlachter, AIP Conf. Proc. **299**, 59 (1993).
10. L. I. Rudakov, A. L. Velikovich, J. Davis, *et al.*, Phys. Rev. Lett. **84**, 3326 (2000).
11. I. F. Kvartskhava, K. N. Kervalidze, G. G. Zukakishvili, *et al.*, Yad. Sint. **3**, 285 (1963).
12. I. F. Kvartskhava, K. N. Kervalidze, Yu. S. Gvaladze, *et al.*, Yad. Sint. **5**, 181 (1965).
13. N. N. Komarov, Yad. Sint. **5**, 174 (1965).
14. N. N. Komarov, I. F. Kvartskhava, and V. M. Fadeev, Yad. Sint. **5**, 192 (1965).
15. G. G. Zukakishvili, I. F. Kvartskhava, and L. M. Zukakishvili, Fiz. Plazmy **4**, 725 (1978) [Sov. J. Plasma Phys. **4**, 405 (1978)].
16. A. B. Kukushkin and V. A. Rantsev-Kartinov, in *Proceedings of 27th European Physical Society Conference on Plasma Physics and Controlling Fusion* (Budapest, 2000).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 754–762.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 840–850.

SOLIDS
Structure

Magnetosensitive Defects Generated during Plastic Straining of NaCl:Eu Crystals

R. B. Morgunov[a],*, S. Z. Shmurak[a], A. A. Baskakov[a], and Y. Tanimoto[b]
[a]Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia
[b]Institute for Molecular Science, Myodaiji, Okazaki 444-8585, Japan
*e-mail: morgunov@issp.ac.ru
Received April 4, 2003

Abstract—The excitation of an equilibrium subsystem of point defects in plastically strained NaCl:Eu crystals has been studied. It is established that the macroplastic straining of crystals containing three-dimensional $EuCl_2$ precipitates leads to a decrease in the amount of these precipitates, accompanied by the appearance of new emission centers of two types. In the easy glide stage, the plastic straining is additionally accompanied by a decrease in the fraction of flat $EuCl_2$ precipitates and an increase in the content of nonequilibrium small-size complexes of impurity–vacancy dipoles sensitive to an external magnetic field. In thoroughly annealed crystals with the point defect subsystem in a nearly equilibrium state, individual dislocations exhibit displacement caused by the action of a pulsed magnetic field and by the joint action of constant and microwave magnetic fields under the electron paramagnetic resonance conditions. *© 2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

The influence of magnetic fields on the mechanical, optical, and electrical properties of diamagnetic solids with structural defects has been extensively studied (see reviews [1, 2]). The main common feature of these phenomena is that a low-energy magnetic field with an induction of ~1 T, capable only of transferring to a paramagnetic particle (or a structural defect) an energy two to three orders of magnitude below the average energy of thermal fluctuations, can lead to a significant (sometimes severalfold) change in the parameters of plasticity and other properties of the initial crystals.

This problem has been solved in the case of magnetoplastic effects (MPEs) related to a change in the structure of point defect clusters in ionic crystals. It was established [3–6] that the external magnetic field affects the structure of nonequilibrium intermediate clusters of point defects arising in the crystals as a result of aggregation of impurity–vacancy dipoles upon quenching. Previously, direct experimental evidence of the participation of spins in MPEs was obtained [5, 6]. This, as well as thermodynamic considerations developed in the physics of spin-dependent processes involving structural defects [7] and in the spin chemistry [8, 9] led to the conclusion that nonequilibrium conditions are necessary to provide for the effect of a weak magnetic field upon the physical properties of materials. At the same time, there are reliable and well-reproduced MPEs that do not require any intentional excitation of the equilibrium subsystem of defects for the manifestation of changes in the plasticity of crystals exposed to a magnetic field [10–18]. As is known, the motion of dislocations is capable of generating new defects [19, 20] and violating the equilibrium distribution of paramagnetic impurity ions over clusters of various types. In principle, this factor can favor the formation of magnetosensitive configurations of such clusters, similar to those appearing upon quenching, or complexes of the dislocation–vacancy type studied in [10–15].

The goals of this investigation were as follows: (i) to find and study the excited states of defects created by moving dislocations in aged NaCl:Eu crystals with an initial equilibrium distribution of Eu over the clusters of various configurations; (ii) to verify the hypothesis that magnetically sensitive Eu clusters can form in the course of plastic straining of aged and thoroughly annealed crystals; and (iii) to establish the possible mechanisms of defect nucleation and identify the types of nonequilibrium defects generated in the course of plastic straining.

2. EXPERIMENTAL METHODS

The choice of NaCl:Eu crystals for this investigation is related to the fact that Eu^{2+} ions are convenient spectroscopic labels for (i) detecting magnetic-field-induced changes via measurements of the photoluminescence (PL) spectra [21], (ii) determining the concentration of free impurity–vacancy dipoles, and (iii) monitoring the time variation of this concentration by means of the electron paramagnetic resonance (EPR) spectroscopy [22]. In addition, it was recently established experimentally [5] that Eu^{2+} ions participate in spin-dependent events influencing the plasticity of NaCl:Eu crystals in an external magnetic field.

The experiments were performed with NaCl:Eu single crystals of three types: (i) crystals aged for five years at room temperature after growth (Eu content, 0.1 at. %); (ii) crystals annealed for two weeks at 700°C (with one-week heating and cooling stages), containing 0.01 at. % Eu; and (iii) quenched samples prepared from aged crystals by heating for 1 h at 700°C and rapid cooling at an average rate of 5 K/s.

The macroplastic straining of crystals was performed either in an Instron machine operating in the strain rate interval from 10^{-5} to 10^{-3} s^{-1} or in a specially constructed hard machine combined with a setup for PL investigations. The PL spectra of the crystals were excited by UV radiation from a xenon lamp (DKSSh-120) passed through a monochromator (MDR-4). The spectra were detected in the wavelength range from λ = 400 to 650 nm using the second monochromator (MDR-6) and a photomultiplier (FÉU-106). Unless otherwise stated, the PL spectra were recorded at room temperature using excitation at a wavelength of 370 nm. For detecting the magnetic-field-induced changes in the PL spectra, the sample was placed into a solenoid in which magnetic field pulses having a half-sine shape with a duration of 10 ms and an amplitude of B = 6 T were generated.

The mobility of individual dislocations was studied by the method of double chemical etching. In the interval between the two etching steps, the sample was placed into a cylindrical cavity of a standard X-band radiospectrometer tuned to an EPR line related to the change in plasticity (for detail, see [4, 5, 23]). After exposure of the crystal in mutually perpendicular constant and microwave magnetic fields, a mean range L of the edge dislocations at a given value B of the magnetic induction was determined from the results of 100–200 measurements. The next analogous sample was used to determine L at a different B value. Using 40–50 samples, we obtained the field dependence $L(B)$ measured under otherwise identical conditions, including the same parameters of the microwave magnetic field.

3. EXPERIMENTAL RESULTS

3.1. *The Effect of Plastic Straining on the Photoluminescence of Crystals*

The typical PL spectrum of an unstrained aged crystal exhibited several bands (Fig. 1, ε = 0). The straining of a crystal led to redistribution of the emission intensity between various regions of the spectrum and to the appearance of a new, previously unreported band at λ = 536 nm (Fig. 1). The strain-induced changes in the PL spectrum were retained upon unloading of the crystal and could be observed during storage of the sample at room temperature over not less than three months. These changes occur in the bulk of the crystal, rather than in a surface layer, since the PL spectra measured from a freshly cleaved surface of a prestrained sample were identical to the spectra of strained crystals in Fig. 1.

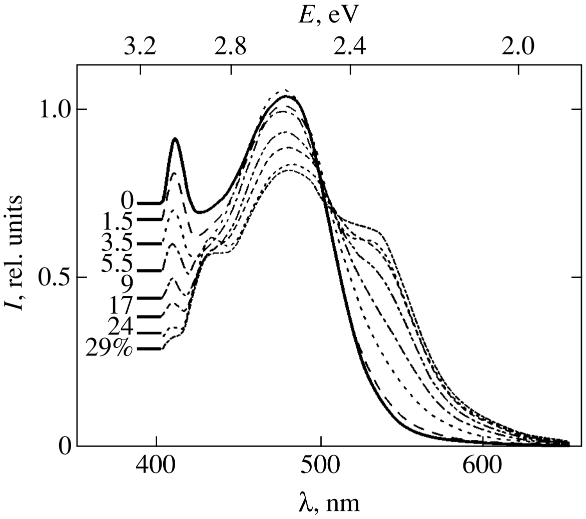

Fig. 1. The PL spectra of aged NaCl:Eu crystals strained to various relative deformations (ε = 0–29%).

In order to determine to what extent the changes observed in the PL spectrum are related to the motion of dislocations, the monochromator in the detection channel was tuned to λ = 410 or 536 nm and variation of the PL intensity at each of these wavelengths was measured as dependent on the relative deformation ε of the crystal, simultaneously with recording of the stress–strain (σ–ε) diagram. It was established that the PL spectrum remains virtually unchanged until reaching the yield point (see the inset in Fig. 2). When the yield point is reached and the intense multiplication of dislocations begins, the PL intensity at λ = 410 nm exhibits a sharp drop, while that at λ = 536 nm increases (Fig. 2). At the end of the easy glide stage, these changes slow down; in the stain hardening stage, the PL intensity in both spectral regions almost ceases to change with further increase in the relative plastic deformation. Thus, the strain-induced variations in the PL spectrum are sensitive to change in the stages of crystal straining, corresponding to the amount of dislocations and the regime of their motion.

The PL spectrum of a strained aged crystal could be decomposed into several Gaussian lines (Fig. 3) with strictly fixed centers of each component. These components include four previously reported emission bands centered at 410, 427, 439, and 480 nm (see, e.g., [24]). These components were characterized by the positions of their centers and by the linewidth known from [24]. Broadening of the left wing of the new component centered at λ = 536 nm, as well as a large rms deviation obtained in attempts at modeling the observed spectrum using only five components, allowed us to suggest that straining of the samples leads to the appearance of two, rather than one, new emission bands. In order to verify this assumption, the PL spectrum was measured using pulsed excitation by a nitrogen laser with λ = 337 nm and a pulse width of 10 ns, instead of the xenon lamp.

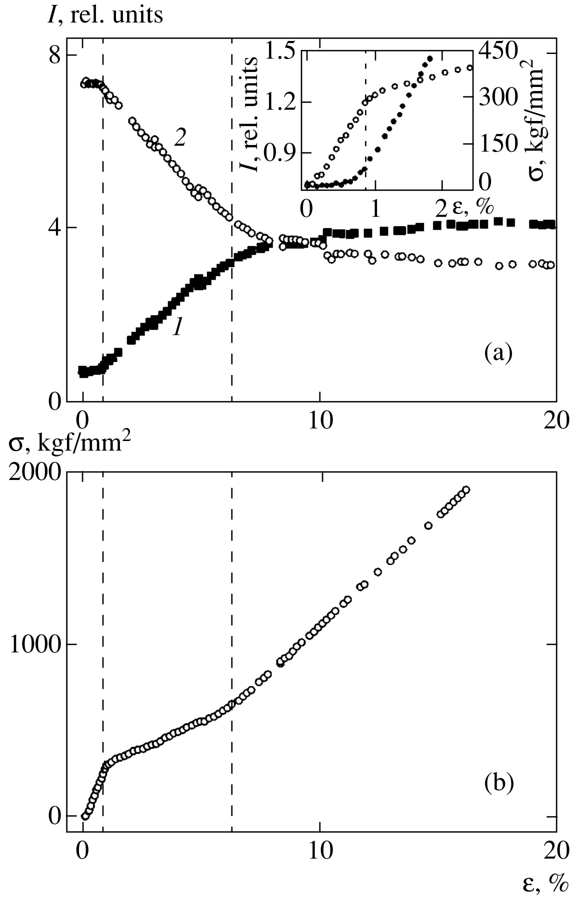

Fig. 2. Plots of (a) PL intensity *I* at λ = 536 (*1*) and 410 nm (*2*) and (b) mechanical stress σ versus relative deformation ε of aged NaCl:Eu crystals. Vertical dashed lines indicate the yield point (left) and the end of the easy glide stage (right). The inset in (a) shows the fragments of curve *1* and plot (b) in the vicinity of the yield point.

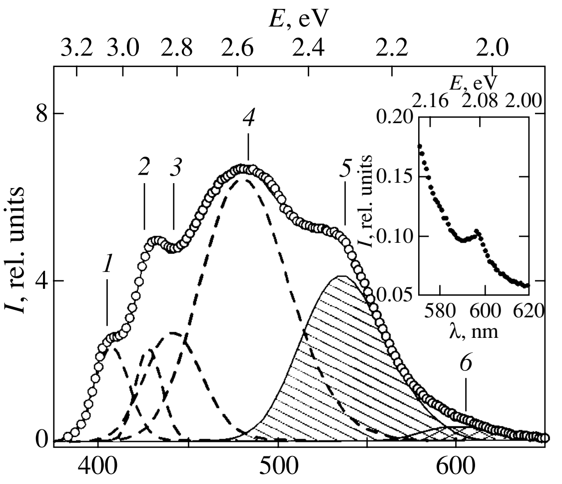

Fig. 3. The PL spectrum of an aged NaCl:Eu crystal strained to ε = 20% decomposed into Gaussian components centered at λ = 410 (*1*), 427 (*2*), 439 (*3*), and 480 nm (*4*) (known from [24]) and two new components centered at λ = 536 (*5*) and 599 nm (*6*). The inset shows a part of the PL spectrum excited by the radiation of a pulsed nitrogen laser with λ = 337 nm and a pulse width of 10 ns.

With this excitation, an additional PL peak was clearly pronounced at λ = 599 nm (see the inset to Fig. 3). Thus, the plastic straining of NaCl:Eu crystals leads to the appearance of two new emission bands. With allowance for these components, the PM spectrum was modeled using six Gaussian lines with a precision of 0.999. In the following investigation, the behavior of each band was traced separately.

In order to determine whether the new emission bands correspond to intrinsic or impurity defects, we studied the effect of straining on the luminescence of quenched NaCl:Eu crystals. The PL spectrum of a quenched crystal is simpler as compared to the spectra of other samples, since coarse precipitates are dissolved at elevated temperatures. The spectrum remained unchanged upon straining a quenched crystal up to ε = 25% (Fig. 4a). The new emission bands appeared neither in nominally pure NaCl crystals nor in the aged or quenched crystals of NaCl:Ca (containing 0.01 at. % Ca), although the dimensions of Ca^{2+} ions, the types of precipitates formed by this impurity, and the conditions of dislocation motion were analogous to those in the aged NaCl:Eu crystals. These results indicated that the new PL bands are related to the formation of centers that contained Eu, rather than to the *F*-like centers such as those reported in [19, 20]. This conclusion is also confirmed by similarity of the luminescence excitation spectrum of the new band to the excitation spectra of other known precipitates and complexes containing Eu (Fig. 4b).

The isochronous annealing of strained aged crystals at various temperatures, followed by cooling in air, leads to the disappearance of all emission bands (both strain-induced and those observed in the spectra of unstrained crystals) in favor of a single band at λ = 427 nm. Upon annealing, the intensity of this band exhibits a more than 20-fold increase (Fig. 5). The annealing temperature necessary for the complete suppression of emission in the bands at λ = 536 and 599 nm is significantly lower than that for the elimination of other bands (Fig. 5).

3.2. The Effect of Magnetic Field on the Photoluminescence of Plastically Strained Crystals

In the next series of experiments, the aged crystals were unloaded between sequential straining steps and subjected to the action of a pulsed magnetic field. The PL spectra were measured before and after the exposure of strained crystals to the magnetic field. The difference between the two spectra was evidence of the magnetic field effect upon the luminescence of samples. It was found that exposure of the NaCl:Eu crystals to magnetic field leads to a redistribution of the emission intensity between various components of the PL spectrum (Fig. 6). The study of variation of each separate component showed that the intensity of emission at λ = 427 nm decreases, while that of the band at 439 nm

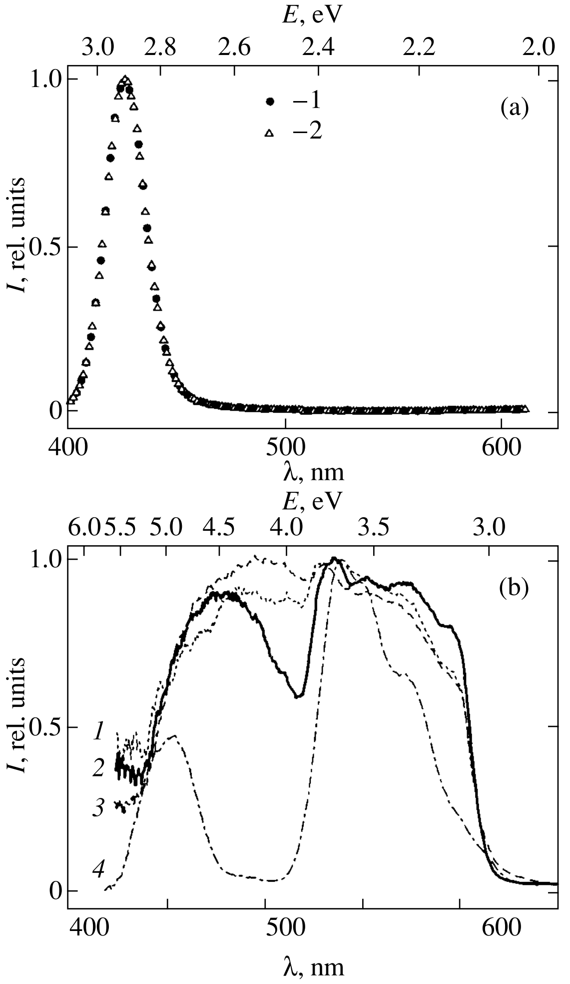

Fig. 4. (a) PL spectra of quenched NaCl:Eu crystals strained to $\varepsilon = 0$ (*1*) and 25% (*2*). (b) The excitation spectra for the PL bands at $\lambda = 410$ (*1*), 536 (*2*), 439 (*3*), and 427 nm (*4*) of an aged crystal strained to $\varepsilon = 20\%$.

increases. The other bands, including the new ones at $\lambda = 536$ and 599 nm, were not affected by the applied magnetic field (Fig. 6).

The external magnetic field influenced the PL spectra of strained crystals only in a narrow range of relative deformations ε corresponding to the easy glide stage (Fig. 7). The maximum effect of the applied magnetic field on the luminescence was always reached in this stage, whereas it was absent in the elastic region and in stages of strain hardening (Fig. 7).

For comparing the effect of straining to that of the magnetic field, the two treatments were applied in different orders to the same sample (Fig. 8). The exposure of the unstrained sample to magnetic field did not change the PL spectrum (Fig. 8). The straining increased the intensity of emission at $\lambda = 427$ nm, while the magnetic field decreased this intensity. After exposure to magnetic field, the crystals exhibited continuous changes accompanied by a slow decrease in intensity of the emission band at $\lambda = 427$ nm (Fig. 8). An analogous behavior was observed upon deformation of the crystal without exposure to magnetic field (Fig. 8). Note that the effect of magnetic field on the luminescence can be repeatedly observed on the same sample subjected to the next straining step after the previous exposure to magnetic field (Fig. 8). In the case of repeated exposures of a strained crystal to magnetic field, the PL spectrum exhibited a change only after the first pulse, while the subsequent pulses left the spectrum virtually unchanged.

3.3. The Effect of Magnetic Field on the Mobility of Individual Dislocations in Annealed Crystals

Previously [5], we reported on the MPE in NaCl:Eu crystals, which was manifested in the same stage of

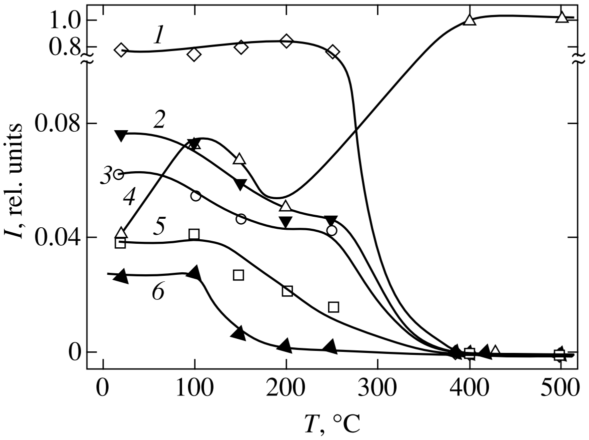

Fig. 5. Plots of the integral intensities I of separate components of the PL spectrum versus temperature T of isochronous annealing (15 min; air cooling) for aged NaCl:Eu crystals strained to $\varepsilon = 25\%$: $\lambda = 480$ (*1*), 439 (*2*), 410 (*3*), 427 (*4*), 536 (*5*), 599 nm (*6*).

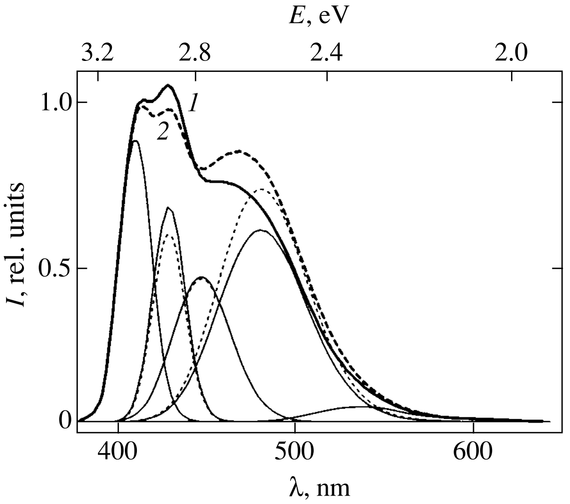

Fig. 6. PL spectra of aged NaCl:Eu crystals strained to $\varepsilon = 2\%$, recorded (*1*) before and (*2*) after exposure to magnetic field and decomposed into Gaussian components (depicted by thin solid and dashed lines, respectively).

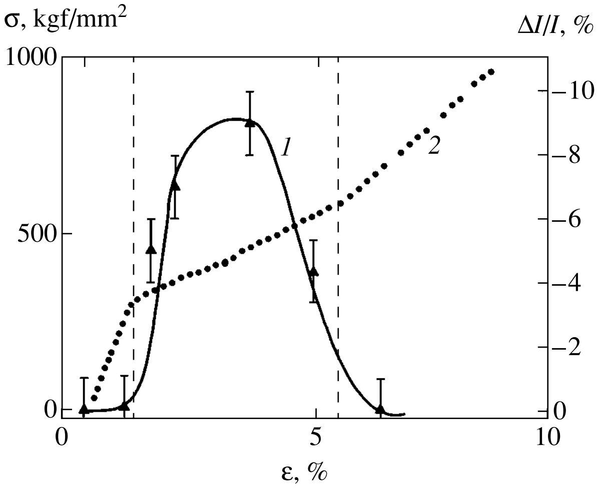

Fig. 7. Plots of (*1*) relative intensity change $\Delta I/I$ of the PL band at λ = 427 nm under the action of magnetic field and (*2*) mechanical stress σ versus the relative deformation ε (before exposure to the field) for aged NaCl:Eu crystals. Vertical dashed lines indicate the yield point (left) and the end of the easy glide stage (right).

impurity aggregation upon quenching as that in which we observe the effect of magnetic field on the luminescence [21]. For this reason, it was of principal interest to study the effect of magnetic field on the mean range of individual dislocations in the crystals upon prolonged annealing, that is, in a material where the distribution of impurity ions over the clusters of various configurations is close to equilibrium.

Upon introducing dislocations into such crystals via scratching of the sample surface followed by chemical etching, exposure to magnetic field (a 10-ms pulse of $B = 6$ T amplitude), and second etching step, it was found that the magnetic field produced a significant displacement of the dislocations as compared to that in the samples not exposed to the field. The exposure of such crystals in the crossed microwave and constant magnetic fields (exciting EPR in the structural defects) produced a shift of the individual dislocations, the range of which reached a maximum at certain discrete values of B (Fig. 9). This effect was observed only for the mutually perpendicular orientation of the constant and microwave magnetic fields, that is, the dislocations exhibited a shift only under the conditions of EPR for the spin-dependent transitions.

Therefore, the mechanism of magnetic field action upon plasticity is related to the field-induced spin transitions in structural defects. The observed MPE is analogous to that reported previously [5, 6], whereby the magnetic field changes the state of nonequilibrium complexes of point defects formed as a result of quenching.

4. DISCUSSION OF EXPERIMENTAL RESULTS

The observed variation of the PL spectra of NaCl:Eu crystals depending on the sample deformation stage (Fig. 2) indicates that the interaction between dislocations and precipitates leads to the formation of defects of new types containing Eu atoms. Therefore, analysis of the obtained experimental data requires knowledge of the atomic structure of Eu complexes present in the crystals prior to straining.

Figure 10 shows schematic diagrams representing some of the known complexes, the structures of which were established by X-ray diffraction in combination with data on contributions to the PL spectra [26]. The

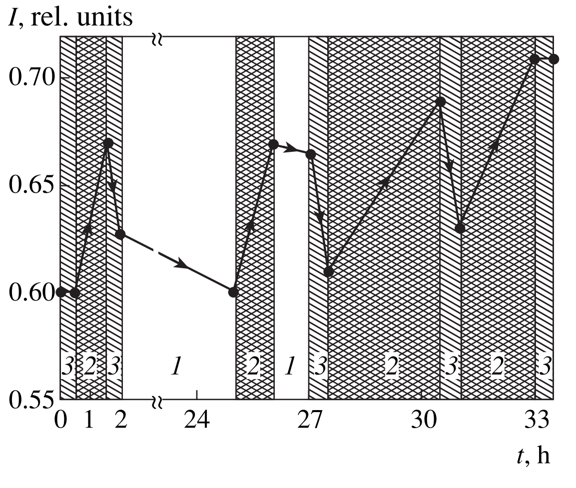

Fig. 8. Plot of integral intensity I of the PL band at λ = 427 nm versus current time t for an aged NaCl:Eu crystal in the course of various sequential treatments: (*1*) no external factors; (*2*) plastic straining; (*3*) magnetic field action (the time scale of exposures to magnetic field is strongly expanded).

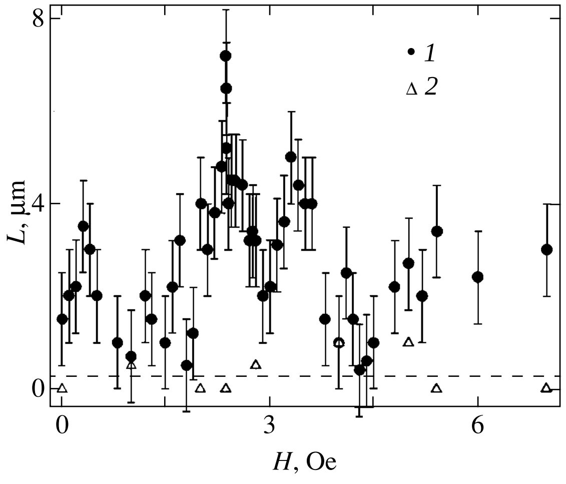

Fig. 9. Plot of mean range L of individual edge dislocations in aged NaCl:Eu crystals versus strength H of a constant magnetic field applied (*1*) simultaneously with an orthogonal microwave (9560 MHz) magnetic field (15 min) and (*2*) in the absence of a microwave field.

corresponding quantum yields were reported in [27]. The three-dimensional $EuCl_2$ precipitates with a fluorite structure are responsible for the emission band centered at λ = 410 nm (quantum yield, ~0.1). The flat $EuCl_2$ precipitates occurring in the (111) plane (Fig. 10b) and (301) plane (Fig. 10c) emit in the bands centered at λ = 439 and 480 nm, respectively, with a quantum yield of about 0.3. Far fewer data are available on the atomic structure of small-size clusters composed of a small number (below ten) of impurity–vacancy dipoles. The possible configurations of such complexes (Fig. 10a) were theoretically calculated by means of the free energy minimization [28]. It was reliably established that these defects, together with separate impurity–vacancy dipoles [26, 28–30], contribute to the PL band centered at λ = 427 nm; the quantum yield of these defects falls within 0.98–1.00 [26].

In order to interpret the observed pattern of changes in the PL spectrum of NaCl:Eu crystals under the action of plastic straining and magnetic fields, it is necessary to consider differences in the dependence of the intensity of various spectral components on the relative deformation ε (Fig. 11). The PL bands related to the three-dimensional precipitates, as well as the new band at λ = 536 nm, exhibit saturation on approaching ε = 10% (Fig. 11a), while changes in the intensity of bands due to the flat precipitates attain plateaus only at $\varepsilon \sim$ 20% and the PL band due to small-size clusters of impurity–vacancy dipoles exhibits no saturation at all (Fig. 11b). Additional grounds for dividing the spectral components into two groups, as presented in Figs. 11a and 11b, was provided by a comparative analysis of the absolute values of changes in the integral intensities of separate PL bands caused by plastic straining. There are strong changes in the intensity of PL bands due to the three-dimensional precipitates and the new band at λ = 536 nm, and less pronounced variations in the other parts of the PL spectrum. The same conclusion follows from an analysis of the effective radius of interaction between dislocations and precipitates, calculated using a method described in [31]. According to these estimates, the effective radius of the interaction of dislocations with the three-dimensional precipitates and the new defects is approximately 9–11 Å, while an analogous value for the emitting defects of other types is about 1–2 Å.

Thus, processes taking place under the action of plastic straining can be divided into two types: (i) transformation of the three-dimensional $EuCl_3$ precipitates into defects emitting at λ = 536 nm and (ii) transformation of the flat $EuCl_2$ precipitates into small-size clusters (precipitates cut by dislocations).

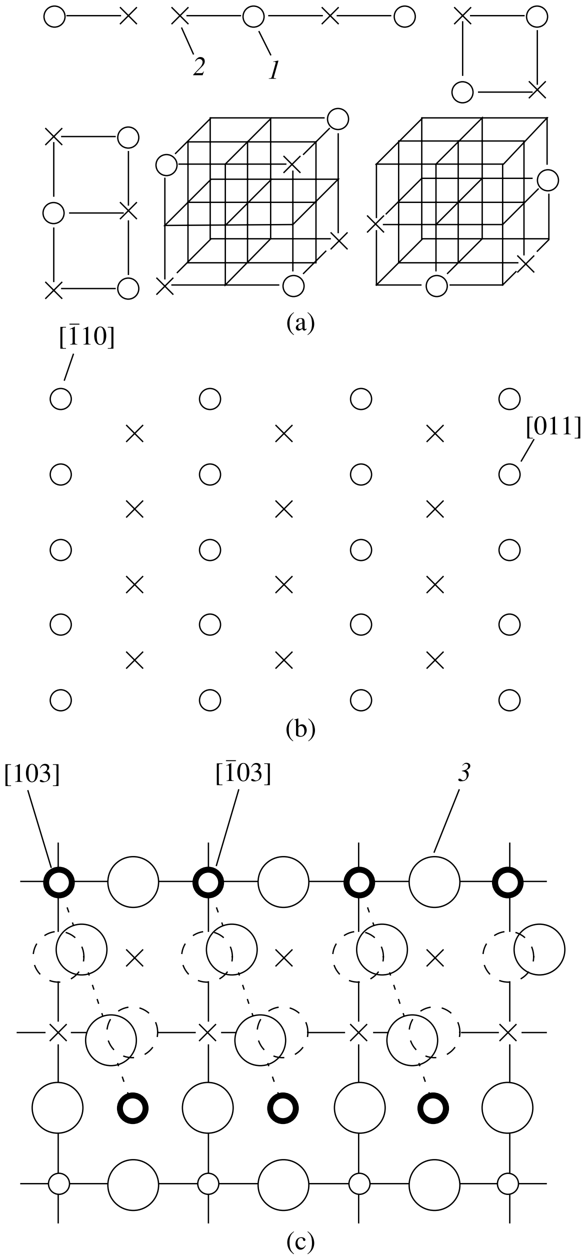

Fig. 10. Schematic diagrams showing the atomic structures of complexes as revealed by X-ray diffraction in NaCl:Eu crystals, representing (*1*) Eu^{2+} ions, (*2*) cationic vacancies V, and (*3*) Cl atoms [25]: (a) impurity–vacancy dipoles and related small-size clusters; (b) flat precipitates occurring in the (111) plane and emitting at λ = 439 nm; (c) flat precipitates occurring in the (310) plane and emitting at λ = 480 nm (depicted in projection onto the (111) plane).

The new emission band centered at λ = 599 nm cannot be unambiguously attributed to either of the above types because of weak intensity and a relatively small contribution to the total PL spectrum. However, the possibility of separating the component at λ = 599 nm by replacing continuous radiation of a xenon lamp with short excitation laser pulses (see Fig. 3 and the inset) is indicative of a difference in the lifetimes of the excited states emitting at 536 and 599 nm. Taking into account that only one type of radiative transitions can be excited in the interval of wavelengths used for the PL excitation in our experiments, we may conclude that each of the new emission bands corresponds to its own type of

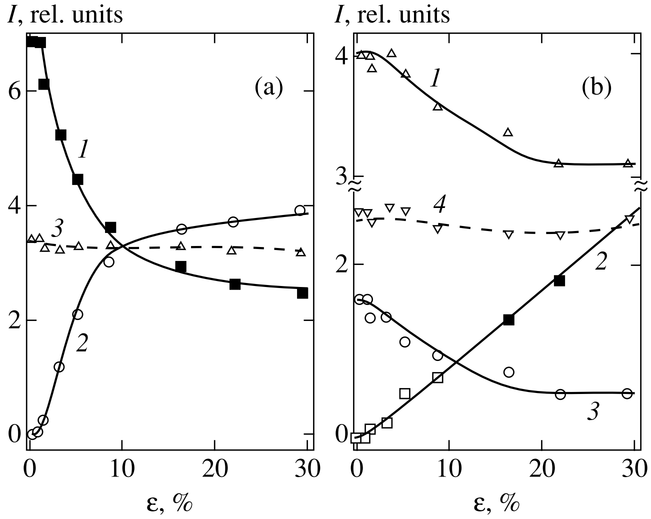

Fig. 11. Plots of the integral intensities I of separate components of the PL spectrum versus relative deformation ε of aged NaCl:Eu crystals: (a) components centered at λ = 410 (*1*) and 536 (*2*), and their average intensity (*3*); (b) components centered at 480 (*1*), 427 (*2*), and 439 (*3*) and their average intensity (*4*) determined taking into account the corresponding quantum yields.

defects. This conclusion is confirmed by the fact that the two bands disappear upon isochronous annealing at different temperatures (Fig. 5). The lower temperatures sufficient for the annealing of the PL bands at λ = 536 and 599 nm as compared to the temperatures required for the dissolution of $EuCl_2$ precipitates is evidence that the new defect clusters are less stable (i.e., they possess a lower binding energy) than the precipitates (Fig. 5).

It was established that the EPR spectra of aged NaCl:Eu crystals remain unchanged (to within a ~10% experimental accuracy) upon straining of the samples. Since, according to [32], only isolated impurity–vacancy dipoles contribute to these spectra, we may conclude that plastic straining does not generate such defects and only leads to the formation of small-size clusters. It should be noted that summation of the PL intensity variations in the separate emission bands, with allowance for their subdivision into two groups and the corresponding quantum yields, gives constant values virtually independent of the relative deformation level (dashed lines in Figs. 11a and 11b). This result confirms the correctness of the notion on precursor defects and the emission centers formed as a result of their interaction with dislocations.

The absence of changes in the PL spectra measured upon straining the quenched NaCl:Eu crystals (Fig. 4a), where the impurity occurs in the form of relatively mobile impurity–vacancy dipoles and small-size clusters (more effectively trapped by dislocations than the immobile $EuCl_2$ precipitates in the aged crystals), suggests that the process of impurity trapping by dislocation cores can hardly explain the strain-induced changes observed in the aged crystals. Moreover, the most intensive multiplication of dislocations and increase in their total length take place in the stage of strain hardening [33], but our experiments showed that the change in the PL spectra in this stage is rather insignificant (Fig. 2). In the region of saturation of the strain-induced changes in the PL spectra (i.e., for $\varepsilon > 10\%$), the emission bands due to precipitates retain rather large intensity. This implies that saturation of the $I(\varepsilon)$ curves is not caused by exhaustion of the source of related defects. Therefore, the amount of new defects formed in the course of plastic straining is determined by the area swept by moving dislocations, rather than by their total amount and length. This suggests that the new defects emitting at λ = 536 nm occur in the vicinity of the slip planes passed by dislocations, rather than in their cores. Then, saturation of the $I(\varepsilon)$ curves can be explained by a change in the regime of dislocation motion and the conditions of interaction between dislocations and precipitates in the stage of strain hardening.

Previously [22], we established that magnetic field is capable of influencing the luminescence of NaCl:Eu crystals via the transformation of small–size clusters of impurity–vacancy dipoles. A change in the structure of these dipoles in the applied magnetic field also leads to a change in the plasticity [3, 4]. For these reasons, it was quite natural to observe the effect of magnetic fields on the state of small–size clusters of impurity–vacancy dipoles formed due to the dislocations “cutting” the flat precipitates (Fig. 6). An increase in the intensity of emission at λ = 480 nm under the action of magnetic field suggests that the field increases the rate of relaxation of the nonequilibrium small-size complexes and stimulates their aggregation into two-dimensional $EuCl_2$ precipitates occurring in the (111) plane. Thus, the effect of the applied magnetic field is opposite to that of the plastic straining.

The dependence of the magnetic-field-induced change in the integral emission intensity $\Delta I/I$ at λ = 427 nm on the relative deformation ε resembles an analogous dependence for an MPE observed during the macroplastic straining of NaCl:Ca crystals [34]. The disappearance of that MPE in the stage of strain hardening was explained [34] by replacement of magnetosensitive obstacles by those of some other types (vacancies and dislocation trees) insensitive to the field. However, the present study revealed long-term (up to within 30 h) variations in the PL spectra of strained crystals after exposure to magnetic field (Fig. 8). This behavior indicates that changes in the PL spectra cannot be explained by the displacements of dislocations related to their spin-dependent interaction with a paramagnetic impurity. This mechanism of the magnetic field action (originally proposed in [10]) requires the presence of a magnetic field in the course of the dislocation motion. In our experiments, the magnetic field apparently influenced the structure of small-size clusters of impurity–vacancy dipoles in the volume of a crystal.

Vanishing of the magnetic field effect in the stage of strain hardening (Fig. 6) can be explained taking into account that the magnetic field pulse applied after each straining step reveals a relative intensity increment $\Delta I/I$ related to the formation of the corresponding portion of magnetosensitive clusters. Under the action of magnetic field, these clusters are irreversibly converted into different configurations incapable of contributing to ΔI during subsequent exposures to the magnetic field. We may suggest that the effective generation of vacancies and dislocation forest in the stage of strain hardening [34] is accompanied by cutting of the $EuCl_2$ precipitates, leading to the formation of clusters insensitive to the magnetic field. In this case, the $\Delta I/I$ curve measured in the course of repeated straining probably reflects an increase in the density of magnetosensitive complexes due to each sequential straining step.

In conclusion, it should be noted that, in contrast to MPEs observed in quenched crystals [6], the displacement of dislocations in annealed crystals exposed to magnetic field (Fig. 9) cannot be related to the formation of magnetosensitive small-size clusters in the course of impurity aggregation. Probably, the aforementioned processes, involving the precipitate being cut by dislocations and leading to changes in the optical properties of aged crystals, can also take place when dislocations are introduced into annealed samples. In such a case, the process of dislocation introduction is just what leads to the nucleation of magnetosensitive clusters. It was difficult to verify this assumption in the present study, because the aged crystals are characterized by a low mobility of dislocations and their displacements cannot be revealed by the etching techniques. In annealed crystals, the distribution of impurity changes in favor of small-size clusters, which hinders monitoring of the state of impurity by PL measurements. Nevertheless, the data presented in Fig. 9 allow us to ascertain that the introduction of dislocations leads to the formation of nonequilibrium magnetosensitive defect configurations, which can be modified by the magnetic-field-induced reorientation of spins in the short-lived states probably generated by thermal fluctuations (for more detail, see [5, 6]). The magnetic-field-induced changes in the structure of these defects leads to an increase in the mean range of dislocations in NaCl:Eu crystals. Such defects can be represented by quasimolecules of the dislocation–paramagnetic impurity center type or by the aforementioned small-size clusters contributing to the PL spectra.

5. CONCLUSIONS

1. It is established that plastic straining affects the PL spectra of NaCl:Eu crystals. Straining of the crystals causes redistribution of the intensity of emission from the defects existing in the initial material and leads to the appearance of new emission bands. It is found that the motion of dislocations initiates the conversion of three-dimensional $EuCl_2$ precipitates, emitting at $\lambda = 410$ nm, into previously unknown centers of emission with $\lambda = 536$ and 599 nm, the same quantum yield as that of the initial precipitates, and a lower binding energy. The flat $EuCl_2$ precipitates occurring in the (111) and (301) planes are broken as a result of the plastic straining and transformed into nonequilibrium small-size complexes of the impurity–vacancy dipole type.

2. The small-size complexes of the impurity–vacancy dipole type formed in the course of plastic straining include a fraction of defects sensitive to a magnetic field. The applied magnetic field accelerates the aggregation of such defects into $EuCl_2$ precipitates. Thus, the magnetic field increases the rate of processes opposite to those related to the plastic straining.

3. In thoroughly annealed NaCl:Eu crystals, where the point defect subsystem occurs in a nearly equilibrium state, freshly introduced edge dislocations exhibit displacement, both under the action of a pulsed magnetic field with an amplitude of 6 T and under the conditions of paramagnetic resonance in crossed constant and microwave fields. This magnetoplastic effect, observed in unquenched NaCl:Eu crystals, is probably related to the generation of nonequilibrium Eu clusters in the course of plastic straining after the introduction of dislocations.

4. The defects of new types and the corresponding emission bands can be used for the development of sensors and dose meters for various types of radiation, based on the strained ionic crystals doped with rare earth impurities [35–38]. The use of silicon-based photodetectors in such devices (see [35]) requires phosphors with the emission bands shifted toward the red part of the spectrum. It was demonstrated that this can be provided by plastic straining.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 01-02-16461) and by the Presidential Program of Support for Young Doctors of Science (project no. 02-15-99302).

REFERENCES

1. Yu. I. Golovin and R. B. Morgunov, Materialovedenie **2–5**, 2 (2000).
2. R. B. Morgunov, Vestn. Ross. Fonda Fundam. Issled. (2003).
3. R. B. Morgunov, A. A. Baskakov, I. N. Trofimova, and D. V. Yakunin, Fiz. Tverd. Tela (St. Petersburg) **45**, 257 (2003) [Phys. Solid State **45**, 270 (2003)].
4. R. B. Morgunov and A. A. Baskakov, Fiz. Tverd. Tela (St. Petersburg) **43**, 1632 (2001) [Phys. Solid State **43**, 1700 (2001)].
5. Yu. I. Golovin, R. B. Morgunov, V. E. Ivanov, and A. A. Dmitrievskiĭ, Zh. Éksp. Teor. Fiz. **117**, 1080 (2000) [JETP **90**, 939 (2000)].

6. Yu. I. Golovin and R. B. Morgunov, Zh. Éksp. Teor. Fiz. **115**, 605 (1999) [JETP **88**, 332 (1999)].
7. Yu. A. Osip'yan, S. I. Bredikhin, V. V. Kveder, N. V. Klassen, V. D. Negriĭ, V. F. Petrenko, I. S. Smirnova, S. A. Shevchenko, S. Z. Shmurak, and É. A. Shteĭnman, *Electronic Properties of Dislocations in Semiconductors* (Editorial URSS, Moscow, 2000), p. 46.
8. B. Ya. Zel'dovich, A. L. Buchachenko, and E. L. Frankevich, Usp. Fiz. Nauk **155** (1), 3 (1988) [Sov. Phys. Usp. **31**, 385 (1988)].
9. K. M. Salikhov, Yu. N. Molin, R. Z. Sagdeev, and A. L. Buchachenko, *Spin Polarization and Magnetic Field Effects in Radical Reactions*, Ed. by Yu. N. Molin (Nauka, Novosibirsk, 1978; Elsevier, Amsterdam, 1984).
10. V. I. Al'shits, E. V. Darinskaya, and O. L. Kazakova, Zh. Éksp. Teor. Fiz. **111**, 615 (1997) [JETP **84**, 338 (1997)].
11. V. I. Al'shits and E. V. Darinskaya, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 749 (1999) [JETP Lett. **70**, 761 (1999)].
12. A. A. Urusovskaya, V. I. Al'shits, A. E. Smirnov, and N. N. Bekkauer, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 470 (1997) [JETP Lett. **65**, 497 (1997)].
13. M. I. Molotskiĭ and V. N. Fleurov, Phys. Rev. B **52**, 15829 (1995).
14. N. A. Tyapunina, V. L. Krasnikov, and É. P. Belozerova, Fiz. Tverd. Tela (St. Petersburg) **41**, 1035 (1999) [Phys. Solid State **41**, 942 (1999)].
15. B. I. Smirnov, N. N. Peschanskaya, and V. I. Nikolaev, Fiz. Tverd. Tela (St. Petersburg) **43**, 2154 (2001) [Phys. Solid State **43**, 2250 (2001)].
16. A. M. Orlov and A. A. Skvortsov, Fiz. Tverd. Tela (St. Petersburg) **43**, 1207 (2001) [Phys. Solid State **43**, 1252 (2001)].
17. V. A. Makara, L. P. Steblenko, N. Ya. Gorid'ko, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **43**, 462 (2001) [Phys. Solid State **43**, 480 (2001)].
18. M. V. Badylevich, Yu. L. Iunin, V. V. Kveder, *et al.*, Zh. Éksp. Teor. Fiz. (in press).
19. V. A. Zakrevskiĭ and A. V. Shul'diner, Fiz. Tverd. Tela (St. Petersburg) **41**, 900 (1999) [Phys. Solid State **41**, 817 (1999)].
20. V. A. Zakrevskiĭ and A. V. Shuldiner, Philos. Mag. **71**, 128 (1995).
21. R. B. Morgunov, S. Z. Shmurak, A. A. Baskakov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 366 (2002) [JETP Lett. **76**, 307 (2002)].
22. R. B. Morgunov and A. A. Baskakov, Fiz. Tverd. Tela (St. Petersburg) **45** (1), 91 (2003) [Phys. Solid State **45**, 94 (2003)].
23. Yu. I. Golovin, R. B. Morgunov, and A. A. Baskakov, Mol. Phys. **100**, 1291 (2002).
24. J. Rubio, J. Phys. Chem. Solids **52**, 101 (1991).
25. A. E. Cordero-Borboa, O. Cano-Corona, A. Clevel-Hernandez, and E. Orozco, J. Phys. C **19**, 7113 (1986).
26. J. A. Munoz, E. Rodriguez, J. O. Tocho, and F. Cusso, J. Lumin. **72–74**, 233 (1997).
27. N. M. Bannon and J. Corish, Philos. Mag. A **51**, 797 (1985).
28. F. J. Lopez, H. S. Murrieta, A. Hernandez, and J. Rubio, Phys. Rev. B **22**, 6428 (1980).
29. J. E. Munoz-Santiuste and J. Garcia-Sole, Phys. Rev. B **38**, 10874 (1988).
30. J. Qi, M. Tanaka, J. S. Ahn, and Y. Masumoto, J. Lumin. **87–89**, 1102 (2000).
31. S. Z. Shmurak, Fiz. Tverd. Tela (St. Petersburg) **41**, 2139 (1999) [Phys. Solid State **41**, 1963 (1999)].
32. P. W. M. Jacobs, J. Phys. Chem. Solids **51**, 35 (1990).
33. B. I. Smirnov, *Structure and Strengthening of Crystals* (Nauka, Leningrad, 1981), p. 325.
34. Yu. I. Golovin and R. B. Morgunov, Pis'ma Zh. Éksp. Teor. Fiz. **61**, 583 (1995) [JETP Lett. **61**, 596 (1995)].
35. F. Jaque, G. Lifante, I. Aguirre de Carcer, *et al.*, Meas. Sci. Technol. **4**, 476 (1993).
36. J. Qi, M. Tanaka, J. S. Ahn, and Y. Masumoto, J. Lumin. **87–89**, 1102 (2000).
37. H. Nanto, Y. Hirai, M. Ikeda, *et al.*, Sens. Actuators A **53**, 223 (1996).
38. V. Chernov, V. Melendrez Ao, T. Piters, and M. Barboza-Flores, Radiat. Meas. **33**, 797 (2001).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 763–772.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 851–861.
Original Russian Text Copyright © 2003 by Geyler, Margulis, Pyataev.

SOLIDS
Electronic Properties

Resonant Tunneling through a Two-Dimensional Nanostructure with Connecting Leads

V. A. Geyler, V. A. Margulis*, and M. A. Pyataev
Mordvinian State University, Saransk, 430000 Russia
**e-mail: theorphysics@mrsu.ru*

Received April 21, 2003

Abstract—Ballistic electron transport through a two-terminal nanodevice is considered. An explicit formula is derived for the transmission coefficient of an electron as a function of its energy. An equation for the parameters of the Breit–Wigner and Fano resonances is obtained. Conditions are found under which a collapse of Fano resonances is observed. The effect that the parameters and the arrangement of contacts have on the transport properties of a system is investigated. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In the last decade, the study of resonant tunneling of electrons in quantum nanostructures has attracted constant interest. A number of theoretical and experimental works have been devoted to the study of resonances and minima of the transmission coefficient. Electron transport through the Aharonov–Bohm ring was considered in [1–4]. The tunneling of electrons through quantum dots connected to conducting leads in various ways was theoretically investigated in [5–7]. The transport properties of a circular ballistic interferometer were studied in [8–10]. A number of studies [11–15] were devoted to the analysis of tunneling in quasi-one-dimensional and two-dimensional channels containing impurities; it was shown that two types of resonances in the transmission coefficient are possible in these channels, ordinary Breit–Wigner resonances and asymmetric Fano resonances. Similar resonances were also observed in three-dimensional channels and constrictions [16, 17].

The Fano resonances are associated with the interference between the localized states of a discrete spectrum and propagating electron waves. Interference phenomena closely related to Fano resonances have recently been intensively studied both theoretically and experimentally. These resonances are of universal nature and have been observed in different systems: during the photoionization of atoms, in electron and ion scattering, in Raman scattering, and in photoabsorption in quantum wells and superlattices. Experimentally, Fano resonances of conductivity were observed in [18–20]. The increase in attention to this problem is due to the fact that resonant states are responsible for the operation of such quantum electron devices as a resonance tunnel diode and a resonance transistor. Recent achievements in the field of nanotechnologies made it possible to fabricate curved conducting two-dimensional nanostructures of different configurations. These include, first of all, fullerenes and fullerene-like structures with metallic-type chemical bonds, as well as metallic nanospheres. In view of this fact, a number of works have been devoted to the study of transport properties of a two-dimensional electron gas on curved surfaces [21, 22]. In these studies, the tunneling through nanostructures in which the motion of particles is finite attracts special interest because these structures may form a basis for a new generation of electron devices.

Until recently, theoretical studies of these nanostructures have mainly been focused on one-dimensional quantum rings; this fact is associated with the difficulties arising when simulating contacts between systems of different dimensions. Certain results in this direction were obtained in [23–25]. A general approach to the simulation of these contacts was developed in [26]. In the present paper, we apply this approach to investigate ballistic electron transport through a two-terminal nanodevice consisting of a nanostructure **S** of finite dimension with two conducting leads $\mathbf{R}_1^+$ and $\mathbf{R}_2^+$ connected to it. The connecting leads are assumed to be one-dimensional and represented by the half-lines $x \geq 0$. As for the nanostructure **S**, a general method for deriving a transmission coefficient that we apply in this study does not impose any special restrictions on its geometry. In particular, it may be a domain of a three-dimensional or a two-dimensional space (a quantum dot), a surface (quantum sphere, quantum torus, etc.), or a one-dimensional quantum ring. The Schrödinger equation that describes the motion of an electron in such a structure may either be a free Hamiltonian or contain potentials, including the vector potential of an external magnetic field. In this paper, we consider in detail a sufficiently general case when the nanostructure **S** represents an arbitrary two-dimensional surface of finite size. It is important that the main results of this paper correspond to the case of nanostructures of arbitrary configuration.

2. THE HAMILTONIAN OF THE SYSTEM AND THE TRANSMISSION COEFFICIENT

The motion of electrons in the one-dimensional conducting lead $\mathbf{R}_j^+$ is described by the free Hamiltonian

$$H_j = p_x^2/2m,$$

where m is the effective mass of an electron, p_x is the operator of the electron momentum in the connecting lead, and $j = 1, 2$ is the lead number. Denote the points at which the conducting leads are connected to the nanostructure by $\mathbf{q}_1$ and $\mathbf{q}_2$. The wave functions of an electron in the nanostructure **S** and in the leads $\mathbf{R}_j^+$ are denoted by ψ_S and ψ_j, respectively.

To derive the Hamiltonian H of the entire system, one has to impose linear boundary conditions on the wave function at the gluing points. As usually, the role of boundary parameters in the one-dimensional connecting leads $\mathbf{R}_j^+$ is played by $\psi_j(0)$ and $\psi'_j(0)$. To obtain the boundary conditions at points $\mathbf{q}_j$ on the surface of **S**, one has to consider functions that have a logarithmic singularity at these points [26]:

$$\psi_S(\mathbf{x}) = -u_j(\psi_S)\frac{m}{\pi\hbar^2}\ln\rho(\mathbf{x}, \mathbf{q}_j) + v_j(\psi_S) + R(\mathbf{x}). \quad (1)$$

Here, $\rho(\mathbf{x}, \mathbf{q}_j)$ is the geodesic distance between points $\mathbf{x}$ and $\mathbf{q}_j$ on the surface, $R(\mathbf{x})$ is a remainder term that tends to zero as $\mathbf{x} \longrightarrow \mathbf{q}_j$, and $u_j(\psi_S)$ and $v_j(\psi_S)$ are complex coefficients that play the role of boundary values for ψ_S at points $\mathbf{q}_j$ on **S**. The necessity of considering singular wave functions follows from the fact that the Green function $G_S(\mathbf{x}, \mathbf{y}; E)$ of the operator H_S has a similar singularity for $\mathbf{x} \longrightarrow \mathbf{y}$.

It is clear that the boundary conditions that should to be imposed on the wave function at the contact points represent linear relations between $\psi_j(0)$, $\psi'_j(0)$ and $u_j(\psi_S)$, $v_j(\psi_S)$. Following [26], we write these relations as

$$\begin{aligned} v_j(\psi_S) &= \sum_{k=1}^{2}\left[B_{jk}u_k(\psi_S) - \frac{\hbar^2}{2m}A_{jk}\psi'_k(0)\right], \\ \psi_j(0) &= \sum_{k=1}^{2}\left[\bar{A}_{kj}u_k(\psi_S) - \frac{\hbar^2}{2m}C_{jk}\psi'_k(0)\right], \\ j &= 1, 2. \end{aligned} \quad (2)$$

Here, the elements A_{jk}, B_{jk}, and C_{jk} form complex 2×2 matrices; the self-adjointness of the operator H requires that matrices B and C should be Hermitian. The physical meaning of the elements of these matrices can easily be revealed by invoking the theory of zero-range potentials [27]. From the viewpoint of this theory, the diagonal elements of matrix B determine the magnitude of point perturbations of operator H_S at points $\mathbf{q}_j$ of nanostructure **S**. The off-diagonal elements of this matrix correspond to the tunnel coupling between points $\mathbf{q}_1$ and $\mathbf{q}_2$; therefore, if the effective size of the contacts is much smaller than the distance between them, these elements should be set equal to zero (see [28] for more details). Similarly, we assume that matrix C is diagonal and interpret its element C_{jj} as the intensity of a zero-range potential at the point $x = 0$ in $\mathbf{R}_j^+$.

Now, consider matrix A. As is clear from formula (2), it characterizes the coupling between the connecting leads and nanostructure **S**. Indeed, if $A = 0$, then boundary conditions (2) decompose into four independent equations, two of which contain only the wave function on **S**, while each of the remaining equations contains a wave function in one of the connecting leads; i.e., there is no coupling between the leads and the nanostructure. If $A_{jk} \neq 0$ for $j \neq k$, then there exist nontrivial boundary conditions that relate lead $\mathbf{R}_j^+$ to point $\mathbf{q}_k$; this corresponds to the tunneling transition of electrons from one connecting lead to another, leaving **S** aside. Therefore, in the case of one-dimensional connecting leads, matrix A should also be diagonal. In view of the aforesaid, the boundary conditions for the wave functions are chosen as

$$\begin{aligned} v_j(\psi_S) &= B_{jj}u_j(\psi_S) - \frac{\hbar^2}{2m}A_{jj}\psi'_j(0), \\ \psi_j(0) &= \bar{A}_{jj}u_j(\psi_S) - \frac{\hbar^2}{2m}C_{jj}\psi'_j(0), \quad j = 1, 2. \end{aligned} \quad (3)$$

According to the theory of scattering by zero-range potentials, the elements of matrix B that describe point perturbations at points $\mathbf{q}_j$ on **S** and the elements of matrix C that describe point perturbations at points $x = 0$ in the connecting leads can be expressed in terms of scattering lengths λ_j^B and λ_j^C on these perturbations:

$$B_{jj} = -\frac{m\ln\lambda_j^B}{\pi\hbar^2}, \quad C_{jj} = -\frac{m\lambda_j^C}{2\hbar^2}.$$

To further facilitate calculations, let us introduce scattering lengths λ_j^A for parameters A_{jj} by the formula

$$A_{jj} = \frac{m}{\hbar^2}\sqrt{\lambda_j^A}e^{i\varphi_j},$$

where ϕ_j is the argument of the complex number A_{jj}.

When energies do not coincide with eigenvalues E_n of Hamiltonian H_S, a solution to the Schrödinger equation, normalized by the incident wave of unit amplitude in $\mathbf{R}_1^+$, is given by

$$\begin{gathered}\psi_S(\mathbf{x}) = \xi_1(E)G_S(\mathbf{x}, \mathbf{q}_1; E) + \xi_2(E)G_S(\mathbf{x}, \mathbf{q}_2; E), \\ \psi_1(x) = e^{-ikx} + r_{11}(E)e^{ikx}, \quad \psi_2(x) = t_{12}(E)e^{ikx}.\end{gathered} \tag{4}$$

Here, $k = \sqrt{2mE}/\hbar$ is the electron wave vector, $\xi_1(E)$ and $\xi_2(E)$ are complex numbers, and $r_{11}(E)$ and $t_{12}(E)$ are the amplitude coefficients of the reflection and transmission of an electron, respectively. Considering asymptotics (1) for the wave function $\psi_S(\mathbf{x})$ from (4) in the neighborhood of points $\mathbf{q}_1$ and $\mathbf{q}_2$, we obtain

$$u_j = \xi_j(E), \quad v_j = Q_{jj}(E)\xi_j(E) + Q_{ji}(E)\xi_i(E).$$

Here, $Q(E)$ is a 2×2 matrix whose elements are expressed in terms of the Green function $G_S(\mathbf{x}, \mathbf{y}; E)$ by the formula

$$Q_{ij}(E) = \begin{cases} G_S(\mathbf{q}_i, \mathbf{q}_j; E), & i \neq j \\ \lim\limits_{\mathbf{q}_i \to \mathbf{q}_j}\left[G_S(\mathbf{q}_j, \mathbf{q}_i; E) + \dfrac{m}{\pi\hbar^2}\ln\rho(\mathbf{q}_j, \mathbf{q}_i)\right], & i = j. \end{cases} \tag{5}$$

Substituting a solution in the form of (4) into boundary conditions (3), we obtain the following system of equations for the coefficients r_{11}, t_{12}, ξ_1, and ξ_2:

$$\begin{gathered}\frac{ik\hbar^2 A_{11}}{2m}r_{11} + (Q_{11}(E) - B_{11})\xi_1 + Q_{12}(E)\xi_2 = \frac{ik\hbar^2 A_{11}}{2m}, \\ \frac{ik\hbar^2 A_{22}}{2m}t_{12} + Q_{21}(E)\xi_1 + (Q_{22}(E) - B_{22})\xi_2 = 0, \\ \left(1 + \frac{ik\hbar^2 C_{11}}{2m}\right)r_{11} - \bar{A}_{11}\xi_1 = \frac{ik\hbar^2 C_{11}}{2m} - 1, \\ \left(1 + \frac{ik\hbar^2 C_{22}}{2m}\right)t_{12} - \bar{A}_{22}\xi_2 = 0.\end{gathered} \tag{6}$$

For convenience, we introduce dimensionless elements of matrix Q by the formula

$$\tilde{Q}(E) = \frac{\hbar^2}{m}[Q(E) - B]$$

and express the elements of matrices A and C in terms of scattering lengths λ_j^A and λ_j^C introduced above. Then, from system (6), we obtain

$$\begin{aligned}r_{11}(E) = \Delta^{-1}(E)[&(k\lambda_1^C - 4i)(k\lambda_2^C + 4i)\det\tilde{Q}(E) \\ &- 2k\lambda_2^A(k\lambda_1^C - 4i)\tilde{Q}_{11}(E) \\ &- 2k\lambda_1^A(k\lambda_2^C + 4i)\tilde{Q}_{22}(E) + 4k^2\lambda_1^A\lambda_2^A],\end{aligned} \tag{7}$$

$$\begin{aligned}t_{12}(E) &= 16ik\exp[i(\phi_1 - \phi_2)] \\ &\times \sqrt{\lambda_1^A\lambda_2^A}\tilde{Q}_{21}(E)\Delta^{-1}(E),\end{aligned} \tag{8}$$

where

$$\begin{aligned}\Delta(E) = {} &(k\lambda_1^C + 4i)(k\lambda_2^C + 4i)\det\tilde{Q}(E) \\ &- 2k\lambda_2^A(k\lambda_1^C + 4i)\tilde{Q}_{11}(E) \\ &- 2k\lambda_1^A(k\lambda_2^C + 4i)\tilde{Q}_{22}(E) + 4k^2\lambda_1^A\lambda_2^A.\end{aligned} \tag{9}$$

One can easily verify that the scattering matrix is unitary; in particular, relation $|r_{11}|^2 + |t_{12}|^2 = 1$ holds, which represents a current conservation law. Note that formulas (7) and (8) are inapplicable when $\rho(\mathbf{q}_1, \mathbf{q}_2) \longrightarrow 0$ because one cannot neglect the off-diagonal components of matrices A, B, and C in this case.

3. RESONANCE STRUCTURE OF THE TRANSMISSION COEFFICIENT

Formulas (8) and (9) show that the transmission coefficient as a function of energy contains six phenomenological parameters: λ_j^A, λ_j^B, and λ_j^C ($j = 1, 2$). Although the scattering lengths have a considerable effect on the function $t_{12}(E)$ in the general case, there are a number of phenomena that are independent of these parameters. We begin our analysis with these phenomena and then discuss the effect of the scattering lengths.

In the general case, the transmission coefficient $T_{12}(E) \equiv |t_{12}(E)|^2$ of an electron from the first connecting lead to the second has zeros of two different types; the zeros of the first type are at the points of spectrum E_n of operator H_S and are attributed to the poles of $Q_{ij}(E)$ at these points. In this case, the denominator in (8) generally has a pole of the second order, while the numerator has a pole of the first order; therefore, transmission coefficient $T_{12}(E)$ vanishes at these points. Note that the positions of the first-type zeros are independent of the distance between contacts. The zeros of the second type are attributed to the numerator in (8) and coincide with the zeros of function $\tilde{Q}_{21}(E)$. The positions of these zeros on the energy axis depend on the arrangement of points $\mathbf{q}_1$ and $\mathbf{q}_2$ on the surface of $\mathbf{S}$.

It follows from boundary conditions (3) that the transmission coefficients vanish if and only if the electron wave function $\psi_S(\mathbf{x})$ vanishes at point $\mathbf{q}_2$ of the second contact. When an incident wave of unit amplitude propagates along the first connecting lead, the coefficients $\xi_1(E)$ and $\xi_2(E)$ in (4) are expressed as

$$\xi_1(E) = \frac{4\hbar^2 k\sqrt{\lambda_1^A}}{m\Delta(E)}\exp(i\phi_1) \times [2k\lambda_2^A - (k\lambda_2^C + 4i)\tilde{Q}_{22}(E)], \tag{10}$$

$$\xi_2(E) = \frac{4\hbar^2 k\sqrt{\lambda_1^A}}{m\Delta(E)}\exp(i\phi_1)(k\lambda_2^C + 4i)\tilde{Q}_{21}(E). \tag{11}$$

Since the Green function represents a wave outgoing from a point source, one can say that each term $\xi_j(E)G_S(\mathbf{x}, \mathbf{q}_j; E)$ in (4) represents a superposition of waves incoming to and outgoing from the point $\mathbf{q}_j$. As is clear from Eqs. (4) and (11), when $\tilde{Q}_{21}(E) = 0$, the wave $\xi_1(E)G_S(\mathbf{x}, \mathbf{q}_1; E)$ outgoing from the point $\mathbf{q}_1$ has a node at the point $\mathbf{q}_2$ since $G_S(\mathbf{q}_1, \mathbf{q}_2; E)$ vanishes in this case. Then, there is no wave incoming to point $\mathbf{q}_2$ since coefficient $\xi_2(E)$ vanishes and the probability that an electron is situated near the second contact also vanishes; as a consequence, the electron does not reach the second connecting lead and the transparency of the system proves to be zero.

To investigate the electron wave function for energies coinciding with eigenvalues E_n of operator H_S, we represent the elements of matrix $\tilde{Q}_{ij}(E)$ as

$$\tilde{Q}_{ij}(E) \approx \frac{\alpha_{ij}}{E - E_n} + \beta_{ij}. \tag{12}$$

Then,

$$\det\tilde{Q}(E) \approx \frac{\det\alpha}{(E - E_n)^2} + \frac{\chi}{E - E_n}, \tag{13}$$

where

$$\chi = \alpha_{11}\beta_{22} + \alpha_{22}\beta_{11} - \alpha_{21}\beta_{21} - \alpha_{21}\beta_{12}.$$

It follows from (5) that coefficients α_{ij} can be expressed as

$$\alpha_{ij} = -\frac{\hbar^2}{m}\sum_{l=1}^{\nu}\varphi_l(\mathbf{q}_i)\varphi_l^*(\mathbf{q}_j), \tag{14}$$

where $\varphi_l(\mathbf{x})$ is an orthonormal basis of the proper subspace of operator H_S that corresponds to a ν-fold degenerate level E_n. Substituting asymptotics (12) of functions $Q_{ij}(E)$ into solution (4) of the Schrödinger equation and passing to the limit as $E \longrightarrow E_n$, we obtain the following expression for $\psi_S(\mathbf{x})$ in the case $\det\alpha \neq 0$:

$$\psi_S(\mathbf{x}) = \frac{4\hbar^2 k\sqrt{\lambda_1^A}e^{i\phi_1}}{m(k\lambda_1^C + 4i)\det\alpha} \times \sum_{l=1}^{\nu}\varphi_l(\mathbf{x})[\alpha_{22}\varphi_l^*(\mathbf{q}_1) - \alpha_{21}\varphi_l^*(\mathbf{q}_2)]. \tag{15}$$

In this case, the wave function has no singularities at the contact points and cannot be interpreted as a superposition of incoming and outgoing waves. One may say that, when $E = E_n$, the electron wave function on the surface of **S** describes a standing wave with a node at point $\mathbf{q}_2$.

In addition to the functions defined by (15), the solution of the Schrödinger equation for H at $E = E_n$ also contains linear combinations of the form

$$\psi_S(\mathbf{x}) = \sum_{l=1}^{\nu} C_l\varphi_l(\mathbf{x}) \tag{16}$$

that vanish at both gluing points. In contrast to the scattering states that contain $\exp(\pm ikx)$, wave functions (16) can be normalized and correspond to a discrete energy level embedded into a continuous spectrum. The electrons in these states are localized on the surface of **S** and do not contribute to the conductivity. In the general case, when $\det\alpha \neq 0$, the multiplicity of degeneracy of the bound states in the spectrum of Hamiltonian H is two times less than the multiplicity of degeneracy ν of the corresponding levels of unperturbed Hamiltonian H_S because the requirement that the wave functions should vanish at the contact points imposes two additional conditions on coefficients C_l of linear combination (16). However, if $\det\alpha$ vanishes for a certain arrangement of contacts, then these conditions prove to be equivalent and the multiplicity of degeneracy decreases only by one. Moreover, there may exist points on a general surface such that all eigenfunctions of an appropriate level may vanish at these points; in this case, the connection of conducting leads to these point does not change the multiplicity of this level.

Thus, when $E = E_n$, two types of wave functions are possible, localized and delocalized. This is associated with the fact that one can prepare two different types of states with the same energy by a different choice of boundary conditions.

Now, consider how the spectral properties of operator H affect the electron transport in the system under investigation. The presence of zeros of the transmission coefficient at certain points and the presence of discrete energy levels embedded into the continuous spectrum provide evidence for the resonant nature of the scattering. It is well known that the resonance peaks of the transmission coefficient are associated with the poles of the scattering matrix in the complex energy plane. Since the scattering matrix is unitary for real values of energy, scattering amplitude $t_{12}(E)$ has no poles on the real axis. However, there are poles on the nonphysical sheet of the Riemann surface; the positions of these poles are determined by equation $\Delta(E) = 0$. In the neighborhoods of these poles, transmission coefficient (8) can be represented as

$$t_{12}(E) \approx \eta \frac{i\Gamma}{E - E_R - i\Gamma}, \tag{17}$$

where E_R is energy that specifies the position of a resonance, Γ is the half-width of the resonance curve, and η is a normalizing constant. Both resonance parameters E_R and Γ can be obtained from the solution of the transcendental equation $\Delta(E) = 0$.

The analysis below shows that, along with the ordinary Breit–Wigner resonances described by formula (17), the system also exhibits asymmetric Fano resonances that are characterized by a closely spaced zero and a peak of the scattering amplitude. The shapes of the Breit–Wigner and Fano resonances in the example considered below—electron transport through a quantum torus—are shown in Fig. 1. A similar shape of resonances is also characteristic of other cases.

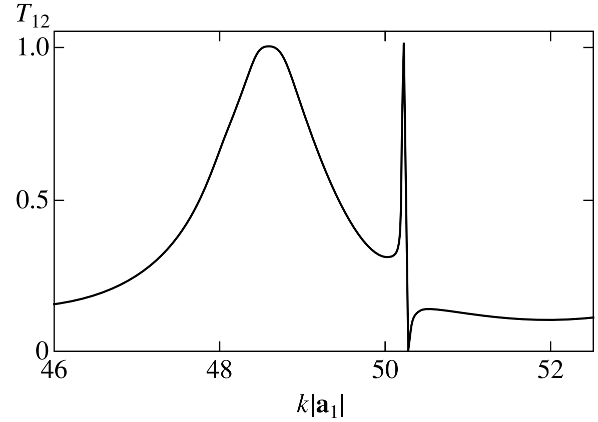

Fig. 1. Characteristic shape of a Breit–Wigner resonance (in the region $47 < k|\mathbf{a}_1| < 49$) and a Fano resonance (in the region $49 < k|\mathbf{a}_1| < 51$) for the case of a quantum torus. Here, $|\mathbf{a}_1|$ is the length of the greater generator of the torus.

In the system considered, the zeros of the transmission coefficient associated with Fano resonances are situated at the points of the spectrum of unperturbed Hamiltonian H_S. Using asymptotics (12) of function $Q_{ij}(E)$ in the neighborhood of E_n and retaining terms linear in $E - E_n$ in the numerator and denominator of (8), we represent transmission coefficient $t_{12}(E)$ in the neighborhood of E_n as

$$t_{12}(E) \approx \eta \frac{E - E_n}{E - E_R - i\Gamma}, \tag{18}$$

where

$$\begin{aligned} E_R + i\Gamma &= E_n - \frac{(k\lambda_1^C + 4i)(k\lambda_2^C + 4i)\det\alpha}{D(E)}, \\ \eta &= \frac{16ik\sqrt{\lambda_1^A\lambda_2^A}\exp[i(\phi_1 - \phi_2)]\alpha_{21}}{D(E)}, \end{aligned} \tag{19}$$

$$\begin{aligned} D(E) &= (k\lambda_1^C + 4i)(k\lambda_2^C + 4i)\chi \\ &- 2k\lambda_2^A(k\lambda_1^C + 4i)\alpha_{11} - 2k\lambda_1^A(k\lambda_2^C + 4i)\alpha_{22}. \end{aligned} \tag{20}$$

It is clear from (18) that the energy dependence of the transmission coefficient in the neighborhood of eigenvalues E_n of operator H_S indeed has the structure of a Fano resonance. Formula (19) shows that Fano parameters E_R and Γ are related to scattering lengths λ^A, λ^B, and λ^C by the formulas

$$\begin{aligned} E_R &= E_n - \det\alpha|D(E)|^{-2}\{|(k\lambda_1^C + 4i)(k\lambda_2^C + 4i)|^2\chi \\ &-2k^2[\lambda_2^A\lambda_2^C((k\lambda_1^C)^2 + 16)\alpha_{11} + \lambda_1^A\lambda_1^C((k\lambda_2^C)^2 + 16)\alpha_{22}]\}, \end{aligned} \tag{21}$$

$$\begin{aligned} \Gamma &= 8k\{\lambda_2^A[(k\lambda_1^C)^2 + 16]\alpha_{11} \\ &+ \lambda_1^A[(k\lambda_2^C)^2 + 16]\alpha_{22}\}\det\alpha|D(E)|^{-2}. \end{aligned} \tag{22}$$

These formulas show that the parameters of the Fano resonances are determined to a considerable degree by residues α_{ij} of functions $Q_{ij}(E)$; if $\det\alpha$ vanishes at a spectral point E_n for a certain arrangement of connecting leads, then the Fano resonance experiences a collapse in the neighborhood of this point. In this case, the pole $E_R + i\Gamma$ and the zero E_n of the transmission coefficient coincide and cancel out.

Using (14), we obtain the following expression:

$$\begin{aligned} \det\alpha &= -\frac{\hbar^4}{m^2}\sum_{l,l'}\overline{\varphi_l(\mathbf{q}_1)\varphi_{l'}(\mathbf{q}_2)} \\ &\times[\varphi_l(\mathbf{q}_1)\varphi_{l'}(\mathbf{q}_2) - \varphi_l(\mathbf{q}_2)\varphi_{l'}(\mathbf{q}_1)] = -\frac{\hbar^4}{2m^2}\sum_{l,l'}|M_{ll'}|^2, \end{aligned} \tag{23}$$

where $M_{ll'} = \varphi_l(\mathbf{q}_1)\varphi_{l'}(\mathbf{q}_2) - \varphi_l(\mathbf{q}_2)\varphi_{l'}(\mathbf{q}_1)$. It follows from (23) that $\det\alpha$ vanishes if and only if all $M_{ll'}$ are equal to zero, i.e., when the following relation holds:

$$\frac{\varphi_1(\mathbf{q}_1)}{\varphi_1(\mathbf{q}_2)} = \frac{\varphi_2(\mathbf{q}_1)}{\varphi_2(\mathbf{q}_2)} = \dots = \frac{\varphi_\nu(\mathbf{q}_1)}{\varphi_\nu(\mathbf{q}_2)}. \tag{24}$$

It is clear from (23) that, when E_n is a nondegenerate level, equality $\det\alpha = 0$ turns into an identity; hence, there are no Fano resonances in the neighborhoods of nondegenerate energy levels of operator H_S.

Now, let us consider the effect of the contacts on the transmission coefficient. Using the relation $|r_{11}|^2 + |t_{12}|^2 = 1$, we rewrite the transmission coefficient $T_{12}(E)$ as

$$T_{12} = \frac{|t_{12}|^2}{|r_{11}|^2 + |t_{12}|^2} = \frac{1}{1 + |r_{11}/t_{12}|^2}. \tag{25}$$

This expression is convenient because it allows one to analyze, instead of function $T_{12}(E)$, auxiliary function $f(E) = r_{11}(E)/t_{12}(E)$, which is expressed via $\tilde{Q}(E)$ in a simpler form. Using formulas (7) and (8), we obtain

$$|f(E)|^2 = \frac{f_1^2(E) + f_2^2(E)}{16k^2\lambda_1^A\lambda_2^A|\tilde{Q}_{21}(E)|^2}, \tag{26}$$

where $f_1(E)$ and $f_2(E)$ are real functions:

$$\begin{aligned} f_1(E) &= (k^2\lambda_1^C\lambda_2^C + 16)\det\tilde{Q}(E) \\ &- 2k^2[\lambda_1^C\lambda_2^A\tilde{Q}_{11}(E) + \lambda_2^C\lambda_1^A\tilde{Q}_{22}(E)] + 4k^2\lambda_1^A\lambda_2^A, \end{aligned} \tag{27}$$

$$\begin{aligned} f_2(E) &= 4k(\lambda_1^C - \lambda_2^C)\det\tilde{Q}(E) \\ &+ 8k[\lambda_2^A\tilde{Q}_{11}(E) - \lambda_1^A\tilde{Q}_{22}(E)]. \end{aligned} \tag{28}$$

Consider a practically important particular case of identical contacts in greater detail; this case corresponds to scalar matrices A, B, and C. In this case, we simply denote $\lambda_j^A \equiv \lambda^A$, $\lambda_j^B \equiv \lambda^B$, and $\lambda_j^C \equiv \lambda^C$. If $\tilde{Q}_{11}(E) \equiv \tilde{Q}_{22}(E)$, which takes place, for example, in the examples considered below, then function $f_2(E)$ identically vanishes, and the expression for $f(E)$ has an especially simple form:

$$\begin{aligned} f(E) &= \frac{1}{16ik\lambda^A\tilde{Q}_{21}(E)}\{[(k\lambda^C)^2 + 16]\det\tilde{Q}(E) \\ &- 2k^2\lambda^C\lambda^A[\tilde{Q}_{11}(E) + \tilde{Q}_{11}(E)] + 4(k\lambda^A)^2\}. \end{aligned} \tag{29}$$

As is clear from (28), if some of the equalities $\lambda_1^A = \lambda_2^A$, $\lambda_1^B = \lambda_2^B$, or $\lambda_1^C = \lambda_2^C$ are not satisfied, then identity $f_2(E) \equiv 0$ no longer holds. In this case, for the transmission coefficient to be strictly equal to one, it is necessary that two different functions $f_1(E)$ and $f_2(E)$ vanish simultaneously, which requires a special choice of parameters. Therefore, in the general case of nonidentical contacts, the maxima of curve $T_{12}(E)$ are less than one. Thus, asymmetry of the contacts results in a decrease in the amplitudes of the transmission peaks.

Now, let us consider the effect of scattering lengths λ^A and λ^C on the transmission coefficient. For brevity, we restrict the analysis to the case of identical contacts. As is clear from (29), as $k|\lambda^C| \longrightarrow \infty$, the transparency of the system tends to zero everywhere except small neighborhoods of the points at which $\det\tilde{Q}(E) = 0$. In this case, the term $(k\lambda^C)^2\det\tilde{Q}(E)$ is dominant in the numerator of (29); therefore, there always exists a value of energy in the neighborhood of the point $\det\tilde{Q}(E) = 0$ at which $f(E) = 0$ and the transmission coefficient is equal to unity. Thus, in the limit of large $|\lambda^C|$, function $T_{12}(E)$ represents a series of very narrow and high (reaching one) resonance peaks whose widths decrease as $|\lambda^C|$ increases. Note that, according to (3), the limit $|\lambda^C| \longrightarrow \infty$ corresponds to breaking the coupling between the nanostructure and the leads; naturally, the transparency of the system in this case tends to zero. Similar phenomena also occur when $\lambda^A \longrightarrow 0$; as is clear from boundary conditions (3), this situation also corresponds to the absence of coupling between the conducting leads and the nanostructure.

In the cases $\lambda^B \longrightarrow 0$ and $\lambda^B \longrightarrow \infty$, which correspond to the absence of point perturbations in a two-dimensional nanostructure, the transparency of the system also represents a series of narrow resonance peaks, beyond which the transmission coefficient tends to zero because the diagonal elements of the matrix $Q(E)$ in this case are large everywhere except small neighborhoods of points E_n, while the denominator of the transmission coefficient is large compared with the numerator.

4. RESONANCES IN A QUANTUM SPHERE AND A QUANTUM TORUS

As an application of the theory developed, consider two closed surfaces of different topologies, a quantum sphere and a quantum torus. Note that, in [25], the authors obtained formulas for the transmission coefficient of an electron passing through a quantum sphere with connecting leads; however, the structure of the

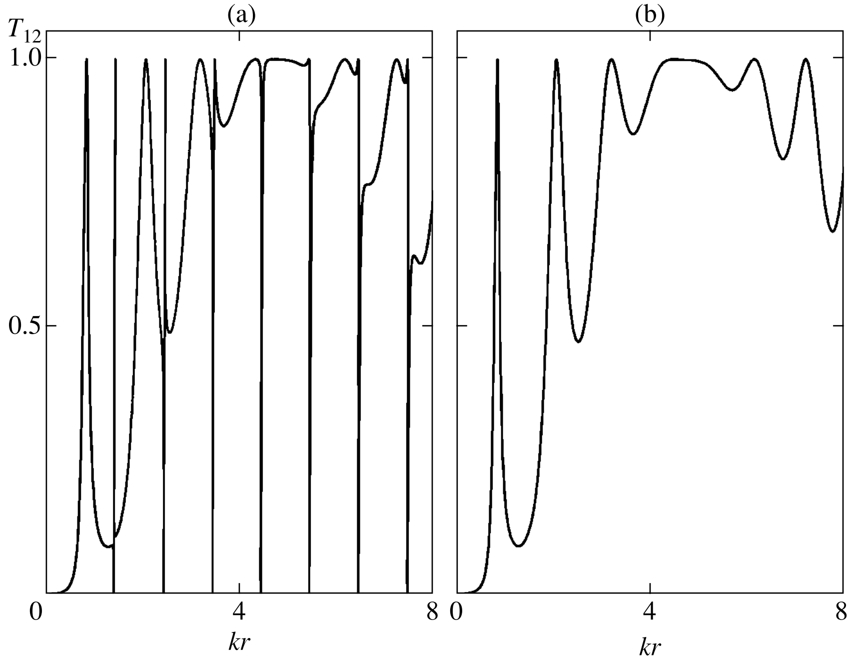

Fig. 2. Resonance structure of the transmission coefficient through a quantum sphere for $\lambda_j^A = \lambda_j^B = \lambda_j^C = 0.2r$; (a) $\rho(\mathbf{q}_2, \mathbf{q}_1) = 0.94\pi r$ and (b) $\rho(\mathbf{q}_2, \mathbf{q}_1) = \pi r$ (a collapse of Fano resonances).

Fano resonances was not studied in [35]. In the case of a sphere of radius r, Hamiltonian H_S has the form

$$H_S = \mathbf{L}^2/2mr^2,$$

where $\mathbf{L}$ is the operator of angular momentum. In this case, the Green function $G_S(\mathbf{x}, \mathbf{y}; E)$ can be expressed as [29]

$$G_S(\mathbf{x}, \mathbf{y}; E) = \frac{m}{2\hbar^2}\frac{1}{\cos(\pi t(E))} \times \mathcal{P}_{t(E)-1/2}\left(-\cos\left(\frac{\rho(\mathbf{x}, \mathbf{y})}{r}\right)\right), \tag{30}$$

where $\mathcal{P}_\nu(x)$ is the Legendre function and $t(E) = \sqrt{r^2k^2 + 1/4}$. Hence, using formula (5), we can easily derive the off-diagonal elements of matrix $Q(E)$. The diagonal elements of this matrix for the case of a sphere [25] are given by

$$Q_{11}(E) = Q_{22}(E) = -\frac{m}{\pi\hbar^2} \times \left[\Psi\left(\frac{1}{2} + t(E)\right) - \frac{\pi}{2}\tan(\pi t(E)) - \ln(2r) + C_\mathrm{E}\right], \tag{31}$$

where $\Psi(x)$ is a logarithmic derivative of the Γ function and C_E is the Euler constant.

Condition (24) for the collapse of Fano resonances in the neighborhood of $E_l = \hbar^2 l(l + 1)/2mr^2$ can be represented as

$$P_l^2\left(-\cos\left(\frac{\rho(\mathbf{q}_1, \mathbf{q}_2)}{r}\right)\right) = 1, \tag{32}$$

where $P_l(x)$ is a Legendre polynomial. This condition is fulfilled simultaneously for all l if $\rho(\mathbf{q}_1, \mathbf{q}_2) = \pi r$ because $P_l(1) \equiv 1$. Thus, when the conducting leads are connected diametrically opposite to the sphere, all Fano resonances experience a collapse. In this case, function $Q_{12}(E)$ does not vanish, zeros of both types on the curve of $T_{12}(E)$ disappear, and the transport regime changes qualitatively (Fig. 2). Note that, when $|x| < 1$, the inequality $|P_l(x)| < 1$ holds for any l [30]; therefore, a collapse of the Fano resonances on the sphere occurs only for a unique configuration of the contacts.

Now, let us consider a quantum torus. The motion of an electron on a torus is described by the Hamiltonian of a two-dimensional free motion with periodic boundary conditions along two directions. To describe these conditions, it is convenient to introduce a rectangular lattice Λ in $\mathbf{R}^2$ with the basis vectors $\mathbf{a}_1$ and $\mathbf{a}_2$:

$$\Lambda = \{n_1\mathbf{a}_1 + n_2\mathbf{a}_2 : n_j \in \mathbf{Z}, j = 1, 2\}.$$

For definiteness, assume that $|\mathbf{a}_1| > |\mathbf{a}_2|$. Let us require that all the functions ψ_S from the domain of definition

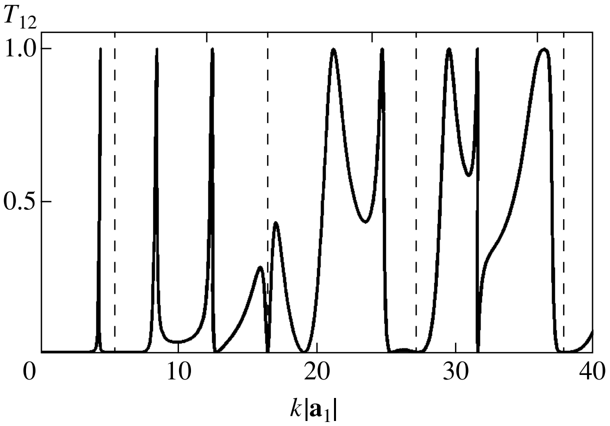

Fig. 3. Transmission coefficient of an electron passing through a quantum torus as a function of the dimensionless parameter $k|\mathbf{a}_1|$ for $|\mathbf{a}_2| = 0.01|\mathbf{a}_1|$, $\mathbf{q}_2 - \mathbf{q}_1 = 0.21\mathbf{a}_1$, and $\lambda_j^A = \lambda_j^B = \lambda_j^C = 0.1|\mathbf{a}_1|$. The dashed lines point to the positions of the second-type zeros.

of the Hamiltonian of the torus satisfy the relation

$$\psi_S(\mathbf{x}+\mathbf{a}) = \psi_S(\mathbf{x}), \quad \forall \mathbf{a} \in \Lambda. \tag{33}$$

Then, the Hamiltonian of the torus represents the operator

$$H_S = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}\right) \tag{34}$$

that acts on the functions satisfying condition (33).

Let us define the explicit form of the function $Q_{ij}(E)$ for a quantum torus. Denote by S the area of a unit cell of Λ and introduce Ω, a lattice dual to Λ, i.e., a lattice with the basis vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ that satisfy the conditions $\mathbf{a}_j\mathbf{b}_k = 2\pi\delta_{jk}$. In this case, the eigenvalues of the operator H_S can be represented as $E_{\mathbf{b}} = \hbar^2\mathbf{b}^2/2m$. The corresponding eigenfunction is

$$\varphi_{\mathbf{b}}(\mathbf{x}) = \frac{1}{S}e^{i\mathbf{b}\cdot\mathbf{x}}.$$

It is clear that all the states of the unperturbed Hamiltonian of the torus, except for the ground state, are degenerate because the wave functions defined by the vectors $\mathbf{b}$ and $-\mathbf{b}$ correspond to the same energy; moreover, in the general case, all the states defined by vectors $\mathbf{b} = n_1\mathbf{b}_1 + n_2\mathbf{b}_2$ with nonzero n_1 and n_2 are fourfold degenerate. If ratio $|\mathbf{a}_1|/|\mathbf{a}_2|$ of the generators of the torus is rational, then the energy levels of Hamiltonian H_S may be more than fourfold degenerate.

Using (5), we can represent matrix $Q_{ij}(E)$ for the torus as [31]

$$\left(Q_{ij}(E) = \frac{2m}{S\hbar^2}\right) \times \lim_{\substack{\omega\to\infty \\ |\mathbf{b}|\le\omega}}\left[\frac{\exp[i\mathbf{b}\cdot(\mathbf{q}_i - \mathbf{q}_j)]}{|\mathbf{b}|^2 - k^2} + \frac{S}{2\pi}\delta_{ij}\ln\omega\right]. \tag{35}$$

Applying the Poisson summation formula, we can rep-

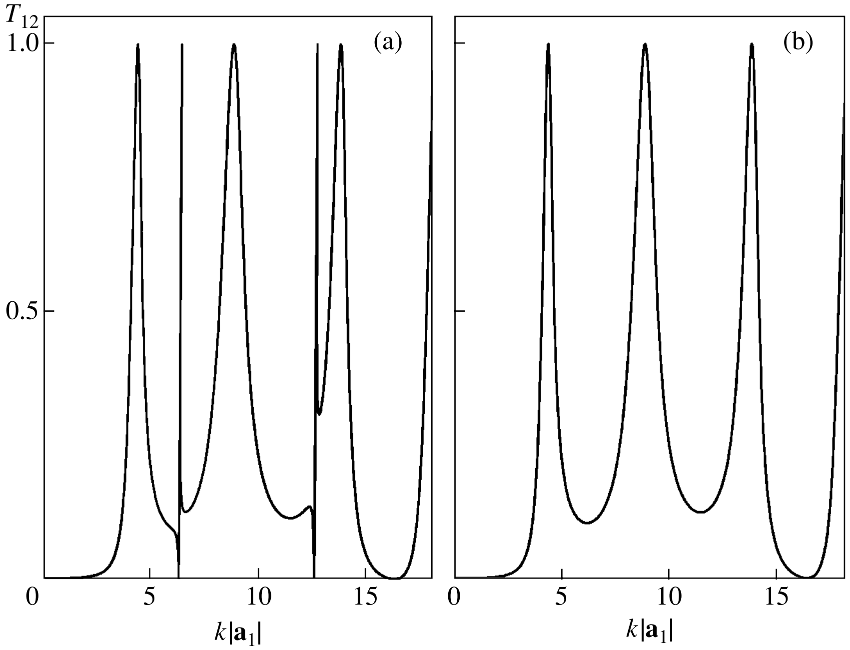

Fig. 4. Resonance structure of the transmission coefficient through a quantum torus for $|\mathbf{a}_2| = |\mathbf{a}_1|/e$ and $\lambda_j^A = \lambda_j^B = \lambda_j^C = 0.02|\mathbf{a}_1|$; (a) $\mathbf{q}_2 - \mathbf{q}_1 = 0.48\mathbf{a}_1$ and (b) $\mathbf{q}_2 - \mathbf{q}_1 = 0.5\mathbf{a}_1$.

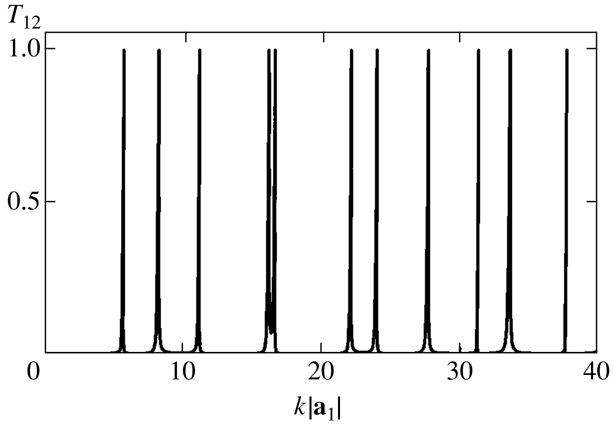

Fig. 5. Resonance structure of the transmission coefficient through a quantum torus for relatively large values of parameter $\lambda^C = |\mathbf{a}_1|$ and $\mathbf{q}_2 - \mathbf{q}_1 = 0.41\mathbf{a}_1 + 0.31\mathbf{a}_2$. Other parameters are the same as in Fig. 3.

resent $Q_{ij}(E)$ as an absolutely converging series

$$Q_{ij}(E) = \frac{2m}{\hbar^2}\left\{\frac{S^{-1}+k^2}{S}\right. \tag{36}$$
$$\left.\times \sum_{\mathbf{b}\in\Omega}\frac{\exp[i\mathbf{b}\cdot(\mathbf{q}_i-\mathbf{q}_j)]}{(|\mathbf{b}|^2-k^2)(|\mathbf{b}|^2+S^{-1})} + \kappa(\mathbf{q}_i-\mathbf{q}_j)\right\},$$

where

$$\kappa(\mathbf{x}) = \begin{cases} \dfrac{1}{2\pi}\sum\limits_{\mathbf{a}\in\Lambda} K_0\left(\dfrac{|\mathbf{x}+\mathbf{a}|}{\sqrt{S}}\right), & \mathbf{x}\notin\Lambda \\ \dfrac{1}{2\pi}\left[\sum\limits_{\substack{\mathbf{a}\in\Lambda\\ \mathbf{a}\neq 0}} K_0\left(\dfrac{|\mathbf{a}|}{\sqrt{S}}\right) + \ln 2 - C_E\right], & \mathbf{x}\in\Lambda. \end{cases}$$

Here, $K_0(x)$ is a Macdonald function. The transmission coefficient of an electron through a quantum torus versus the wave number k is displayed in Fig. 3 for a general position of contacts. One can see that the curve contains Fano and Breit–Wigner resonances, as well as zeros that are not associated with the resonances.

In the case of a torus, a state of the system is uniquely defined by the reciprocal-lattice vector $\mathbf{b}$ and collapse condition (24) for the Fano resonances can be written as

$$\begin{aligned}\exp[i\mathbf{b}_1\cdot(\mathbf{q}_2-\mathbf{q}_1)] &= \exp[i\mathbf{b}_2\cdot(\mathbf{q}_2-\mathbf{q}_1)] \\ &= \dots = \exp[i\mathbf{b}_\nu\cdot(\mathbf{q}_2-\mathbf{q}_1)],\end{aligned} \tag{37}$$

where $\mathbf{b}_i$ are vectors corresponding to the sates of a ν-fold degenerate level E_n: $|\mathbf{b}_i|^2 = 2mE_n/\hbar^2$. For twofold degenerate levels, this condition has the simplest form, $\mathbf{b}\cdot(\mathbf{q}_2-\mathbf{q}_1) = \pi n$, $n\in\mathbf{Z}$. One can easily verify that, when $\mathbf{q}_2-\mathbf{q}_1 = 0.5\mathbf{a}_1$, condition (37) holds for all twofold and fourfold degenerate levels. Therefore, when the generators of the torus are incommensurate for $\mathbf{q}_2 - \mathbf{q}_2 = 0.5\mathbf{a}_1$, all Fano resonances experience a collapse; however, the second-type zeros related to $Q_{12}(E)$ are conserved (Fig. 4).

Note that, in both examples considered above, equality $Q_{11}(E) = Q_{22}(E)$ is fulfilled for any arrangement of points $\mathbf{q}_1$ and $\mathbf{q}_2$; therefore, in the case of identical contacts, the resonance peaks may reach one.

5. CONCLUSIONS

An explicit formula has been obtained for the transmission coefficient of an electron passing through a two-dimensional nanostructure **S** placed between two one-dimensional conducting leads. We have shown that, in the general case, the system exhibits two types of resonances, Breit–Wigner and Fano resonances. We have found a collapse condition (24) for the Fano resonances under which the resonance widths become zero and discrete energy levels embedded into a continuous spectrum arise. We have established that, for Fano resonances to occur in the system, the energy levels of the original Hamiltonian H_S should be degenerate.

In contrast to [5–7], the method developed in this paper for deriving the scattering matrix parameters makes it possible to take into account explicitly the effect of the geometry of the nanostructure on the electron transport in a device. The geometric features of the nanodevice lead to a difference in the behavior of the resonances and zeros of function $T_{12}(E)$. In particular, the symmetry of the sphere leads to a collapse of all Fano resonances and the vanishing of all zeros for a diametrically opposite configuration of connecting leads; in other cases, a collapse is not observed. In contrast to a sphere, a collapse on a torus occurs under condition (37), which is satisfied for many configurations of contacts; however, a part of the zeros on the graph of $T_{12}(E)$ is always conserved in this case.

Consider the case when one of the generators of a torus is much greater than the other, $|\mathbf{a}_1| \gg |\mathbf{a}_2|$; this situation is close to the geometry of a nanotube coiled into a torus. In this case, for energies lower than $\hbar^2\mathbf{b}_2^2/2m$, the main contribution to $Q_{ij}(E)$ is made by low-energy states defined by vectors $\mathbf{b} = n\mathbf{b}_1$ with $n < |\mathbf{a}_1|/|\mathbf{a}_2|$. Therefore, for the values of energy considered here, function $Q_{ij}(E)$ for a torus approximates a similar function for a ring up to a dimensional factor. Such a situation is associated with the fact that the electron modes corresponding to the motion along the smaller generator of the torus are not excited when $E < \hbar^2\mathbf{b}_2^2/2m$. The results obtained in this limit are in good agreement with the results obtained for a one-dimensional ring [4].

In this work, we have considered the effect of the parameters of contacts and demonstrated that, in the limiting cases corresponding to a weak coupling between connecting leads and a nanostructure **S**, function $T_{12}(E)$ represents a series of narrow resonance peaks between which the transmission coefficient is close to zero (Fig. 5). As the coupling becomes weaker, the widths of the resonance peaks decrease, although their amplitudes remain unchanged provided that the contacts are identical. It is important that the positions of the zeros of the transmission coefficient in the system under investigation are independent of the scattering lengths and are determined only by the arrangement of contacts and the geometry of nanostructure **S**.

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project nos. 01-02-16564 and 02-01-00804), by INTAS (grant no. 00-257), and by DFG.

REFERENCES

1. C.-M. Ryu and S. Y. Cho, Phys. Rev. B **58**, 3572 (1998).
2. Q. Sun, J. Wang, and T. Lin, Phys. Rev. B **60**, R13981 (1999).
3. N. T. Bagraev, A. D. Buravlev, V. K. Ivanov, *et al.*, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 846 (2000) [Semiconductors **34**, 817 (2000)].
4. V. A. Geĭler, V. V. Demidov, and V. A. Margulis, Zh. Tekh. Fiz. **73** (6), 1 (2003) [Tech. Phys. **48**, 661 (2003)].
5. A. A. Clerk, X. Waintal, and P. W. Brouwer, Phys. Rev. Lett. **86**, 4636 (2001).
6. B. R. Bulka and P. Stefanski, Phys. Rev. Lett. **86**, 5128 (2001).
7. M. E. Torio, K. Hallberg, A. H. Ceccatto, *et al.*, Phys. Rev. B **65**, 085302 (2002).
8. O. A. Tkachenko, V. A. Tkachenko, D. G. Baksheev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 366 (2000) [JETP Lett. **71**, 255 (2000)].
9. A. A. Bykov, D. G. Baksheev, L. V. Litvin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 631 (2000) [JETP Lett. **71**, 434 (2000)].
10. A. A. Bykov, D. G. Bakarov, L. V. Litvin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 300 (2000) [JETP Lett. **72**, 209 (2000)].
11. Ch. S. Kim and A. M. Satanin, Zh. Éksp. Teor. Fiz. **115**, 211 (1999) [JETP **88**, 118 (1999)].
12. Ch. S. Kim, A. M. Satanin, Yong S. Joe, *et al.*, Zh. Éksp. Teor. Fiz. **116**, 263 (1999) [JETP **89**, 144 (1999)].
13. C. S. Kim and A. M. Satanin, Physica E (Amsterdam) **4**, 211 (1999).
14. Ch. S. Kim, A. M. Satanin, and V. B. Shtenberg, Zh. Éksp. Teor. Fiz. **118**, 413 (2000) [JETP **91**, 361 (2000)].
15. Ch. S. Kim, O. N. Roznova, A. M. Satanin, *et al.*, Zh. Éksp. Teor. Fiz. **121**, 1157 (2002) [JETP **94**, 992 (2002)].
16. V. A. Geĭler, V. A. Margulis, and L. I. Filina, Zh. Éksp. Teor. Fiz. **113**, 1376 (1998) [JETP **86**, 751 (1998)].
17. N. G. Galkin, V. A. Geĭler, and V. A. Margulis, Zh. Éksp. Teor. Fiz. **118**, 223 (2000) [JETP **91**, 197 (2000)].
18. J. Göres, D. Goldhaber-Gordon, S. Heemeyer, *et al.*, Phys. Rev. B **62**, 2188 (2000).
19. I. G. Zacharia, D. Goldhaber-Gordon, G. Granger, *et al.*, Phys. Rev. B **64**, 155311 (2001).
20. K. Kobayashi, H. Aikawa, S. Katsumoto, *et al.*, Phys. Rev. Lett. **88**, 256806 (2002).
21. C. L. Foden, M. L. Leadbeater, and M. Pepper, Phys. Rev. B **52**, R8646 (1995).
22. L. I. Magarill, D. A. Romanov, and A. V. Chaplik, Zh. Éksp. Teor. Fiz. **113**, 1411 (1998) [JETP **86**, 771 (1998)].
23. A. Kiselev, J. Math. Anal. Appl. **212**, 263 (1997).
24. B. S. Pavlov, I. Yu. Popov, V. A. Geyler, *et al.*, Europhys. Lett. **52**, 196 (2000).
25. J. Brüning, V. A. Geyler, V. A. Margulis, *et al.*, J. Phys. A **35**, 4239 (2002).
26. J. Brüning and V. A. Geyler, J. Math. Phys. **44**, 371 (2003).
27. Yu. N. Demkov and V. N. Ostrovskiĭ, *Zero-Range Potentials and Their Applications in Atomic Physics* (Leningr. Gos. Univ., Leningrad, 1975; Plenum, New York, 1988).
28. L. Dabrowski and H. Grosse, J. Math. Phys. **26**, 2777 (1985).
29. C. Grosche and F. Steiner, *Handbook of Feynman Path Integrals* (Springer, Berlin, 1998).
30. G. Szego, *Orthogonal Polynomials* (Am. Math. Soc., New York, 1959; Fizmatgiz, Moscow, 1962).
31. Yu. E. Karpeshina, Teor. Mat. Fiz. **57**, 304 (1983).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 773–780.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 862–870.
Original Russian Text Copyright © 2003 by Borisov, Gavrichkov, Ovchinnikov.

SOLIDS
Electronic Properties

Temperature and Concentration Dependences of the Electronic Structure of Copper Oxides in the Generalized Tight Binding Method

A. A. Borisov*, V. A. Gavrichkov, and S. G. Ovchinnikov
Kirenskiĭ Institute of Physics, Siberian Division, Russian Academy of Sciences,
Akademgorodok, Krasnoyarsk, 660036 Russia
**e-mail: alex@sme.krk.ru*

Received December 5, 2002

Abstract—The electronic structure of *p*-type doped HTSC cuprates is calculated by explicitly taking into account strong electron correlations. The smooth evolution of the electronic structure from undoped antiferromagnetic to optimally and heavily doped paramagnetic compositions is traced. For a low doping level, in-gap impurity-type states are obtained, at which the Fermi level is pinned in the low-doping region. These states are separated by a pseudogap from the valence band. The Fermi surfaces calculated for the paramagnetic phase for various concentrations of holes are in good agreement with the results of ARPES experiments and indicate a gradual change in the Fermi surface from the hole type to the electron type. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Knowledge of the low-energy electronic structure, i.e., the Fermi surface and the dispersion of bands both in the superconducting and in the normal state, is essential for a better understanding of the properties of high-temperature superconductors (layered cuprates). This is necessary to determine the superconductivity mechanism and to interpret the thermodynamic and transport properties. In addition, this information can be directly obtained using angular-resolution photoelectron spectroscopy (ARPES) for the compounds $Bi_2Sr_2CaCu_2O_{8+y}$ (Bi2212) [1–9], $Bi_2Sr_2CuO_{6+y}$ (Bi2201) [10, 11], and $YBa_2Cu_3O_{7-y}$ (YBCO) [12]. The most comprehensive ARPES studies have been carried out for $La_{2-x}S_xCO_4$ (LSCO) and Bi2212 in the HTSC state. Weakly doped LSCO exhibits in-gap states and the absence of a shift (pinning) of the Fermi level. As the doping level increases to above the optimal value, the Fermi surface changes from the hole type to the electron type [13]. Unfortunately, the data on weakly doped Bi2212 are scarce in view of the absence of stable materials. However, all other properties of Bi2212 (in particular, the *d* symmetry and the concentration dependence of the pseudogap and the evolution of the Fermi surface (FS) from the hole type to the electron type in the region of strong doping [13]) are similar to those observed for LSCO. Since the electron properties of high-temperature cuprates strongly depend on the doping level, it is essential to trace the evolution of the band structure of the ARPES spectra upon doping for an understanding of the main features of high-temperature superconductivity.

The conventional band theory successfully reproduces the shape of the Fermi surface in optimally doped compounds [14] such as Bi2212. However, problems exist in HTSC systems, which cannot be solved in the framework of the band theory. Indeed, in accordance with the observed phase diagram, La_2CuO_4 is an antiferromagnetic (AFM) insulator. The dielectric nature of the ground state is ensured by a strong electron–electron Coulomb interaction at a site. However, in accordance with the Wilson criterion, the band theory demonstrates that La_2CuO_4 is a paramagnetic (PM) metal, and there are no "in-gap" states of the corresponding pinning of the Fermi level in the region of weakly doped compositions. Thus, it can be concluded that the advances in the band theory in the local density functional approximation are limited to optimally and strongly doped HTSC compounds. For weakly doped and undoped compounds, it is necessary to calculate the electronic structure taking into account strong electron correlations. Earlier, we proposed a generalized tight binding method (GTBM) [15] combining the exact diagonalization of the multielectron Hamiltonian in a cell with perturbation theory for hopping between unit cells. For an undoped CuO_2 layer, the GTBM calculations correctly reproduced not only the width of the dielectric gap, but also the dispersion at the top of the valence band for $Sr_2CuO_2Cl_2$ [16].

In this study, we analyze the spectral density of states at the top of the valence band as well as the FS cross sections for doped cuprates for various hole concentrations. The results of numerical calculations and their analysis and comparison with experiments will be given below.

In Section 2, the GTBM will be described briefly and basic relations for dispersion and spectral density will be given. In Section 3, we describe the numerical calculations for the density of states and the dependence of the position of the chemical potential for various concentrations of the doping component both for the PM and for the AFM phase using only the parameters determined from the calculation of the dielectric state [16]. In Section 4, the cross sections of the Fermi surface and their dependence on the doping levels are considered and compared with experimental results. In Section 5, we will show that the nontrivial temperature dependence of the spectral density in cuprates can also be reproduced in computations. The results are compared with the ARPES data obtained for chlorides $Sr_2CuO_2Cl_2$ and $Ca_2CuO_2Cl_2$.

2. BRIEF DESCRIPTION OF THE GENERALIZED TIGHT BINDING METHOD

The initial Hamiltonian of the multiband *p–d* model can be written in the form [17]

$$H = H_d + H_p + H_{pd} + H_{pp}, \quad H_d = \sum_r H_d(r),$$

$$H_d(r) = \sum_{\lambda\sigma}\Bigg[(\varepsilon_\lambda - \mu)\overset{+}{d}_{\lambda r\sigma} d_{\lambda r\sigma} + \frac{1}{2}U_\lambda n^{\alpha}_{\lambda r} n^{-\sigma}_{\lambda r}$$

$$+ \sum_{\lambda'\sigma'}\Bigg(- J_d \overset{+}{d}_{\lambda r\sigma} d_{\lambda r\sigma'} \overset{+}{d}_{\lambda' r\sigma'} d_{\lambda' r\sigma} + \sum_{r'} V_{\lambda\lambda'} n^{\sigma}_{\lambda r} n^{\sigma'}_{\lambda' r'}\Bigg)\Bigg],$$

$$H_p = \sum_i H_p(i),$$

$$H_p(i) = \sum_{\alpha\sigma}\Bigg[(\varepsilon_\alpha - \mu)\overset{+}{p}_{\alpha i\sigma} p_{\alpha r\sigma} + \frac{1}{2}U_\alpha n^{\sigma}_{\alpha i} n^{-\sigma}_{\alpha i} \tag{1}$$

$$+ \sum_{\alpha'\sigma'}\Bigg(- J_p \overset{+}{p}_{\alpha i\sigma} p_{\alpha i\sigma'} \overset{+}{p}_{\alpha' i\sigma'} p_{\alpha' i\sigma} + \sum_{i'} V_{\alpha\alpha'} n^{\sigma}_{\alpha i} n^{\sigma'}_{\alpha' i'}\Bigg)\Bigg],$$

$$H_{pd} = \sum_{\langle i, r\rangle} H_{pd}(i, r),$$

$$H_{pd}(i, r) = \sum_{\alpha\lambda\sigma\sigma'} (t_{\lambda\alpha} \overset{+}{p}_{\alpha i\sigma} d_{r\lambda\sigma} + V_{\alpha\lambda} n^{\sigma}_{\alpha i} n^{\sigma'}_{\lambda r}),$$

$$H_{pp} = \sum_{\langle i, r\rangle} \sum_{\alpha\beta\sigma} (t_{\alpha\beta} \overset{+}{p}_{\alpha i\sigma} \beta_{\beta j\sigma} + \text{H.c.}),$$

where $n^{\sigma}_{\lambda i} = \overset{+}{d}_{\lambda i\sigma} d_{\lambda i\sigma}$ and $n^{\sigma}_{\alpha i} = \overset{+}{p}_{\alpha i\sigma} p_{\alpha i\sigma}$. Indices *r* and *i* run through the positions of copper ($d_{x^2-y^2} \equiv d_x$, $d_{3z^2-r^2} \equiv d_z$) and oxygen (p_x, p_y, p_z) sets of localized atomic orbitals. Similarly, $\varepsilon_\lambda = \varepsilon_{d_x}$ ($\lambda = d_x$), $\varepsilon_\lambda = \varepsilon_{d_x}$ ($\lambda = d_z$) and $\varepsilon_\alpha = \varepsilon_p$ ($\alpha = p_x, p_y$), $\varepsilon_\alpha = \varepsilon_{p_z}$ ($\alpha = p_z$) are the energies of the corresponding atomic orbitals; $t_{\lambda\alpha} = t_{pd}$ ($\lambda = d_x$; $\alpha = p_x, p_y$), $t_{\lambda\alpha} = t_{pd}/\sqrt{3}$ ($\lambda = d_z$, $\alpha = p_x, p_y$) are the matrix elements of the copper–oxygen hopping; $t_{\alpha\beta} = t_{pp}$ are the matrix elements of the hopping between the nearest oxygen ions; $U_\lambda = U_d$ ($\lambda = d_x, d_z$) and $U_\alpha = U_p$ ($\alpha = p_x, p_y, p_z$) are the intraatomic Coulomb interactions at copper and oxygen; $V_{dd}(V_{pp})$ and $J_{dd}(J_{pp})$ are the energies of the intraatomic Coulomb and exchange interactions of copper (oxygen) electrons on different orbitals; and $V_{\alpha\lambda} = V_{pd}$ ($\alpha = p_x, p_y$; $\lambda = d_x, d_z$) and $V_{\alpha\lambda} = V'_{pd}$ ($\alpha = p_z$, $\lambda = d_x, d_z$) are the energies of the Coulomb repulsion between copper and oxygen. All matrix elements of the Coulomb and exchange interactions are assumed to be independent of the form of the *d*- and *p*-plane orbitals. The prime corresponds to the interaction with apical oxygen in the CuO_6 cluster.

All calculations were made using the GTBM [16] for the CuO_2 plane divided into unit CuO_6 cells. In this version, the cell symmetry coincides with the symmetry of the crystal, but there arises the problem of common oxygen belonging to two cells simultaneously.

The problem of nonorthogonality of molecular orbitals of neighboring clusters was solved explicitly by constructing the corresponding Vanier functions on the $d_{x^2-y^2}$, $d_{3z^2-r^2}$, p_x, p_y, p_z five-orbital initial basis of atomic states. In the new symmetric basis, the one-cell part of the Hamiltonian is factorized, permitting the classification of all possible one-particle excitations in the CuO_2 plane according to symmetry. The subsequent exact diagonalization of the unit cell Hamiltonian and a transition to the representation of the Hubbard operators of the unit cell make it possible to take into account the intercellular-hopping part of the Hamiltonian in perturbation theory. As a result of exact diagonalization, the cell Hamiltonian H_c for the antiferromagnetic phase assumes the form

$$H_c = \sum_{p f_G \sigma} (\varepsilon_{1pG} - \mu) X^{pp}_{f_G\sigma} + \sum_{d f_G \sigma} (\varepsilon_{2qG} - 2\mu) X^{qq}_{f_G\sigma},$$
$$f_G = \begin{cases} f_A, & f \in A \\ f_B, & f \in B. \end{cases} \tag{2}$$

Here, *p* and *q* denote one-hole and two-hole terms of the cell and $X^{pq}_f = |p\rangle\langle q|$ are the Hubbard operators constructed on exact states of the unit cell. The energy levels of sublattices are split by the molecular field of the antiferromagnetic state: $\varepsilon_{1pA} = \varepsilon_{1p} - \sigma h$ and $\varepsilon_{1pB} = \varepsilon_{1p} - \sigma h$. The quantity $h \propto J\langle S_z\rangle$, where *J* is the effective exchange interaction between nearest neighbors. With increasing doping level, the value of h decreases, vanishing in the paramagnetic phase. In this study, we confine the analysis to non-self-consistent calculation in which the magnetic (antiferromagnetic or paramagnetic) state is assumed to be preset.

The Hamiltonian of hopping between unit cells can be written in matrix form:

$$H_{cc} = \begin{pmatrix} H_{AA} & H_{AB} \\ H_{BA} & H_{BB} \end{pmatrix} = \sum_{\lambda\lambda'\sigma}\sum_{kmn} \gamma^{*}_{\lambda\sigma}(m)\gamma_{\lambda'\sigma}(n)$$

$$\times \begin{pmatrix} T^{AA}_{\lambda\lambda'}(k) X^{+m}_{k\sigma} X^{n}_{k\sigma} & T^{AB}_{\lambda\lambda'}(k) X^{+m}_{k\sigma} Y^{n}_{k\sigma} \\ T^{BA}_{\lambda\lambda'}(k) Y^{+m}_{k\sigma} X^{n}_{k\sigma} & T^{BB}_{\lambda\lambda'}(k) Y^{+m}_{k\sigma} Y^{n}_{k\sigma} \end{pmatrix} + \text{H.c.},$$

$$T^{AA}_{\lambda\lambda'}(k) = T^{BB}_{\lambda\lambda'}(k) = \frac{2}{N}\sum_{R_1} T^{AA}_{\lambda\lambda'}(\mathbf{R}_1) e^{ikR_1},$$

$$T^{AB}_{\lambda\lambda'}(k) = T^{BA}_{\lambda\lambda'}(k) = \frac{2}{N}\sum_{R_2} T^{AB}_{\lambda\lambda'}(\mathbf{R}_2) e^{ikR_2}$$

($X^{m}_{k\sigma}$ and $Y^{n}_{k\sigma}$ are the Fourier transforms of the Hubbard operators over the A and B sublattices, respectively; here, index m labels the quasiparticle band and is determined by the pair (p, q) of indices of the initial and final multielectron states $|p\rangle$ and $|q\rangle$ in Hubbard operator X^{pq} (in terms of [18], m denotes the root vector $\boldsymbol{\alpha}_m = \boldsymbol{\alpha}(p, q)$).

The corresponding dispersion relations for the band structure of quasiparticles were derived using the equations of motion for the two-time Green temperature functions constructed on the Hubbard operators,

$$\begin{aligned} G^{\lambda\lambda'}_{k\sigma} &= \langle\langle c_{k\lambda\sigma} | \overset{+}{c}_{k\lambda'\sigma} \rangle\rangle_E \\ &= \sum_{mn} \gamma_{\lambda\sigma}(m) \overset{+}{\gamma}_{\lambda'\sigma}(n) D^{mn}_{k\sigma}, \end{aligned} \tag{3}$$

where

$$\hat{D}_{k\sigma} = \begin{pmatrix} \hat{D}_{k\sigma}(AA) & \hat{D}_{k\sigma}(AB) \\ \hat{D}_{k\sigma}(BA) & \hat{D}_{k\sigma}(BB) \end{pmatrix},$$

$$D^{mn}_{k\sigma}(AB) = \langle\langle X^{m}_{k\sigma} | Y^{n}_{k\sigma} \rangle\rangle_E.$$

In the long run, in the Hubbard I approximation, the dispersion relations are defined by the equation

$$\left\| \frac{(E - \Omega^{G}_{m})\delta_{mn}}{F^{G}_{\sigma}(m)} - 2\sum_{\lambda\lambda'} \gamma^{*}_{\lambda\sigma}(m) T^{PG}_{\lambda\lambda'}(\mathbf{k}) \gamma_{\lambda'\sigma}(n) \right\| = 0. \tag{4}$$

This equation is an analog of the ordinary one-electron equation in the tight binding method and differs from it in the following two aspects. First, the local energy levels $\boldsymbol{\Omega}_m$ are defined as resonances between multielectron states and, hence, take into account explicitly the strong correlations due to exact diagonalization of the Hamiltonian of the CuO$_6$ cluster. Second, the filling factors $F^{G}_{\sigma}(m) = \langle X^{pp}_{f_G\sigma} \rangle + \langle X^{qq}_{f_G\sigma} \rangle$, which are calculated self-consistently, lead to a doping dependence of both the dispersion relations and the amplitude of the quasiparticle peak in the spectral density. From the mathematical point of view, we are dealing with a generalized eigenvalue problem, in which the inverse matrix of the corresponding filling factors appears instead of the conventional "nonorthogonality matrix." Each root vector $\boldsymbol{\alpha}_m$ describes a Fermi quasiparticle with charge e and a spin of 1/2, but with a fractional spectral weight of the order of $F(m)$; their local energies are equal to $\Omega^{G}_{m} = \varepsilon_{2qG} - \varepsilon_{1pG}$.

Formula (4) is convenient for calculating the dispersion relation in the sense that it enables us to obtain all possible quasiparticle states. However, not all of these states can be observed in experiments. It is well known that the spectrum in ARPES experiments is proportional to the spectral density of quasiparticles, which can be calculated by the GTBM in the form

$$\begin{aligned} A_{\sigma}(\mathbf{k}, E) &= -\frac{1}{\pi}\sum_{\lambda} \mathrm{Im}(G^{\lambda\lambda}_{k\sigma}) \\ &= -\frac{1}{\pi}\sum_{\lambda mn} \gamma_{\lambda\sigma}(m) \overset{+}{\gamma}_{\lambda\sigma}(n) \mathrm{Im}(D^{mn}_{k\sigma}(AA) + D^{mn}_{k\sigma}(BB)). \end{aligned} \tag{5}$$

Owing to corresponding filling factors, the spectral density for some types of quasiparticles may be simply negligibly low or even equal to zero. Consequently, the corresponding quasiparticle peak can be missing in the experiment. The numerical calculation of spectral density by formula (5) was carried out along the principal symmetric directions of the Brillouin zone at $T = 0$. For the PM phase, the dispersion relation and the spectral density can be obtained using one-sublattice analogs of formulas (4) and (5). The band structure of the undoped CuO$_2$ layer in the PM phase was calculated in [19] and is shown in Fig. 1. In this case, the Fermi level lies in the forbidden band and the dielectric gap width $E_g \sim 0.2$ eV is close to the experimentally observed value. It should be noted that the dielectric ground state is due to strong electron correlations, while the gap width is mainly determined by charge transfer processes. The complex structure of the valence band is due to a large number of quasiparticles with the participation of two-

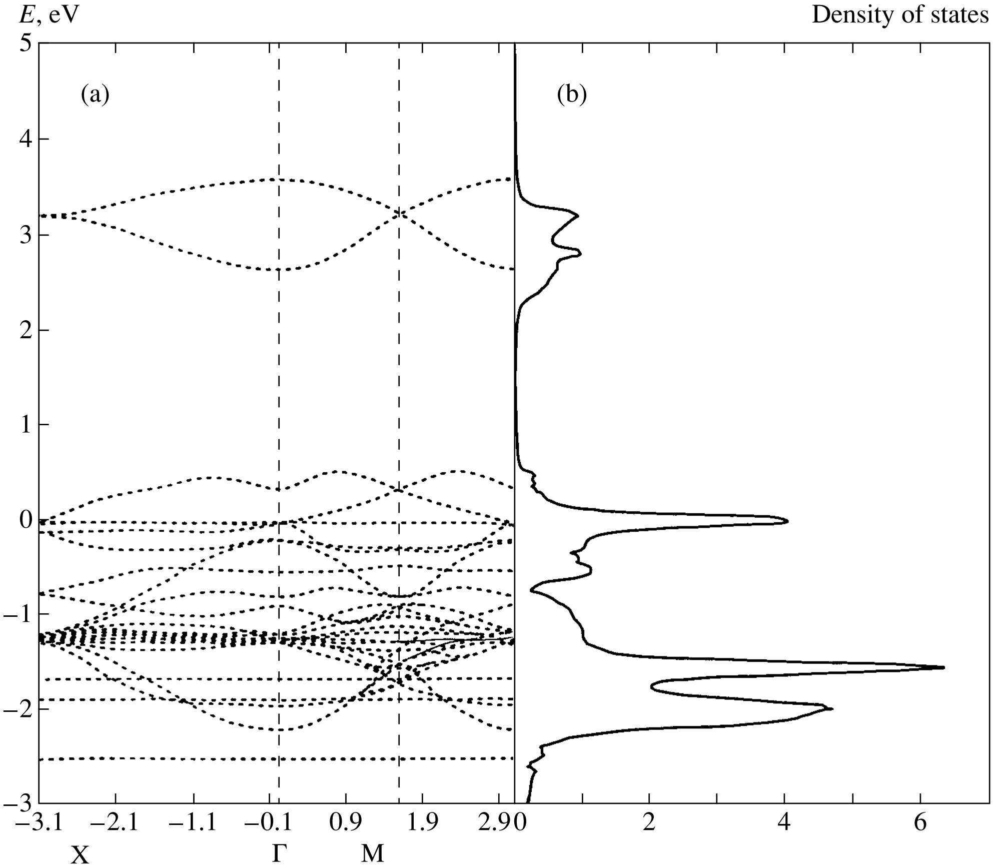

Fig. 1. Dispersion relation and the density of states of an undoped CuO_2 layer in the PM phase (borrowed from [19]).

hole states of the unit cell; the width of the valence band is approximately equal to 3 eV in conformity with the experimentally observed value for La_2CuO_4.

Following publication [19], numerous experimental studies of the band structure of cuprates were carried out using ARPES methods, which enabled us to refine the model parameters [16] and to achieve a quantitative agreement for the dispersion relation for the top of the valence band in $Sr_2CuO_2Cl_2$. We obtained the following model parameters in units of t_{pd} ($\varepsilon_{d_x} = 0$):

$$\begin{gathered}\varepsilon_{d_z} = 2, \quad \varepsilon_p = 1.6, \quad \varepsilon_{p_z} = 0.5, \\ t_{pp} = 0.46, \quad t'_{pp} = 0.42, \quad U_d = 9, \\ U_p = 4, \quad U_{pd} = 1.5, \quad J_d = 1.\end{gathered} \tag{6}$$

Here, we describe the doping dependence of the electronic structure without introducing additional fitting parameters.

3. DENSITY OF STATES

Figure 2 shows the dispersion relations for the bottom of the conduction band and the top of the valence band of an undoped CuO_2 layer, calculated for the AFM phase at $T = 0$ with parameters (6). The Fermi level lies in the gap. Nondispersed levels at the bottom of the conduction band and at the top of the valence band

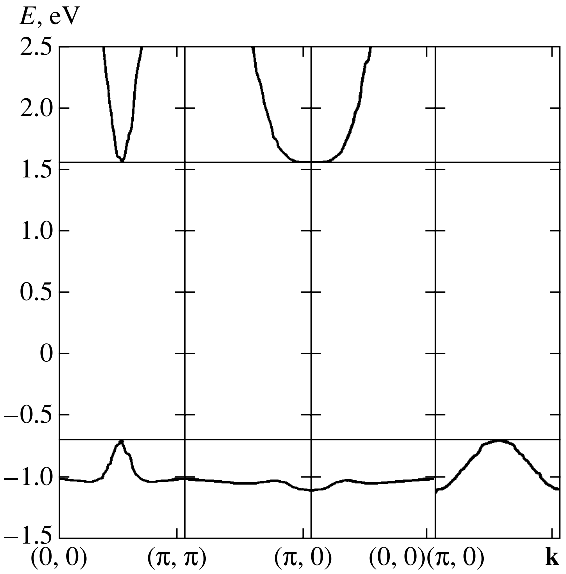

Fig. 2. Dispersion relation for the top of the valence band and the bottom of the conduction band for an undoped CuO_2 layer in the AFM phase.

show the virtual states with zero spectral weight, which were obtained in [16]. It is these levels that form, as a result of doping, the in-gap states with a spectral weight proportional to the doping level.

Figure 3 shows a part of the density of states near the top of the valence band of the CuO_2 layer doped with holes in the AFM phase. We denote by $N_1(E)$ the contribution to $N(E)$ from the top of the valence band; this contribution is characterized by a narrow peak associated with Van Hove singularities. As a result of doping, a virtual level acquires dispersion, the density $N_2(E)$ within the gap states being low in view of the smallness of the hole concentration. The in-gap states are responsible for a red shift of the top of the valence band; however, this shift is difficult to observe in experiments in view of the low spectral density. The main spectral weight $N_1(E)$ of the peak experiences a blue shift approximately equal to 0.1 eV. One more peak appearing at the "shoulder" of the main peak corresponds to a Van Hove singularity at point $\mathbf{k} = (\pi, 0)$. Naturally, only an insignificant blue shift is observed against the background of the density of states from the entire valence band (see Fig. 1) since the value of $N_1(E)$ is much smaller than the total density of states. Thus, the absorption edge at low doping levels in the AFM phase is formed with the participation of states whose presence at the top of the valence band is associated with strong electron correlations. The spectral density of such an individual state may change in the limits determined by the sum rule determined by the presence of a "constraint" (inaccessibility of a part of multielectron states due to strong electron correlations),

$$\int A_1(\mathbf{k}, E)dk = N_1(E),$$

$$\int A_2(\mathbf{k}, E)dk = N_2(E),$$

$$\int N_1(E)dE + \int N_2(E)dE = 1,$$

which were used in calculating the position of the Fermi level. Thus, the position of the Fermi level can be determined from the equation

$$x = \int_{\mu}^{\infty} N_1(E)dE + \int_{\mu}^{\infty} N_1(E)dE,$$

where the spectral density and the density of states depend on the doping component concentration. In other words, the situation takes place opposite to that in the hard band model. It can be seen from Fig. 3 that the density of states $N_2(E)$ of an impurity-type band has dips in the vicinity of $E \approx -0.8$ eV, which correspond to the pseudogap between the valence band and the virtual level. The main contribution to the formation of the pseudogap comes from the states in the vicinity of point $(\pi, 0)$ of the Brillouin zone. For $x \sim 0.01$, for which the AFM phase is preserved after doping, the spectral weight of the in-gap state is small in view of the smallness of x. For this reason, we show in Fig. 3 for better visualization the density of states for large values of x from the range of weakly doped compositions. For $x = 0.10$, the long-range AFM order is not observed any longer, but a short-range order of the AFM type takes place in this range with a correlation length ξ_{AFM} larger than the length l of formation of the electron structure (analogous to the mean free path); $l = 2$–3 lattice parameters. Consequently, as a reasonable interpolation between the AFM ordering with a long-range order ($x < 0.03$) and the strongly doped paramagnetic region ($x > 0.18$), we disregard the difference between the long-range and short-range AFM order in weakly and optimally doped samples and calculate the band structure in the AFM phase; for $x > x_{\text{opt}}$, the band structure in the PM phase is calculated. The vertical dashed line in Fig. 3 shows the position of the Fermi level. The evolution of the position of the Fermi level upon doping (Fig. 4) indicates the possibility of Fermi level pinning as a result of simultaneous increase in the carrier concentration and the spectral density at a virtual level. Precisely such a behavior of the Fermi level was observed in [20] for $La_{2-x}Sr_xCuO_4$. In the PM phase, the impurity band merges with the valence band and the Fermi level for $x > 0.15$ is displaced towards the bottom of the valence band.

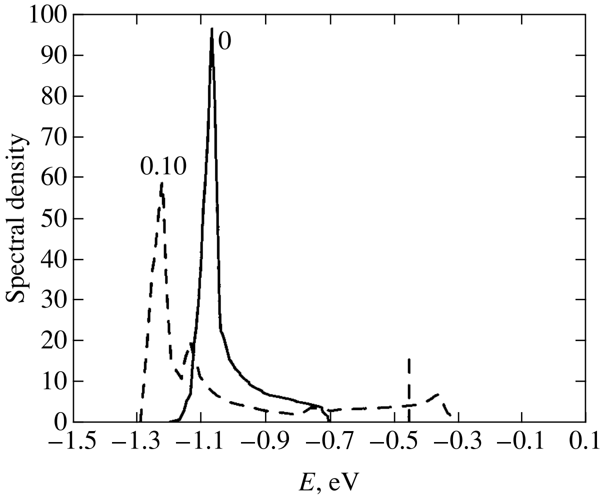

Fig. 3. Density of states of a CuO_2 layer doped with holes with concentration x in the AFM phase.

4. FERMI SURFACE

Figure 5 shows the dynamics of the Fermi surface cross section upon a change in the doping level for the PM phase. For lower doping levels, no agreement with the experimental data [21] could be reached. This can be explained by the necessity of a more detailed

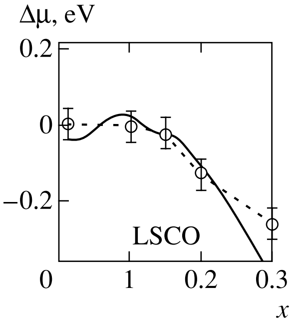

Fig. 4. Chemical potential shift as a function of the doping level. The solid curve describes the results of GTBM computations ($x < 0.15$ for the AFM phase and $x > 0.15$ for the PM phase). The experimental data for $La_{2-x}Sr_xCuO_4$ are borrowed from [20].

description of the spin system for lower doping levels. The Fermi surface is subjected more strongly to the effect of various parameters of the system and not only the chemical potential, which is an integrated characteristic. In our computations, the optimal doping level for which the Fermi level reaches a Van Hove singularity in the PM phase is observed at $x^T_{\rm opt} \approx 0.55$, while the experimental value is $x^{\rm exp}_{\rm opt} \approx 0.18$. Assuming that we were incorrect only while determining the scale of the doping level and the evolution of the Fermi surface corresponds to the experimental results, we describe our results in units of relative doping levels $x^T = x^* x^T_{\rm opt}/x^{\rm exp}_{\rm opt}$. Indeed, it can be seen in Fig. 5 that the results of calculations are in good qualitative agreement with experimental data. Thus, the evolution of the FS in LSCO is reproduced qualitatively in the concentration range of $x > 0.15$. Quantitative discrepancies are observed for the absolute value of $x_{\rm opt}$. A departure from the Hubbard I approximation in which Eq. (4) was derived and the inclusion of short-range order spin correlations will apparently narrow the band and will lead to correct values of concentrations. An analogous conclusion was drawn for the *t–J* model in [22, 23].

5. TEMPERATURE DEPENDENCE OF ARPES SPECTRA

Kim *et al.* [24] obtained the results of a peculiar variation of ARPES spectra for $Sr_2CuO_2Cl_2$ with temperature. With increasing temperature, the height of spectral density peaks decreased and the peaks were shifted towards lower energies. For finite temperatures, the intensity of the ARPES spectra can be described as

$$I(\mathbf{k}, \omega) \propto f_F(\omega) A(\mathbf{k}, \omega),$$

where $A(\mathbf{k}, \omega)$ was calculated by formula (5) with the help of two-time Green temperature functions. Usually, the temperature dependence of the photoelectron spectra in the vicinity of the Fermi level is determined by the spread of the Fermi distribution function. However, in a system with strong correlations, additional, stronger mechanisms of temperature dependence associated with the spectral density itself come into play. In our case of the undoped CuO_2 layer, the filling factors depend on temperature in view of the temperature dependence of the population of the upper spin level $\varepsilon^{-\sigma}_{1,A}$. In the absence of spin fluctuations, we would have $\langle S^Z_A \rangle = 1/2$ and $n^{-\sigma}_1 = 0$. However, zero-point spin fluctuations take place in the AFM phase even at $T = 0$, and $\langle S^Z_A \rangle = 1/2 - n^{-\sigma}_1(0)$ and $n^{-\sigma}_1 = n(T)$. As the temperature increases, the occupational number $n(T)$ for the upper spin layer increases. The value of $n(T)$ can be calculated quite easily in the spin-wave approximation (which gives $n^{-\sigma}_1 \approx 2$ for $T = 0$); however, experiments [24] were performed for rather high temperatures of $T \le T_N$, at which the evaluation of the magnon concentration is a complicated problem, which is beyond the framework of this paper. For this reason, instead of $n(T)$, we calculate the spectral density at three points corresponding to

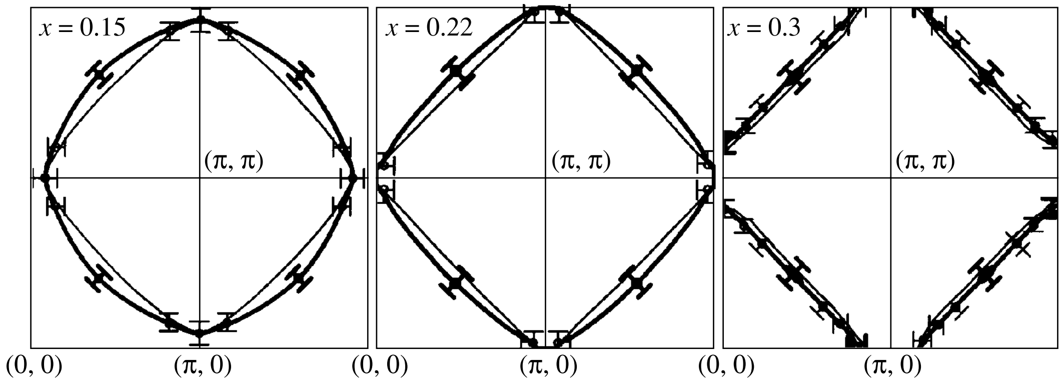

Fig. 5. Evolution of the Fermi surface upon an increase in the hole concentration in $La_{2-x}Sr_xCuO_4$. Thin curves describe the results of GTBM computations, while bold curves are ARPES spectroscopic data from [21].

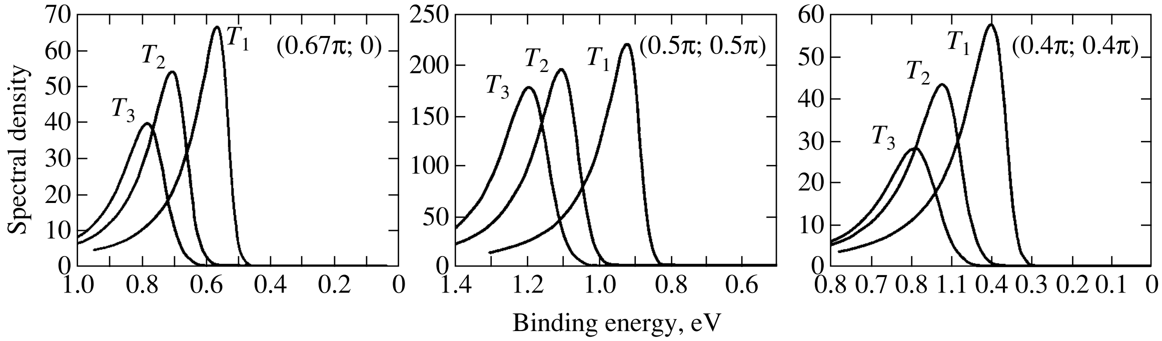

Fig. 6. Temperature dependence of the spectral density for points (2π/3; 0), (π/2; π/2), and (0.4π; 0.4π).

temperatures $T_1 < T_2 < T_3$ such that $n(T_1) = 0.3$, $n(T_2) = 0.4$, and $n(T_3) = 0.5$, point T_3 lying above T_N. It can be seen from Fig. 6 that for all three points of the Brillouin zone at which the temperature dependence of the ARPES spectra, we obtained qualitative agreement with the experiment; for $Sr_2CuO_2Cl_2$, the spectral density peak amplitude at point $\mathbf{k} = (0.67\pi; 0)$ decreases upon an increase in temperature faster than at point $\mathbf{k} = (0.5\pi; 0.5\pi)$, in accordance with the experimental data.

6. CONCLUSIONS

A systematic analysis of the evolution of the electronic structure with doping in LSCO within the GTBM reveals the following three distinct features.

First, the Fermi level in the range of low-doped *p*-type HTSC compounds is pinned by the states of a virtual level in the region of the dielectric gap and exhibits no shift. Indeed, in view of the presence of the undoped virtual CuO_2 level at the top of the valence band, which acquires a finite spectral density and dispersion as a result of doping, the Fermi level is pinned by these states and not by the valence band itself.

Second, as the dopant concentration increases in the PM phase, the evolution from the hole-type FS with the center at $k = (\pi, \pi)$ to the electron-type FS with the center at $k = (0, 0)$ takes place [21]. This scenario is reproduced in calculations. However, a discrepancy is observed between the theoretical and experimental values of x_{opt} in this case. Since the calculated values of x_{opt} for the AFM phase are lower ($x_{\text{opt}} \approx 0.28$–$0.3$), we believe that this discrepancy is due to the simplified description of spin correlations and hope that the correct value of x_{opt} can be obtained by taking into account the short-range spin order.

Third, in the AFM phase of HTSC compounds, our computations reproduce the pseudogap between the "impurity" band and the top of the valence band. The pseudogap is destroyed together with the impurity band upon an increase in the doping level since the dispersion of electrons in the PM phase is similar to dispersion of optimally doped HTSC compounds.

In addition, the band structure of quasiparticles and the spectral density strongly depend on the temperature in the AFM phase; the nature of this dependence is associated with the redistribution of the spectral weight between various quasiparticles upon a change in temperature. Neither the temperature nor the concentration dependence of the band structure can be obtained using the traditional one-electron band approach; these dependences are consequences of the specific band structure of quasiparticles in strongly correlated electronic systems [25].

ACKNOWLEDGMENTS

This study was supported financially by the Federal Special Program "Integration" (grant no. B0017), the Russian Foundation for Basic Research and Krasnoyarsk Regional Science Foundation "Enisei" (grant no. 02-02-97705), Russian Foundation for Basic Research (project no. 03-02-16124), INTAS (grant no. 01-0654), and the program "Quantum Macrophysics" of the Russian Academy of Sciences.

REFERENCES

1. Z.-X. Shen and D. S. Dessau, Phys. Rep. **253**, 1 (1995).
2. D. S. Marshall, D. S. Dessau, A. G. Loeser, *et al.*, Phys. Rev. Lett. **76**, 4841 (1996).
3. H. Ding, M. R. Norman, T. Yokoya, *et al.*, Phys. Rev. Lett. **78**, 2628 (1997).
4. H. Ding, M. R. Norman, J. C. Campuzano, *et al.*, Phys. Rev. B **54**, R9678 (1996).
5. H. Ding, T. Yokoya, J. C. Campuzano, *et al.*, Nature **382**, 51 (1996).
6. P. J. White, Z.-X. Shen, C. Kim, *et al.*, Phys. Rev. B **54**, R15669 (1996).
7. A. G. Loeser, Z.-X. Shen, D. S. Dessau, *et al.*, Science **273**, 325 (1996).
8. J. M. Harris, Z.-X. Shen, P. J. White, *et al.*, Phys. Rev. B **54**, 15665 (1996).
9. M. R. Norman, H. Ding, M. Randeria, *et al.*, Nature **392**, 157 (1998).
10. J. M. Harris, P. J. White, Z.-X. Shen, *et al.*, Phys. Rev. Lett. **79**, 143 (1997).

11. D. M. King, Z.-X. Shen, D. S. Dessau, *et al.*, Phys. Rev. Lett. **73**, 3298 (1994).
12. M. C. Schabel, C.-H. Park, A. Matsuura, *et al.*, Phys. Rev. B **55**, 2796 (1997).
13. A. Damascelli, D. H. Zu, and Z.-X. Shen, J. Electron Spectrosc. Relat. Phenom. **117–118**, 165 (2001).
14. E. G. Maksimov, Usp. Fiz. Nauk **170**, 1033 (2000) [Phys. Usp. **43**, 965 (2000)].
15. S. G. Ovchinnikov and I. S. Sandalov, Physica C (Amsterdam) **161**, 607 (1989).
16. V. A. Gavrichkov, S. G. Ovchinnikov, A. A. Borisov, and E. G. Goryachev, Zh. Éksp. Teor. Fiz. **118** (2), 422 (2000) [JETP **91**, 369 (2000)].
17. Ya. B. Gaididei and V. M. Loktev, Phys. Status Solidi B **147**, 307 (1988).
18. R. O. Zaĭtsev, Zh. Éksp. Teor. Fiz. **70**, 1100 (1976) [Sov. Phys. JETP **43**, 574 (1976)].
19. S. G. Ovchinnikov, Phys. Rev. B **49**, 9891 (1994).
20. N. Harima, J. Matsuno, A. Fujimori, *et al.*, Phys. Rev. B **64**, 220507 (2001).
21. A. Ino, C. Kim, M. Nakamura, *et al.*, Phys. Rev. B **65**, 094504 (2002).
22. N. M. Placida and V. S. Oudovenko, Phys. Rev. B **59**, 11949 (1999).
23. A. Sherman and M. Schreiber, Phys. Rev. B **65**, 134520 (2002).
24. C. Kim, F. Ronning, A. Damascelli, *et al.*, Phys. Rev. B **65**, 174516 (2002).
25. V. V. Val'kov and S. G. Ovchinnikov, *Quasi-Particles in Strongly Correlated Systems* (Sib. Otd. Ross. Akad. Nauk, Novosibirsk, 2001).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 781–787.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 871–877.
Original Russian Text Copyright © 2003 by Khachaturov.

SOLIDS
Electronic Properties

Characteristic Features of Tunnelling in Small-Area Junctions

A. I. Khachaturov

Galkin Donetsk Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 83114 Ukraine
e-mail: khach@sts.dipt.donetsk.ua

Received December 27, 2002

Abstract—The possibility of manifestation of quantization in the transverse component of the electron wave vector in tunnel junctions with a small transverse size is considered. It is shown that the quantization of the transverse component of the wave vector of tunnelling electrons becomes noticeable against the total background of the dependence of the differential tunnel conductivity on voltage when the Fermi energy of one of the electrodes of a tunnel junction is low (on the order of 1 eV), and one of its transverse dimensions is small (on the order of a few hundreds Ångströms). It is noted that the possibility of standing waves emerging in tunnel junctions should be taken into account in analyzing spectroscopic quasiparticle excitations, which is based on an analysis of the second derivative d^2I/dV^2 as a function of voltage V. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Size-quantization effects in thin metallic films have been studied by the electron tunnelling method quite extensively [1–9]. However, in all these publications (both theoretical and experimental), a traditional configuration of a tunnel junctions was considered; in this configuration, two-dimensional regions of a thin-film electrode were parallel to the tunnel barrier plane. It should be noted that such an arrangement is not the only possible one. Tunnel junction of small cross section [spring-loaded junctions, in which one electrode is a sharp tip; contacts formed by breaking a thin film on a flexible substrate (break junctions); and junctions in which the end face of a thin-film electrode is in contact with the opposite bulk electrode] are widely used for studying high-temperature superconductors and magnetoresistive materials. In all these cases, quantization of one of the transverse components of the wave vector can be expected under certain conditions; in this case, the spectrum of tunnelling electrons may be in the form of a set of two-dimensional zones arranged at right angles to the plane of the tunnel junctions. It should be noted that quantization of the transverse component of the electron wave vector may also take place in the insulating layer itself. Indeed, the transverse component of the electron wave vector, in contrast to its longitudinal component, remains real-valued in the barrier (in the case of specular tunnelling, this component preserves its value in all the three parts of the tunnel junction). If the electron wave function attenuates exponentially in the direction of tunnelling, it continues to be of the wave type in directions parallel to the barrier plane. Consequently, for small transverse dimensions of a tunnel junction, the emergence of standing waves in the barrier can be expected. It will be shown here that quantization of the spectrum of tunnelling electrons in such junctions must be manifested in the differential tunnel conductivity $\sigma(V) = dI/dV$ for a favorable combination of the barrier parameters and the electron characteristics of the electrodes.

2. FORMULATION OF THE MODEL

For definiteness, we consider a tunnel junction formed by a thin metallic film whose end face adjoins a bulk metallic electrode (see the inset to Fig. 1). Junctions with such a configuration were used, for example, in [10]. The tunnel current will be calculated using an algorithm similar to the one employed in [7]. We will proceed from the same assumptions on the nature of the tunnel barrier and the mode of passage of the tunnel current; namely, we assume that the potential barrier has a trapezoidal shape,

$$\varphi(z, V) = \varphi_1 + (\varphi_2 - eV - \varphi_1)z/d,$$

and the total electron energy E and component $k_\parallel$ of the electron wave vector along the barrier plane are conserved during tunnelling. In other words, we assume that tunnelling is elastic and specular by nature. We assume that the film and the bulk electrodes are not made of traditional Fermi metals with a quadratic dispersion relation and the barrier transparency can be determined using the BCS approximation (see, for example, formula (4) in [7]). For simplicity, we assume that the effective mass is equal to the mass of a free electron in all three parts of the tunnel junction. First, we consider the case when the Fermi energy of the initial electrode is lower than the Fermi energy of the final

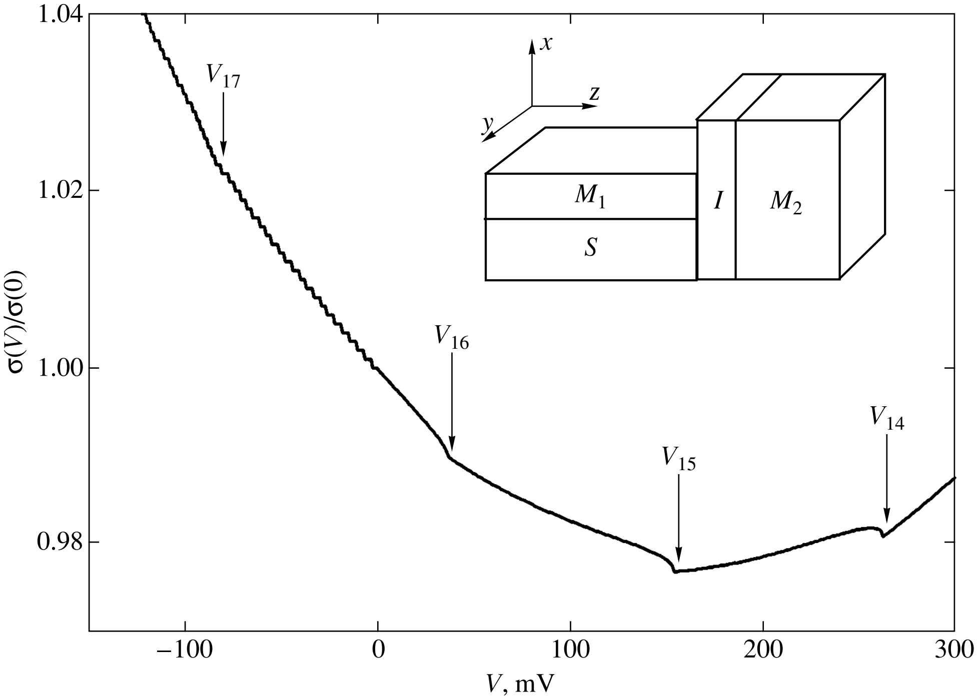

Fig. 1. Differential conductivity of a tunnel junction with the quantized transverse wave vector component in one of the electrodes. The Fermi energies E_{F_1} and E_{F_2} of the electrodes are equal to 1 and 1.3 eV, respectively. The thickness L of the size-quantized electrode is equal to 100 Å. The parameters of the rectangular potential barrier: height $\varphi_1 = \varphi_2 = 4$ meV and thickness $d = 10$ Å. The inset schematically shows the junction under investigation: M_1 is the size-quantized thin-film electrode, M_2 is the ordinary metallic electrode, I is the insulating layer, and S is the substrate.

electrode, $E_{F_1} < E_{F_2}$. We assume at the beginning that the thin-film electrode is homogeneous in thickness L. At $T = 0$, the number N of filled two-dimensional bands of the quasicontinuous spectrum of this electrode is equal to the integral part of $\sqrt{E_F/W}$, where $W = (\pi\hbar)^2/2mL^2$ is the energy corresponding to the first allowed value $k_{x_1} = \pi/2L$ of the transverse component of the wave vector. Figure 2 shows one of such subbands. We apply bias voltage $eV < E_{F_1} - E_n$ to the tunnel structure under study. In this case, electrons of the initial electrode, which are located in the half-ring formed by the lines of intersection of the constant-energy surfaces $E_{F_1} = \text{const}$ and $E_{F_1} - eV = \text{const}$ with plane $E_x = E_n = \text{const}$ and moving towards the barrier $k_z > 0$ can tunnel to the opposite electrode (see Fig. 2). The contribution from these electrons to the tunnel current can be regarded as the difference in the current contributions from the half-ring of radius

$$k_1 = \sqrt{2m(E_{F_1} - E_n)}/\hbar$$

and the half-ring of radius

$$k_2 = \sqrt{2m(E_{F_1} - E_n - eV)}/\hbar.$$

In contrast to [7], the electrons belonging to one subband in this case possess different group velocities and different tunnelling probabilities $P(E_z, V)$. The values of these quantities are the same for particles located on chords AB and $A'B'$ (see Fig. 2). The contribution to the tunnel current from the electrons located on chord AB can be determined by multiplying the number of these electrons by the group velocity, tunnelling probability $P(E_z, V)$, and electron charge:

$$J_{AB} = \frac{ek_zP(E_z, V)}{\pi m}\sqrt{2m(E_{F_1} - E_n - E_z)}.$$

Integrating with respect to k_z and passing from an integral with respect to momenta to integrals with respect to energies E_z, we obtain the tunnel current introduced by the nth subband:

$$J_n(V) = K\left[\int_0^{E_{F_1} - E_n} P(E_z, V)\sqrt{E_{F_1} - E_n - E_z}\,dE_z - \int_0^{E_{F_1} - eV - E_n} P(E_z, V)\sqrt{E_{F_1} - E_n - E_z - eV}\,dE_z\right], \tag{1}$$

$$eV < E_{F_1} - E_n,$$

where $K = e\sqrt{2m}/(\pi\hbar)^2$. For $eV \geq E_{F_1} - E_n$, all the electrons of the *n*th subband can take part in tunnelling; consequently,

$$J_n(V) = K\int_0^{E_{F_1}-E_n} P(E_z, V)\sqrt{E_{F_1}-E_n-E_z}\,dE_z, \qquad eV > E_{F_1}-E_n. \tag{2}$$

Differentiating $J(V)$ with respect to V, we obtain analytic expressions for contribution $\sigma_n(V)$ by the *n*th subband to the tunnel conductivity $\sigma(V)$ in the case when $eV < E_{F_1} - E_n$,

$$\sigma_n(V) = K[A_n(V) - B_n(V) + C_n(V) + D_n(V)], \tag{3}$$

where

$$A_n(V) = \int_0^{E_{F_1}-E_n} \frac{\partial P}{\partial V}\sqrt{E_{F_1}-E_n-E_z}\,dE_z,$$

$$B_n(V) = \int_0^{E_{F_1}-E_n-eV} \frac{\partial P}{\partial V}\sqrt{E_{F_1}-E_n-E_z-eV}\,dE_z,$$

$$C_n(V) = \int_0^{E_{F_1}-E_n-eV} e\frac{\partial P}{\partial E_z}\sqrt{E_{F_1}-E_n-E_z-eV}\,dE_z,$$

$$D_n(V) = eP(0, V)\sqrt{E_{F_1}-E_n-eV}.$$

For $eV > E_{F_1} - E_n$, we have

$$\sigma_n(V) = KA_n(V). \tag{4}$$

Partial derivative $\partial P/\partial E_z$ in the expressions for A_n, B_n, and C_n appears as a result of integration by parts, which was carried out to get rid of the improper integral. In our calculations, most of the energy quantities (E_{F_1}, E_z, and E_n), as well as ϕ_1 and ϕ_2, are measured from the bottom of the conduction band of the initial electrode. The only exceptions are the Fermi energy E_{F_2}, which is measured from the bottom of the conduction band of the final electrode and potential difference eV. It should be noted that the values of ϕ_1 and ϕ_2 are connected with the conventionally used barrier heights measured from the Fermi level via the relations $\varphi_1 = \phi_1 - E_{F_1}$ and $\varphi_2 = \phi_2 - E_{F_1}$.

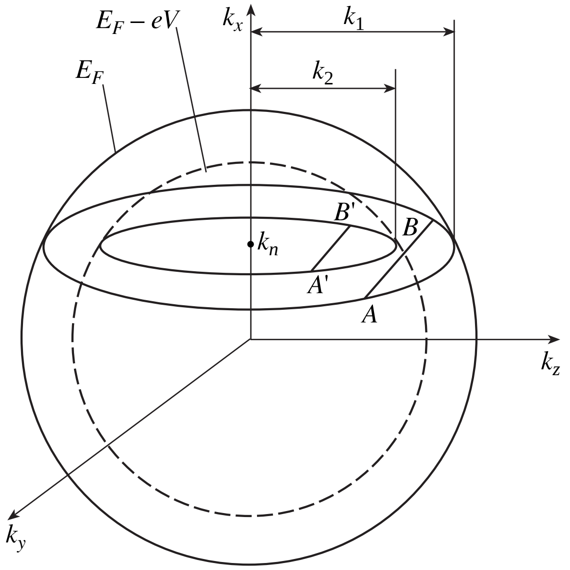

Fig. 2. Reciprocal space of a size-quantized thin-film electrode for $E_{F_1} < E_{F_2}$. In the case of quantization of transverse wave vector component k_z, two-dimensional subbands are perpendicular to the plane of the tunnel junction.

The total differential conductivity can be determined by summation over all two-dimensional subbands:

$$\sigma(V) = 2\sum_{n=1}^{N}\sigma_n(V).$$

If the Fermi energy of the initial electrode exceeds the corresponding energy of the opposite electrode, the situation is slightly complicated. As a matter of fact, not all electrons lying in the energy interval from E_{F_1} to $E_{F_1} - eV$ can satisfy the specular tunnelling conditions in this case; i.e., not all electrons can simultaneously preserve the values of total energy E and parallel momentum $k_{\parallel}$ during tunnelling. For this reason, the bands for which $E_n > E_{F_2}$ make zero contribution to the tunnel current for $eV < E_n - E_{F_2}$. Tunnelling from these bands is possible only for $eV > E_n - E_{F_2}$. Tunnelling electrons in this case lie in the hatched segment of the circle of radius k_1, formed by the secant

$$k_z = k_{\min} = \sqrt{2m(E_{F_1} - E_{F_2} - eV)}/\hbar$$

(Fig. 3a). The contributions to the tunnel current from

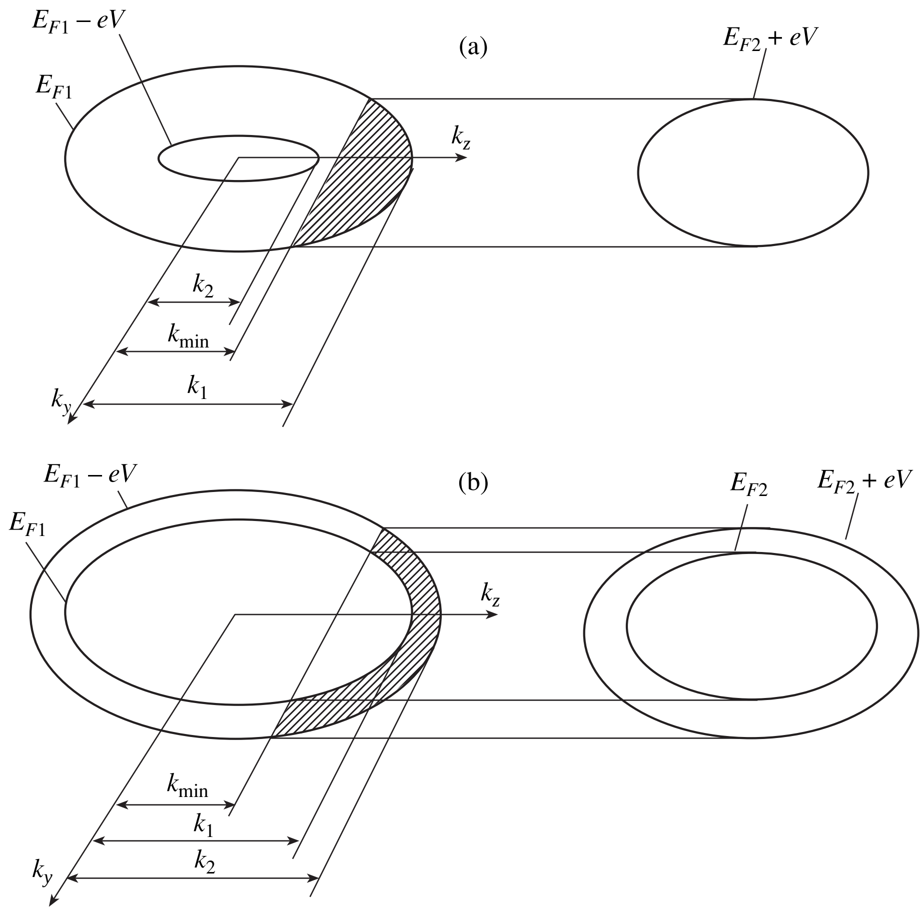

Fig. 3. Reciprocal space of a size-quantized thin-film electrode for $E_{F_1} > E_{F_2}$: (a) for $E_n > E_{F_2}$, electrons from the *n*th subband can participate in the tunnelling process starting only from voltages with $eV \ge E_{F_2} - E_n$; all such electrons are located in the hatched segment; (b) for $E_n > E_{F_2}$, tunnelling is possible starting from zero voltages; tunnelling electrons lie in the hatched region of the ring.

these bands are given by

$$J_n(V) = 0, \quad eV < E_n - E_{F_2}.$$

$$J_n(V) = K \int\limits_{E_{F_1} - E_{F_2} - eV}^{E_{F_1} - E_n} P(E_z, V)\sqrt{E_{F_1} - E_n - E_z}\,dE_z, \quad (5)$$

$$E_n - E_{F_2} < eV < E_{F_1} - E_{F_2}.$$

The corresponding expressions for differential conductivity have the form

$$\sigma_n(V) = 0, \quad eV < E_n - E_{F_2}, \quad (6)$$

$$\sigma_n(V) = K[A_{1_n}(V) + D_{1_n}(V)],$$

$$E_n - E_{F_2} < eV < E_{F_1} - E_{F_2},$$

where

$$A_{1_n}(V) = \int\limits_{E_{F_1} - E_{F_2} - eV}^{E_{F_1} - E_n} \frac{\partial P}{\partial V}\sqrt{E_{F_1} - E_n - E_z}\,dE_z,$$

$$D_{1_n}(V) = eP(E_{F_1} - E_{F_2} - eV, V) \times \sqrt{E_{F_2} + eV - E_n}.$$

For $eV \ge E_{F_1} - E_{F_2}$, the expressions for tunnel current and differential conductivity of these bands coincide with expressions (2) and (4), respectively.

Energy bands satisfying the condition $E_n < E_{F_2}$ contribute to the tunnel current starting from zero voltages. In this case, the electrons participating in the tunnelling are located in the annular segment (Fig. 3b):

$$J_n(V) = K\left[\int\limits_{E_{F_1}-E_{F_2}-eV}^{E_{F_1}-E_n} P(E_z, V)\sqrt{E_{F_1}-E_n-E_z}dE_z\right.$$

$$\left. - \int\limits_{E_{F_1}-E_{F_2}-eV}^{E_{F_1}-eV-E_n} P(E_z, V)\sqrt{E_{F_1}-E_n-E_z-eV}dE_z\right], \tag{7}$$

$$eV < E_{F_1}-E_{F_2}.$$

Differentiating this expression with respect to voltage V we obtain

$$\sigma_n(V) = k[A_{1_n}(V) - B_{1_n}(V) + C_{1_n}(V) + D_{1_n}(V)],$$
$$eV < E_{F_1}-E_{F_2}, \tag{8}$$

where

$$B_{1_n}(V) = \int\limits_{E_{F_1}-E_{F_2}-eV}^{E_{F_1}-E_n-eV} \frac{\partial P}{\partial V}\sqrt{E_{F_1}-E_n-E_z-eV}dE_z,$$

$$C_{1_n}(V) = \int\limits_{E_{F_1}-E_{F_2}-eV}^{E_{F_1}-E_n-eV} e\frac{\partial P}{\partial E_z}\sqrt{E_{F_1}-E_n-E_z-eV}dE_z.$$

The expression for the tunnel current for these bands coincides with Eq. (1) for a voltage of $eV > E_{F_1} - E_{F_2}$ and with Eq. (2) for $eV > E_{F_1} - E_n$. Accordingly, the differential tunnel conductivity is calculated by formula (3) for $eV > E_{F_1} - E_{F_2}$ and by formula (4) for $eV > E_{F_1} - E_n$.

3. DISCUSSION

Our calculations show that singularities associated with the presence of standing waves in the thin-film electrode are not manifested in the differential tunnel conductivity for ordinary metals with a Fermi energy on the order of several electronvolts. However, these singularities become quite noticeable for electrodes with the Fermi energy on the order of an electronvolt. For example, Fig. 1 shows the results of calculation of the differential conductivity for a tunnel junction L = 100 Å thick for the size-quantized electrode and with the Fermi energies of the electrodes equal to E_{F_1} = 1 eV and E_{F_2} = 1.3 eV. The potential barrier parameters were assumed to be as follows: d = 10 Å and $\varphi_1 = \varphi_2$ = 4 eV. Positive voltages in Fig. 1 correspond to tunnelling from the electrode with a lower Fermi energy to the electrode with a higher Fermi energy. In this case, starting from zero voltages, all two-dimensional bands participate in tunnelling. However, while the tunnel current of the nth subband increases due to an increase both in the barrier transparency and in the number of electrons capable of participating in the tunnelling process for $eV \le E_{F_1} - E_n$, the latter factor disappears for $eV > E_{F_1} - E_n$. The band turns out to be completely open, and the increase in the tunnel current slows down. This is reflected in the tunnel conductivity in Fig. 1 for voltages $V_n = (E_{F_1} - E_n)/e$. The negative polarity corresponds to the case when the Fermi energy E_{F_1} of the initial electrode is higher that the corresponding energy E_{F_2} of the final electrode. In this case, part of the energy bands for which $E_{F_1} < E_n < E_{F_2}$ at initial voltages do not participate in tunnelling. The inclusion of these bands is reflected in the region of negative voltages for $V_n = (E_{F_1} - E_n)/e$ in the form of kinks on the $\sigma(V)$ curve in Fig. 1.

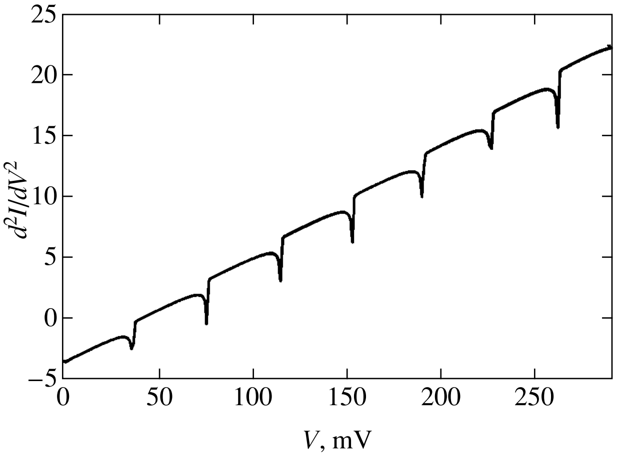

Fig. 4. Dependence of the second derivative d^2I/dV^2 of the tunnel current with respect to voltage on bias voltage V across the junction for an “end face tunnel junction” in which one electrode is a homogeneous size-quantized film L = 250 Å thick. Barrier parameters: $\varphi_1 = \varphi_2$ = 4 meV and thickness d = 10 Å. The Fermi energies E_{F_1} and E_{F_2} of the initial and final electrodes are equal to 1 and 1.3 eV, respectively.

Figure 4 shows the dependence of the second derivative d^2I/dV^2 of the tunnel current on voltage V, which was obtained by numerical differentiation of the curve describing the differential tunnelling conductivity. It can be seen from the figure that size quantization in the homogeneous film leads to the emergence of a distinct structure of almost periodically repeated sharp singularities in the dependence of d^2I/dV^2 on V, which are arranged, generally speaking, asymmetrically relative to zero voltage. For a nonhomogeneous film, the contri-

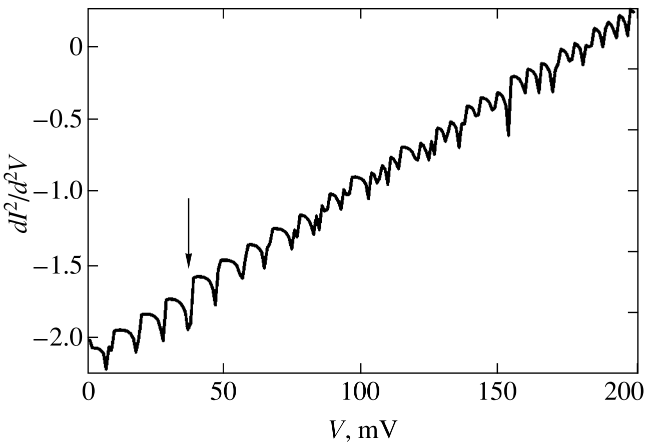

Fig. 5. Second derivative of the tunnel current with respect to voltage of the tunnel junction under investigation with a nonhomogeneous size-quantized electrode as a function of the bias voltage. It is assumed that the junction consists of ten blocks of different thickness L from 245 to 255 Å, which appear with identical weights. The lattice constant a = 2.5 Å. The value of $V_{\rm res}$ = 37 meV (indicated by the arrow) satisfies the geometrical resonance condition for $s/q = 2/5$. The remaining parameters of calculations coincide with the corresponding values in Fig. 1.

bution from the segments of different thickness add up to give a complex spectrum of the dependence of d^2I/dV^2 on V.

One of the main difficulties encountered in the method of electron tunnelling used for reconstructing the electron–phonon interaction spectra for high-temperature superconductors is that the dependence of the second derivative d^2I/dV^2 on V contains, in addition to phonon singularities whose origin is attributed to the electron–phonon interaction, a structure of singularities extending far beyond the phonon spectrum of the materials under study. The presence of singularities of unknown nature on the experimental curve considerably lowers the confidence level of the results; for this reason, rather peculiar hypotheses are used for explaining such singularities. For example, it was proposed in [11] that the singularities in question are due to multiphonon processes. Since, as mentioned above, small-area tunnel junctions in which the transverse wave vector component can be quantized are used for studying high-temperature superconductors, it is quite natural, in our opinion, to assume that such quantization is responsible for the emergence of high-energy singularities on the curve describing the dependence of d^2I/dV^2 on V.

Our calculations show that, if the quantized electrode is not homogeneous over thickness, but consists of individual segments of various thickness, the dependence of d^2I/dV^2 on V has a spectrum rich in singularities in the entire calculated region of voltages. The general form of this spectrum is found to be sensitive to computation parameters such as the average thickness of a quantized film as well as the barrier height and thickness. Nevertheless, the position of several singularities remains unchanged in this case because these singularities reflect the so-called commensurate energy levels. Moreover, with an appropriate choice of computation parameters in the dependence of d^2I/dV^2 on V, it is possible to simulate the so-called geometrical resonance. Such a resonance is possible since the thickness of a nonhomogeneous polycrystalline electrode can vary only discretely: $L = Na$, where a is the unit cell size in the direction of the x axis. Consequently, states exist whose energies are independent of thickness. For example, two-dimensional bands corresponding to a state with a wavelength of $\lambda = 2a$ are present in quasi-two-dimensional spectra of all crystallites irrespective of their thickness. In order to determine the commensurate states, we can use the relation

$$k_{\rm res} = \frac{s\pi}{qa},$$

where s/q is an irreducible fraction (it can easily be verified that, for example, wavelength $\lambda_{\rm res} = 4a$ for $s/q = 1/2$ and, hence, the corresponding two-dimensional bands are present in all crystallites with an even number of layers) [9]. In complete agreement with the above considerations, Fig. 5 shows the resonant structure for voltages,

$$V_{\rm res} = \frac{1}{e}\left(\frac{(\hbar k_{\rm res})^2}{2m} - E_F\right),$$

which corresponds to $s/q = 2/5$, $V_{\rm res} = 37$ mV.

Assuming that the value of L is infinitely large and replacing the summation over n by integration with respect to E_x, we arrive at the expression for the tunnel current through an ordinary junction with nonquantized metallic electrodes,

$$J_m(V) = K_m\left(\int_0^{E_F}\frac{dE_x}{\sqrt{E_x}}\int_0^{E_F-E_x}P(E_z, V)\sqrt{E_F-E_x-E_z}\,dE_z \right. \tag{9}$$
$$\left. - \int_0^{E_F-eV}\frac{dE_x}{\sqrt{E_x}}\int_0^{E_F-eV-E_x}P(E_z, V)\sqrt{E_F-E_x-E_z-eV}\,dE_z\right),$$

where $K_m = em/(\pi\hbar)^3$. Naturally, the results of calculations made in accordance with this formula completely coincide with the results obtained by using the standard formulas for tunnel current and tunnelling probability [9].

ACKNOWLEDGEMENTS

The author is grateful to V.M. Svistunov for all-around support.

REFERENCES

1. V. N. Lutskiĭ, D. N. Korneev, and M. I. Elinson, Pis'ma Zh. Éksp. Teor. Fiz. **4**, 267 (1966) [JETP Lett. **4**, 179 (1966)].
2. G. A. Gogadze and I. O. Kulik, Fiz. Tverd. Tela (Leningrad) **7**, 432 (1965) [Sov. Phys. Solid State **7**, 345 (1965)].
3. R. C. Jaklevic and J. Lambe, Phys. Rev. B **12**, 4146 (1975).
4. L. C. Davis, R. C. Jaklevic, and J. Lambe, Phys. Rev. B **12**, 798 (1975).
5. V. M. Svistunov and V. Yu. Tarenkov, Pis'ma Zh. Éksp. Teor. Fiz. **26**, 34 (1977) [JETP Lett. **26**, 30 (1977)].
6. A. A. Galkin, V. M. Svistunov, A. I. D'yachenko, and V. Yu. Tarenkov, Pis'ma Zh. Éksp. Teor. Fiz. **21**, 259 (1975) [JETP Lett. **21**, 118 (1975)].
7. A. I. Khachaturov, Zh. Éksp. Teor. Fiz. **118**, 622 (2000) [JETP **91**, 541 (2000)].
8. Yu. F. Komnik, *Physics of Metal Films* (Atomizdat, Moscow, 1979).
9. E. L. Wolf, *Principles of Electron Tunneling Spectroscopy* (Oxford Univ. Press, New York, 1985; Naukova Dumka, Kiev, 1990).
10. J. S. Tsai, I. Takeuchi, J. Fujita, *et al.*, Physica C (Amsterdam) **157**, 537 (1989).
11. R. Aoki, H. Murakami, T. Kita, *et al.*, Physica B (Amsterdam) **219–220**, 172 (1996).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 788–793.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 878–885.
Original Russian Text Copyright © 2003 by Misochko, Kaidashev, Georgiev, Dekorsy, Zakharchenko.

SOLIDS
Electronic Properties

A Time-Resolved Optical Study of the Paramagnetic Dielectric–Ferromagnetic Metal Transition in $La_{0.7}Ca_{0.3}MnO_3$

O. V. Misochko[a],*, E. M. Kaidashev[b], N. Georgiev[c], T. Dekorsy[c], and I. N. Zakharchenko[d]

[a]*Institute of Solid-State Physics, Russian Academy of Sciences,*
Chernogolovka, Moscow oblast, 142432 Russia
[b]*Institute of Mechanics and Applied Mathematics, Rostov State University,*
pr. Stachki 200/1, Rostov-on-Don, 344090 Russia
[c]*Institute for Ion Beam Physics and Materials Research Forschungszentrum Rossendorf,*
P.O. Box 510119, D-01314, Dresden, Germany
[d]*Research Institute of Physics, Rostov State University, pr. Stachki 194, Rostov-on-Don, 344090 Russia*
**e-mail: misochko@issp.ac.ru*

Received February 20, 2003

Abstract—The relaxation dynamics of charge carriers in an epitaxial $La_{0.7}Ca_{0.3}MnO_3$ film was studied by the light reflection method with a femtosecond time resolution in a wide temperature range. The relaxation of time-resolved photoinduced reflections was shown to be two-component in the whole temperature range including the paramagnetic dielectric–ferromagnetic metal transition at $T \approx 150$–160 K. The fast relaxation component had a maximum lifetime ($\tau \sim 500$ ps) in the transition region, and its contribution to relaxation increased as temperature decreased. The lifetime of the slow component was minimum ($\tau \sim 15$ fs) in the transition region. In addition, fast oscillations assigned to coherent phonons appeared in photoinduced responses at $T < 200$ K. The dephasing time of these oscillations increased as temperature decreased, whereas their frequency changed insignificantly. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The discovery of giant magnetoresistance in perovskite manganese oxides (such as $La_{1-x}A_xMnO_3$, where A = Sr, Ca, or Ba) has stimulated extensive studies of these systems. This interest in manganites is caused by at least two reasons. First, the giant magnetoresistance effect is potentially important for several technical applications, such as the design of magnetic recording heads and devices for storing information. Secondly, perovskite manganites are of interest for basic science. Unlike standard (elemental and oxide) ferromagnets, in which the electron spin is weakly coupled with the crystal lattice, manganites are characterized by mutually related spin, charge, orbital, and lattice degrees of freedom, which results in a fine balance of interactions responsible for their complex phase diagram [1]. This phase diagram includes metal–dielectric transitions, charge and orbital ordering, and electronic phase separation, which manifests itself by the formation of spin and charge stripes and other inhomogeneous states. The phase diagram of manganites is in many respects similar to the diagram of high-T_c superconductors, because materials of both classes are strongly correlated systems in which the potential energy of carriers is comparable with or larger than their kinetic energy. The conductivity (charge transport) of manganites can to a great extent be determined by the spin subsystem and caused by the double exchange mechanism, whereas the Jahn–Teller distortion of $Mn^{3+}O_6$ octahedra couples the spin and lattice degrees of freedom. The ferromagnetic transition in manganites is unusual and, as a rule, coincides with the metal–dielectric transition, which is responsible for the giant magnetoresistance effect. The results of experimental studies of these materials are summarized in review [1], and the most important theoretical models are described in reviews [2–5]. An analysis of these works shows that, currently, no consensus has been attained in understanding the physics of manganites. This makes further inquiries capable of providing new information a topical task.

Advances in laser technology, which allowed the laser pulse width to be reduced to several femtoseconds, opened up possibilities for studying the dynamics of excitation in solids in real time. At present, such studies are being actively performed for condensed media [6, 7]. Understanding the mechanism of restoring equilibrium disturbed by a pumping pulse allows the special features of the dynamics of elementary excitations and their interaction to be revealed. Quite recently, the potential of optical time-resolved methods has been demonstrates in studies of such strongly correlated systems as high-T_c superconductors, one- and two-dimensional systems with charge density waves, and one-dimensional perovskites [7–11]. This prompted us to undertake a study of one of the lanthanum manganites ($La_{0.7}Ca_{0.3}MnO_3$) by pumping–probing with femtosecond laser pulses to track the relaxation dynamics of quasi-particles in this material in a wide temperature range including the metal–insulator transition. Note that the creation of a nonequilibrium photoinduced hole population in high-T_c superconductors gave additional information about Hubbard constants in the copper and oxygen bands and allowed the charge gap width to be determined [12]. It follows from these results that a systematic study of manganites under the conditions when the electronic subsystem is nonequilibrium offers much promise.

2. THE SAMPLES AND PROCEDURE FOR MEASUREMENTS

The epitaxial $La_{0.7}Ca_{0.3}MnO_3$ films were prepared by off-axis laser spraying on chemically purified $LaAlO_3$ substrates 1.0 × 1.0 × 0.5 cm^3 in size. The scheme for laser spraying was similar to that described in [13]. YAG : Nd^{3+} laser radiation (λ = 1.06 μm, τ = 15 ns, and f = 12.5 Hz) was focused into a 0.2 × 0.2 cm^2 spot on the surface of a rotating ceramic $La_{0.7}Ca_{0.3}MnO_3$ target of the stoichiometric composition at an angle of 45°. The energy density on the surface of the target was about 2 J/cm^2. The oxygen pressure during deposition was 10–30 Pa, and the distance between the target and the substrate was 3.5–4.5 cm. The substrate temperature during film growth was varied in the temperature range 700–760°C. The film thickness exceeded 1000 Å, and the lattice constant was 3860 Å. The X-ray patterns of the films only contained (001) $La_{0.7}Ca_{0.3}MnO_3$ reflections and were free of traces of *hkl*-type orientations with $h, k \neq 0$ or impurity phases. The [100] $La_{0.7}Ca_{0.3}MnO_3$ || [100] $LaAlO_3$ orientation was observed in the substrate plane. The temperature dependence of the resistance of the samples was studied to find that, at T = 140–160 K, the $\partial\rho/\partial T$ derivative changed sign from positive at low temperatures to negative at T > 150–160 K, see inset to Fig. 1. The electric conductivity, although small in magnitude, increased as the temperature rose. The conductivity changed by one to two orders of magnitude after the transition and decreased toward its low-temperature value as the temperature increased further. Measurements performed in an H = 0.9 T magnetic field showed that, at the highest film magnetoresistance, $(\rho_H - \rho_0)/\rho_H \approx 15\%$.

Measurements in the time range were performed using a system for rapid scanning [9] and sapphire titanate laser pulses (λ = 780 nm) width 70 fs in width with a 78-MHz repetition frequency. The train of laser pulses was split into two beams. The probing beam was approximately 50 times weaker than the pumping beam (the probing beam power was 2 mW). The sample was excited by a pumping pulse and probed by the second (probing) pulse with a controlled time lag. The polarizations of the exciting and probing pulses were mutually orthogonal and lay in the film plane. The experimentally measured value was the difference reflection of the excited and unexcited samples $\Delta R(t) = R(t) - R_0$ as a function of delay time t between the pumping and probing pulses. The zero point of the time axis was determined from the autocorrelation function. The sample was placed into an optical helium cryostat, which allowed us to perform measurements in the temperature range 4–330 K.

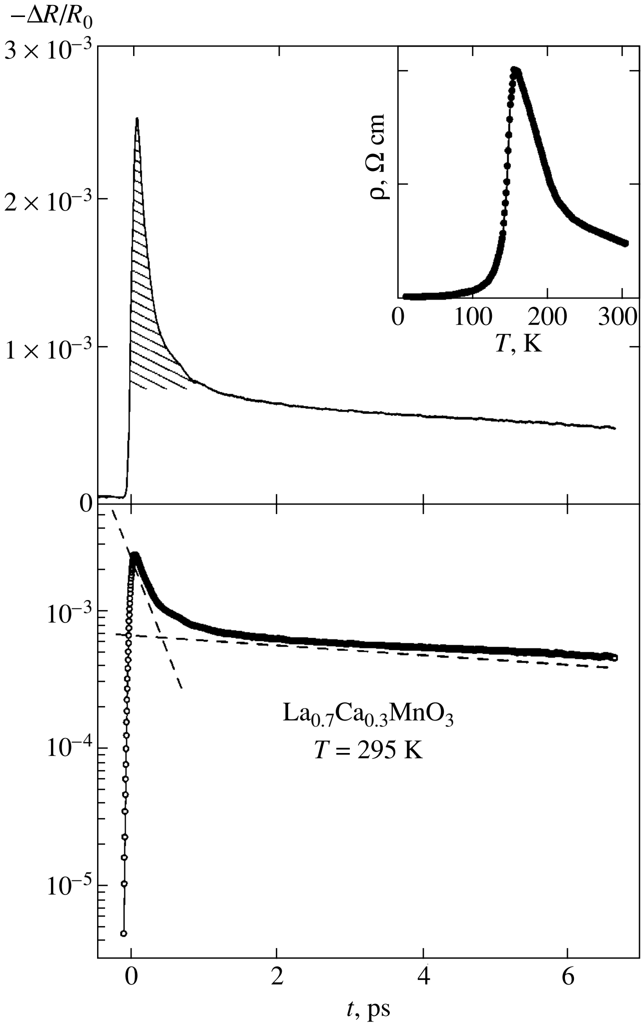

Fig. 1. Time-resolved differential reflection $|\Delta R/R_0|$ for a $La_{0.7}Ca_{0.3}MnO_3$ film at room temperature. The fast relaxation component is hatched. Given in the inset is the temperature dependence of resistance. The dashed lines refer to the fast and slow relaxation channels.

3. EXPERIMENTAL RESULTS AND DISCUSSION

A typical optical response recorded by exciting and subsequently probing a $La_{0.7}Ca_{0.3}MnO_3$ epitaxial film

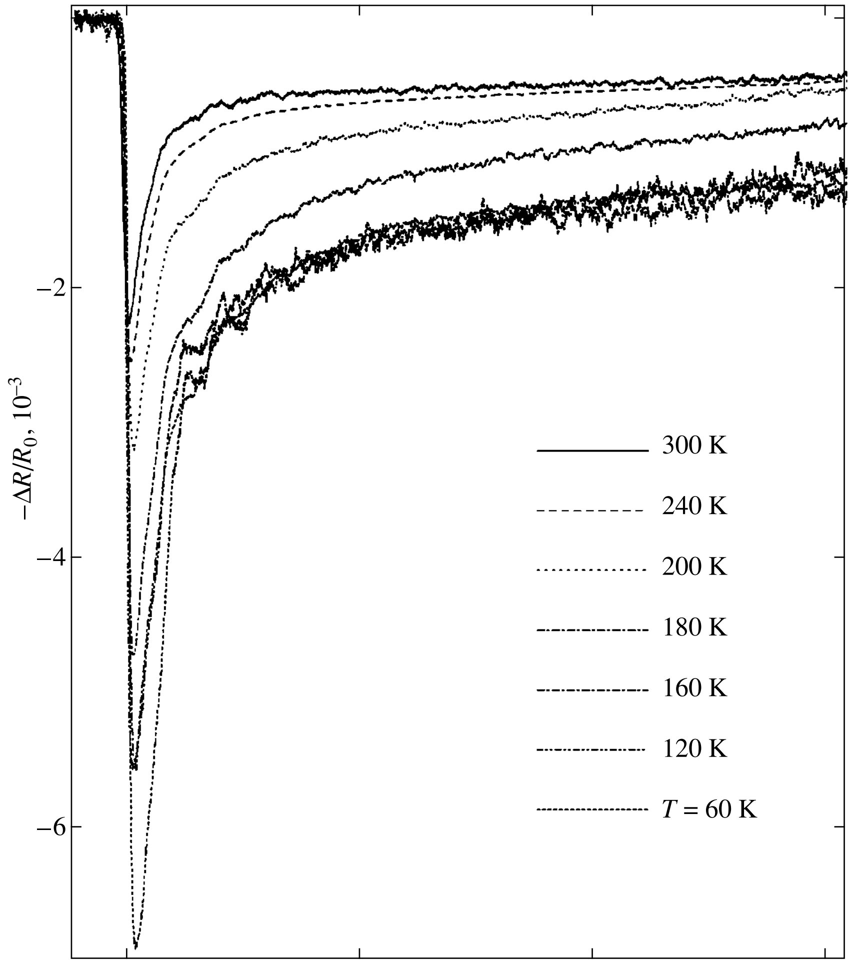

Fig. 2. Temperature dependences of time-resolved differential reflection $\Delta R/R_0$ at several specified temperatures.

by a femtosecond pulse is shown in Fig. 1. The film experiences a transition to the excited state in times on the order of 100–150 fs after the arrival of the pumping pulse. Simultaneously, its reflection decreases to its minimum value. This excited state on the system relaxes to the equilibrium state in times of the order of dozens of picoseconds. As the material under study has hole conductivity, its reflection decreases after the action of a pumping pulse. The relaxation of the photoinduced differential reflection is nonexponential in character and can be approximated by the sum of two (slow and fast) exponential functions,

$$-\Delta R/R_0 = A(T)\exp(-t/\tau_A) + B(T)\exp(-t/\tau_B).$$

This two-component relaxation dynamics is clearly seen in Fig. 1, in which the differential reflection is plotted on the linear and logarithmic scales. The fast component relaxes in times on the order of hundreds of femtoseconds, whereas the lifetime of the slow component amounts to dozens picoseconds.

The relaxation dynamics is shown in Fig. 2 for several temperatures in both the dielectric (paramagnetic) and metallic (ferromagnetic) states. This figure illustrates the temperature dependence of the time-resolved photoinduced response. The two-component relaxation dynamics is observed at all temperatures of our experiments, and the differential photoinduced response $|\Delta R/R_0|$ monotonically increases in magnitude as temperature decreases. In addition, the ratio between the fast and slow relaxation dynamics components changes. The lifetime of the fast component and its contribution to relaxation increase as the temperature approaches that of the largest resistance. Simultaneously, the lifetime of the slow component decreases and becomes minimum in the region of the paramagnetic dielectric–ferromagnetic metal transition temperature. In the ferromagnetic metallic phase, cooling the sample decreases the lifetime of the fast component and slightly increases the lifetime of the slow component. The results that describe temperature-induced changes in the dynamics of relaxation are summarized in Fig. 3. Note that ΔR reflection changes induced by the pump-

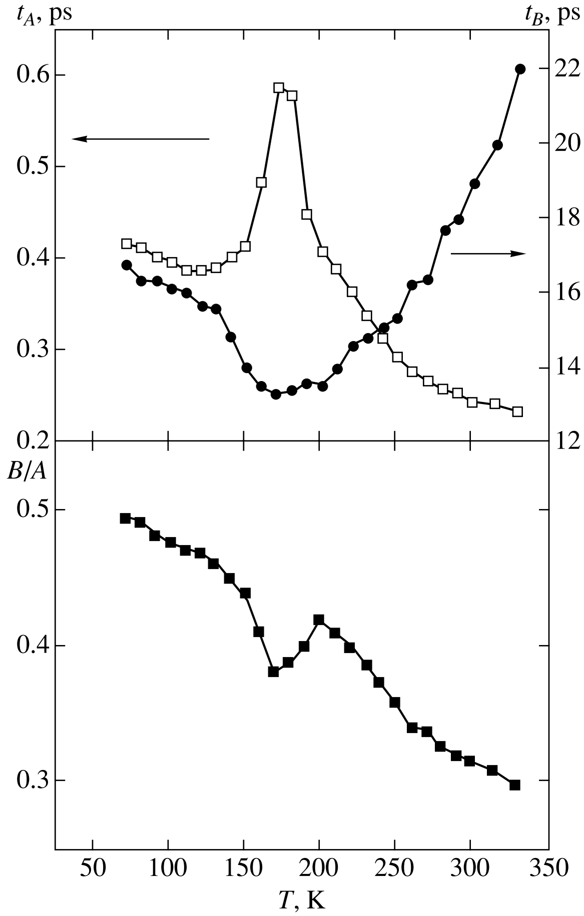

Fig. 3. Temperature dependences of the dynamic characteristics of excitation relaxation. Fast τ_A and slow τ_B relaxation components are given at the top, and the ratio between the fast and slow component contributions, at the bottom.

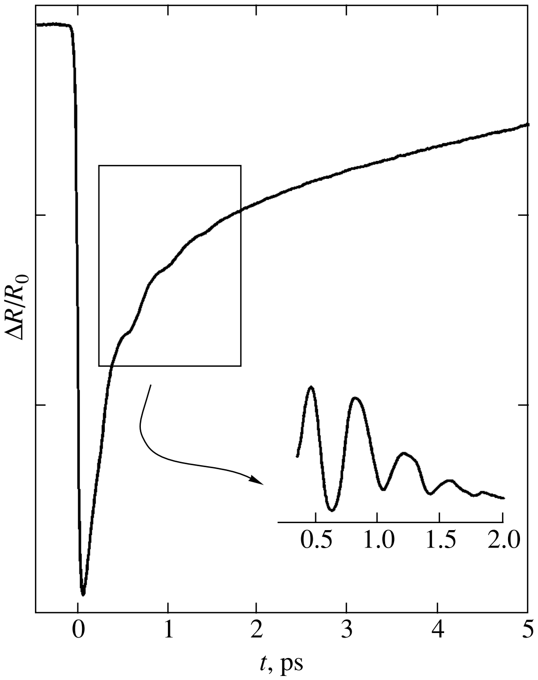

Fig. 4. Time resolved differential reflection $\Delta R/R_0$ of a $La_{0.7}Ca_{0.3}MnO_3$ film at $T = 150$ K. Shown in the inset are the oscillations assigned to coherent phonons.

ing pulse are virtually independent of temperature. Changes in the $\Delta R/R_0$ differential reflection caused by sample cooling are largely determined by changes in the unexcited sample reflection R_0.

Somewhat above the temperature at which resistance is maximum, fast relaxation curve oscillations appear. These oscillations are shown in Fig. 4, where the nonoscillating relaxation component is subtracted. The oscillations can be attributed to coherent phonons excited by ultranarrow laser pulses [7]. The phonons recorded using subpicosecond pulses are called coherent because they have a well-defined phase [7], unlike thermal phonons, which are studied by Raman and infrared reflectance spectroscopy. Such phonons can appear because the exciting pulse width Δt is smaller than the inverse phonon mode frequency Ω^{-1}, $\Delta t < \Omega^{-1}$. Coherent phonons are recorded for $La_{0.7}Ca_{0.3}MnO_3$ films at all temperatures below $T = 200$ K and are absent at the higher temperatures. The phonon modes were identified by numerically Fourier-transforming the time range data into the frequency range; the result is shown in Fig. 5. The period of the oscillations insignificantly decreases and their lifetime considerably increases as the temperature lowers. For instance, at $T = 190$ K, the oscillations are damped during one cycle, whereas the frequency of the oscillations at $T = 150$ K is much higher than the inverse damping time. This follows from a comparison of both time-resolved responses and periodograms of the oscillating response component. One of such periodograms is shown in Fig. 5.

It is expedient to precede a possible interpretation of the results with a presentation of the generally accepted views on the system under study. The $La_{0.7}Ca_{0.3}MnO_3$ compound can be treated as a solid solution formed by $LaMnO_3$ and $CaMnO_3$. Doping $LaMnO_3$ with divalent Ca yields a mixed-valence compound, which contains Mn^{3+} and Mn^{4+} ions. In the double exchange model, the width of the hole band, which depends on the level of doping and temperature, controls the metal–dielectric transition. Exchange coupling alone is, however, insufficient for obtaining the real characteristics of manganites. As distortions caused by the cooperative Jahn–Teller effect already exist in the initial $LnMnO_3$ compound and lower its symmetry from cubic, it has been suggested that taking into account lattice excitations and strong electron–phonon coupling may give a correct description of the system. The Jahn–Teller effect lowers crystal symmetry and removes degeneracy of the e_g d Mn orbitals. Depending on the theoretical model, it is assumed that doping results in the appearance of a hole either on the manganese d orbital [4] or

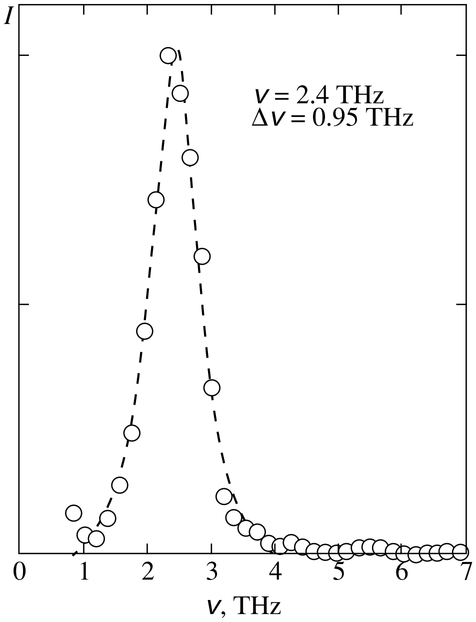

Fig. 5. Fourier transform of the time response of the oscillating $\Delta R/R_0$ signal part at T = 150 K. The dashed line was obtained by fitting the Lorentzian function with the parameters shown in the figure to the experimental curve.

on the oxygen *p* orbital [3]. Irrespective of the model, it is, however, believed that the conductivity of this system is polaron in character (that is, the charge carrier is “dressed” in a phonon or magnetic coat).

We begin the discussion with the temperature dependence of the differential reflection, which increases as the temperature lowers. An increase in $|\Delta R/R_0|$ is, as mentioned above, caused by a decrease in R_0, that is, in reflection at negative delay times t. A decrease in R_0 at low temperatures is determined by changes in the optical spectra. It is known that, in the dielectric phase, the optical conductivity spectra have a peak at ≈1.0 eV, which increases in intensity and shifts to the lower frequencies as the temperature decreases. This peak is also recorded in the ferromagnetic phase and transforms into a Drude response at helium temperatures [14, 15]. The temperature dependence of the optical conductivity was described using the double exchange model including the dynamic Jahn–Teller effect; it was attributed to changes in the Jahn–Teller polaron, which had a small radius at high temperatures and a large radius at low temperatures [14]. The appearance of the quasi-Drude response in the temperature range in which the resistance is fairly large was, however, explained in [15] by phase separation (according to this model, charge carriers experienced segregation in high-conductivity regions, which had the form of ferromagnetic drops in a dielectric matrix). Inelastic neutron scattering measurements gave experimental evidence for phase stratification in manganites; in these experiments, two spin excitation branches with quadratic and linear dispersion laws were recorded [16–18]. The first branch was associated with magnons in the ferromagnetic phase, and the second branch was assigned to magnetic excitations of the antiferromagnetic phase.

If, following [15], we relate the fast relaxation component to the relaxation of carriers in the ferromagnetic phase (which are characterized by high mobility and low concentration) and the slow component to carriers in the antiferromagnetic phase (low mobility and high concentration), then the fraction of the ferromagnetic phase should decrease as the temperature lowers because the B/A ratio increases. Such a behavior is physically meaningless, because the fraction of the ferromagnetic phase at $T < T_c$ cannot be smaller than at $T > T_c$.

One can attempt to relate the fast component of the decay of the excited state to the relaxation of photoexcited band carriers, whereas the slow component may be ascribed to the relaxation of the polaron coat, which is responsible for the large mass of polarons and consists of localized spins and/or phonons. With such an assignment of the components, a change in the temperature dependences of the components in the region of the transition temperature is, most likely, evidence of the crossover of the nature of the polarons. This can be the transition from large-sized to small-sized polarons or a bipolaron–polaron-type transition [3].

We emphasize that we do not assert that the two-component response to the action of ultranarrow laser pulses cannot be explained within the framework of the phase separation model. The optical method that we use is, however, an integral method, and the question of the characteristic size of nonuniformities and the topology of the nonuniform state therefore remains open. We can only note that the films were macroscopically homogeneous and their relaxation dynamics was not influenced by observation point displacements (the size of the probed region was 300 μm).

Note that several experimental and theoretical works give evidence of the thermally activated character of conductivity for optimally doped manganites in the paramagnetic phase [19, 20]. A detailed time-resolved study of manganites at temperatures $T \geq T_c$ can in principle be used to estimate the activation energy.

The fast photoinduced response oscillations occur at a frequency of 2.4 THz (=80 cm^{-1}). The low oscillation frequency favors the assignment of this phonon to either La-ion stretching motions or oxygen octahedron rotations (Mn ions make no contribution to even phonons, and the stretching and bending oxygen motions have much higher frequencies). Unfortunately, we cannot directly compare our results with Raman spectra because data on the Raman frequencies below 100 cm^{-1} are lacking [21]. In principle, an arbitrary Raman-active (even) phonon can be excited by femtosecond pulses provided the excitation mechanism is stimulated Raman scattering. If coherent phonons are generated by the shift mechanism, totally symmetrical modes can only be observed [16]. As the coherent phonon with a 2.4 THz frequency exists both above and below the transition temperature, it should be suggested that the transition occurs without a change in lattice symmetry. The fast photoinduced response oscillations are absent at temperatures $T > 200$ K because the decay of this mode is a strong function of temperature and the phonon is predamped in this temperature range. At the same time, there is no indication of transition-induced phonon frequency softening. Some increase in the frequency at low temperatures is most likely caused by lattice anharmonicity.

4. CONCLUSIONS

The temperature dependence of relaxation dynamics is evidence that two contributions with substantially different lifetimes exist in $La_{0.7}Ca_{0.3}MnO_3$ lanthanum manganites. The fast relaxation component with lifetimes of the order of hundreds of femtoseconds reaches a maximum in the region of the dielectric–metal transition, and its contribution increases as temperature decreases. The slow component with lifetimes of the order of dozens of picoseconds also has an extremum (minimum) in the transition region. The two-component character of relaxation, which includes contributions whose ratio changes in the transition, is evidence of a change in the character of the polaron responsible for conductivity in this material.

The fast photoinduced response oscillations assigned to coherent phonons exist both above and below the transition temperature. It follows that the structure of the lattice remains unchanged in the transition. However, note that, although the frequency of these oscillations changes insignificantly as temperature decreases, their damping (dephasing) considerably weakens.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic research (project no. 01-02-1640), a NATO grant for joint scientific research (PST.GLG.978935), and the Humboldt Foundation (Germany).

REFERENCES

1. J. M. D. Coey, M. Virtel, and S. von Molnar, Adv. Phys. **48**, 167 (1999).
2. É. L. Nagaev, Usp. Fiz. Nauk **166**, 577 (1996) [Phys. Usp. **39**, 539 (1996)].
3. A. S. Alexandrov and A. M. Bratkovsky, J. Phys.: Condens. Matter **11**, 1989 (1999).
4. Yu. A. Izyumov and Yu. N. Skryabin, Usp. Fiz. Nauk **171**, 121 (2001) [Phys. Usp. **44**, 109 (2001)].
5. M. Yu. Kagan and K. I. Kugel’, Usp. Fiz. Nauk **171**, 833 (2001).
6. J. Shah, *Ultrafast Spectroscopy of Semiconductors and Semiconducting Nanostructures* (Springer, Telos, 1998).
7. O. V. Misochko, Zh. Éksp. Teor. Fiz. **119**, 285 (2001) [JETP **92**, 246 (2001)].
8. J. Demsar, L. Forro, H. Berger, and D. Mihailovic, Phys. Rev. B **66**, 041101 (2002).
9. T. Ogasawara, M. Ashida, N. Motoyama, *et al.*, Phys. Rev. B **63**, 113105 (2001).
10. J. Demsar, K. Biljakovic, and D. Mihailovic, Phys. Rev. Lett. **83**, 800 (1999).
11. O. V. Misochko, N. Georgiev, T. Dekorsy, and M. Helm, Pis’ma Zh. Éksp. Teor. Fiz. **75**, 768 (2002) [JETP Lett. **75**, 642 (2002)].
12. V. I. Kudinov, A. I. Kirilyuk, N. M. Kreines, *et al.*, Phys. Lett. A **151**, 358 (1990).
13. E. Kaidashev, V. Dneprovski, D. Breus, and R. N. Sheftal, J. Supercond. **13**, 407 (2000).
14. M. Quijada, J. Cerne, J. R. Simpson, *et al.*, Phys. Rev. B **58**, 16093 (1998).
15. N. N. Loshkareva, Yu. P. Sukhorukov, V. E. Arkhipov, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 475 (1999) [Phys. Solid State **41**, 426 (1999)].
16. E. O. Wollan and W. C. Koehler, Phys. Rev. **100**, 545 (1955).
17. M. Hennion, F. Moussa, G. Biotteau, *et al.*, Phys. Rev. Lett. **81**, 1957 (1998).
18. Y. Moritomo, A. Machida, S. Mori, *et al.*, Phys. Rev. B **60**, 9220 (1999).
19. A. L. Rakhmanov, K. I. Kugel, Ya. M. Blanter, and M. Yu. Kagan, Phys. Rev. B **63**, 174424 (2001).
20. J. H. Zhao, H. P. Kunkel, X. Z. Zhou, and Gwyn Williams, Phys. Rev. B **66**, 184428 (2002).
21. L. Martin-Carron, A. De Andres, J. Martinez-Lope, *et al.*, Phys. Rev. B **66**, 174303 (2002).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 794–805.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 886–897.
Original English Text Copyright © 2003 by Falkovsky.

SOLIDS
Electronic Properties

Electron–Phonon Interaction and Coupled Phonon–Plasmon Modes[¶]

L. A. Falkovsky
Landau Institute for Theoretical Physics, Russian Academy of Sciences,
ul. Kosygina 2, Moscow, 117337 Russia
e-mail: falk@itp.ac.ru

Received March 24, 2003

Abstract—The theory of Raman scattering by the coupled electron–phonon system in metals and heavily doped semiconductors is developed with Coulomb screening and the electron–phonon deformation interaction taken into account. The Boltzmann equation for carriers is applied. Phonon frequencies and optic coupling constants are renormalized due to interactions with carriers. The k-dependent semiclassical dielectric function is employed instead of the Lindhard–Mermin expression. The results of calculations are presented for various values of the carrier concentration and the electron–phonon coupling constant. *© 2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Recently, there has been considerable interest in the effect of electron–phonon interactions on optical-phonon dispersion. This interest is stimulated by contradictions between different approaches to the electron–phonon interaction. The strong phonon renormalizations were first obtained by Migdal [1] (see also [2]) within a consistent many-body approach based on the Frölich Hamiltonian. An extremely large dispersion of optical phonons was predicted in [3] also using the Frölich model. These results contradict the Born–Oppenheimer (adiabatic) concept [4] according to which the phonon renormalizations should be small in terms of the nonadiabatic parameter $\sqrt{m/M}$, where m and M are the electron and ion masses, respectively (see also [5]). Theoretical investigations [6] of the sound velocity and acoustic attenuation in metals confirm the adiabatic concept. In a recent paper, Reizer [7] emphasized the importance of taking the screening effect into account. To our knowledge, the Coulomb screening effect on LO phonons was first studied in [8]. Using the Boltzmann equation, we found in [9] that the electron–phonon interaction results more significantly in optical-phonon damping than in the dispersion law. In any case, the Frölich model has evident shortcomings.

From the experimental standpoint, the best opportunity for the investigation of interactions between electrons and optical phonons is provided by coupled phonon–plasmon modes in doped semiconductors (see, e.g., [10]). Two such modes, $L^{\pm}$, have been observed in Raman experiments for many semiconductors. At the early stage, the Raman results were compared with the theory in [11] based on the Drude model (see, e.g., [12]), but the Lindhard–Mermin expression for the dielectric function has been used more recently [13].

The Lindhard–Mermin expression [14] represents a sophisticated generalization of the Lindhard function with the help of the electron relaxation time. The Lindhard approach is very useful, while the momentum transfer k in the Raman scattering is compared with the Fermi momentum p_F. The most significant effect of the carriers should be expected for $kv_F \sim \omega$, where v_F is the Fermi velocity and ω is the phonon frequency. For solids with metallic conductivity, the Fermi velocity can be estimated using the argument of stability under the Coulomb interaction $e^2/\pi\hbar v_F \le 1$. This condition gives $v_F \sim 0.7 \times 10^8$ cm/s. For the typical value of the optical phonon frequency $\omega = 500$ cm^{-1}, the values of interest are $k \le \omega/v_F \approx 10^6$ cm^{-1}. Therefore, the condition $k < p_F$ is satisfied for a carrier concentration larger than 3×10^{17} cm^{-3}. In experiments, heavily doped semiconductors with a large carrier concentration are used in order to obtain a visible carrier effect. The condition $k \ll p_F$ is then satisfied, and we can apply the Boltzmann equation in calculating the electronic susceptibility and in the evaluating the Raman cross section. The method of the Boltzmann equation is valid for the anisotropic electron plasma in solids at arbitrary temperatures. In the present paper, we obtain the Raman efficiency, applying the Boltzmann equation for generated carriers in heavily doped semiconductors at a temperature lower than the Fermi energy, $T \ll \varepsilon_F$.

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

2. EFFECTIVE HAMILTONIAN AND LIGHT SCATTERING

For the electron–phonon system in solids, we use the operator of particle numbers $\hat{n}$, the phonon displacements $\hat{b}_j$, and the macroscopic electric field E, which accompanies vibrations in polar semiconductors and acts on the electron and ion charges. The effective Hamiltonian describing inelastic light scattering in solids can be written in the semiclassical Wigner representation as

$$\mathcal{H} = \frac{e^2}{mc^2}\int d^3r\,\mathcal{N}(\mathbf{r}, t)U(\mathbf{r}, t), \qquad (1)$$

where

$$\mathcal{N}(\mathbf{r}, t) = \gamma\hat{n}(\mathbf{r}, t) + g_j\hat{b}_j(\mathbf{r}, t) + g_E E(\mathbf{r}, t) \qquad (2)$$

is a linear form in the variables $\hat{n}$, $\hat{b}_j$, and E. The subscript j denotes the various phonon modes, longitudinal (LO) or transverse (TO). More precisely, the subscript j designates the different phonon representations, which can be degenerate. The transformation properties of the coupling constants g_j are determined by this representation. The notation $U(\mathbf{r}, t)$ is introduced for a product of the vector potentials of the incident and scattered photons,

$$\begin{aligned} A^{(i)}(\mathbf{r}, t)A^{(s)}(\mathbf{r}, t) &= U(\mathbf{r}, t) \\ &= \exp[i(\mathbf{k}\cdot\mathbf{r} - \omega t)]U(\mathbf{k}, \omega), \end{aligned}$$

where the momentum and frequency transfers are $\mathbf{k} = \mathbf{k}^{(i)} - \mathbf{k}^{(s)}$ and $\omega = \omega^{(i)} - \omega^{(s)}$. The polarization vectors of $\hat{\mathbf{b}}_j(\mathbf{r}, t)$, $\mathbf{E}(\mathbf{r}, t)$, $\mathbf{A}^{(i)}(\mathbf{r}, t)$, and $\mathbf{A}^{(s)}(\mathbf{r}, t)$ are included in the coupling constants.

The first term in the right-hand side of Eq. (2) describes the light scattering by electron–hole pairs with the vertex

$$\gamma(\mathbf{p}) = e_\alpha^{(i)}e_\beta^{(s)}\left[\delta_{\alpha\beta} + \frac{1}{m}\right.$$
$$\left.\times\sum_n\left(\frac{p^\beta_{fn}p^\alpha_{nf}}{\epsilon_f(\mathbf{p}) - \epsilon_n(\mathbf{p}) + \omega^{(i)}} + \frac{p^\beta_{fn}p^\alpha_{nf}}{\epsilon_f(\mathbf{p}) - \epsilon_n(\mathbf{p}) + \omega^{(s)}}\right)\right],$$

where the resonant term is included; $e_\alpha^{(i)}$ and $e_\beta^{(s)}$ are the polarization vectors of the incident, $\mathbf{A}^{(i)}(\mathbf{r}, t)$, and scattered, $\mathbf{A}^{(s)}(\mathbf{r}, t)$, photons. The quantum-mechanical and statistical average of the first term in Eq. (2),

$$\langle\langle\gamma\hat{n}(\mathbf{r}, t)\rangle\rangle = \int\frac{2d^3p}{(2\pi)^3}\gamma(\mathbf{p})f_p(\mathbf{r}, t) \qquad (3)$$

can be expressed in terms of the electron distribution function $f_p(\mathbf{r}, t)$. The constants g_j and g_E are the deformation-optic and electrooptic couplings with the phonon displacements and the macroscopic electric field, respectively. The estimation gives $g_j \sim 1/a^4$, $g_E \sim 1/ea$, and $g(\mathbf{p}) \sim m/m^*$, where a is the lattice parameter and m^* is the effective mass.

The variable $U(\mathbf{r}, t)$ can be considered as an external force. The generalized susceptibility $\chi(\mathbf{k}, \omega)$ is then introduced as the linear response to this force,

$$\langle\langle\mathcal{N}(\mathbf{k}, \omega)\rangle\rangle = -\chi(\mathbf{k}, \omega)U(\mathbf{k}, \omega). \qquad (4)$$

According to the fluctuation–dissipation theorem, the function

$$K(\mathbf{k}, \omega) = \frac{2}{1 - e^{-\omega/T}}\mathrm{Im}\chi(\mathbf{k}, \omega)$$

is the Fourier component of the correlation function

$$K(\mathbf{r}, t; \mathbf{r}', t') = \langle\langle\mathcal{N}^\dagger(\mathbf{r}, t)\mathcal{N}(\mathbf{r}', t')\rangle\rangle, \qquad (5)$$

which depends only on the differences $\mathbf{r} - \mathbf{r}'$ and $t - t'$. The Raman cross section is given by

$$\frac{d\sigma}{d\omega^{(s)}d\Omega^{(s)}} = \frac{k_z^{(s)}\omega^{(s)}}{\pi c}\left(\frac{2e^2}{c\hbar m\omega^{(i)}}\right)^2 K(\mathbf{k}, \omega)|U(\mathbf{k}, \omega)|^2, \qquad (6)$$

where $k_z^{(s)}$ is the normal to the sample surface component of the scattered wave vector in vacuum.

A note should be made. Evidently, any sample has a surface. The surface effects in the Raman scattering were considered in our paper [15]; they are omitted in the derivation of Eq. (6). Because of the skin effect, moreover, the incident and scattered fields do not penetrate the bulk. For the optical range of the incident light, we have normal skin-effect conditions. We then integrate the distribution $|U(\mathbf{k}, \omega)|^2$ in Eq. (6) over the normal component k_z. As shown in [15], the integration of $|U(\mathbf{k}, \omega)|^2$ gives a factor of $1/\zeta_2$, where ζ_2 is expressed in terms of the wave-vector components inside the semiconductor, $\zeta_2 = \mathrm{Im}(k_z^{(i)} - k_z^{(s)})$. The Raman cross section (6) is dimensionless. It represents the ratio of the energy of inelastic scattered light to the incident energy.

3. BOLTZMANN EQUATION FOR CARRIERS

The problem of the evaluation of the Raman cross section consists in the calculation of generalized susceptibility (4). We apply the Boltzmann equation for the electron distribution function:

$$\frac{\partial f_p(\mathbf{r}, t)}{\partial t} + \mathbf{v}\frac{\partial f_p(\mathbf{r}, t)}{\partial \mathbf{r}} + \dot{\mathbf{p}}\frac{\partial f_p(\mathbf{r}, t)}{\partial \mathbf{p}} = -\frac{1}{\tau}[f_p(\mathbf{r}, t) - \langle f_p(\mathbf{r}, t)\rangle]. \tag{7}$$

The angular brackets denote the average over the Fermi surface,

$$\langle \ldots \rangle = \frac{1}{\nu_0}\int(\ldots)\frac{2dS_{\mathrm{F}}}{v(2\pi)^3},$$

where the integral is performed in the momentum space over the Fermi surface and ν_0 is the density of electron states, defined by the condition $\langle 1\rangle = 1$. We use the τ-approximation, which is correct for the electron scattering by defects in metals and heavily doped semiconductors. The collision integral in form (7) conserves the number of electrons in collisions. Therefore, the charge density satisfies the equation of continuity. This ensures the correct ω-dependence of the dielectric function at low frequencies.

In accordance with Eqs. (1), (2), and (3), instead of the unperturbed electron spectrum $\varepsilon_0(\mathbf{p})$, we introduce the local electron spectrum in the presence of the external force $U(\mathbf{r}, t)$ as

$$\varepsilon(\mathbf{p}, \mathbf{r}, t) = \varepsilon_0(\mathbf{p}) + \gamma(\mathbf{p})U(\mathbf{r}, t) + \zeta_j(\mathbf{p})b_j(\mathbf{r}, t),$$

where the last term represents the electron–optical-phonon deformation potential and $b_j(\mathbf{r}, t) = \langle\langle \hat{b}_j(\mathbf{r}, t)\rangle\rangle$. We use this form of the electron–phonon interaction instead of the Frölich polarization type $\zeta(\mathbf{p})\mathrm{div}\,\mathbf{b}(\mathbf{r}, t)$ because the first is larger by the parameter $1/ka$ for optical phonons.

We linearize Eq. (7), seeking its solution in the form

$$f_p(\mathbf{r}, t) = f_0[\varepsilon(\mathbf{p}, \mathbf{r}, t) - \mu] - \frac{df_0}{d\varepsilon}\delta f_p(\mathbf{r}, t), \tag{8}$$

where $f_0[\varepsilon(\mathbf{p}, \mathbf{r}, t) - \mu]$ is the Fermi–Dirac local distribution function. It is important that the collision term in the Boltzmann equation is canceled by the local-equilibrium term in Eq. (8).

We impose the number conservation condition on the chemical potential,

$$\int\frac{d^3p}{(2\pi)^3}f_0[\varepsilon(\mathbf{p}, \mathbf{r}, t) - \mu] = \int\frac{d^3p}{(2\pi)^3}f_0(\varepsilon_0 - \mu_0),$$

and obtain

$$\mu = \mu_0 + \langle\gamma(\mathbf{p})\rangle U(\mathbf{r}, t) + \langle\zeta_j(\mathbf{p})\rangle b_j(\mathbf{r}, t).$$

This condition implies the renormalization of vertices

$$\begin{aligned}\gamma(\mathbf{p}) &\longrightarrow \gamma(\mathbf{p}) - \langle\gamma(\mathbf{p})\rangle,\\ \zeta_j(\mathbf{p}) &\longrightarrow \zeta_j(\mathbf{p}) - \langle\zeta_j(\mathbf{p})\rangle,\end{aligned} \tag{9}$$

and this substitution is to be made in what follows.

The linearized Boltzmann equation in the Fourier components is given by

$$-i(\omega - \mathbf{k}\cdot\mathbf{v} + i/\tau)\delta f_p(\mathbf{k}, \omega) = \psi_p(\mathbf{k}, \omega) + \langle\delta f_p(\mathbf{k}, \omega)\rangle/\tau,$$

where

$$\psi_p(\mathbf{k}, \omega) = e\mathbf{v}\cdot\mathbf{E}(\mathbf{k}, \omega) - i\omega[\gamma(\mathbf{p})U(\mathbf{k}, \omega) + \zeta_j(\mathbf{p})b_j(\mathbf{k}, \omega)].$$

The solution to this equation is easily obtained as

$$\delta f_p(\mathbf{k}, \omega) = i[\psi_p(\mathbf{k}, \omega) + \langle\delta f_p(\mathbf{k}, \omega)\rangle/\tau]/\Delta_p, \tag{10}$$

where we designate $\Delta_p = \omega - \mathbf{k}\cdot\mathbf{v} + i/\tau$. We now obtain

$$\langle\delta f_p(\mathbf{k}, \omega)\rangle = \frac{i\langle\psi_p(\mathbf{k}, \omega)/\Delta_p\rangle}{1 - i\langle\tau^{-1}/\Delta_p\rangle}. \tag{11}$$

Notice that in accordance with the adiabatic concept, no additional contribution comes from the local equilibrium distribution function $f_0[\varepsilon(\mathbf{p}, \mathbf{r}, t)]$ in Eq. (8).

4. EQUATION OF MOTION FOR PHONONS INTERACTING WITH CARRIERS

In the long-wave approximation ($k \ll 1/a$, where a is the lattice parameter), we write the equation of motion for the phonon displacement field as

$$(\omega_k^2 - \omega^2)b_j(\mathbf{k}, \omega) = \frac{Z}{M'}E_j(\mathbf{k}, \omega) - \frac{g_jU(\mathbf{k}, \omega)}{M'N} - \frac{1}{M'N}\int\frac{2dS_F}{v(2\pi)^3}\zeta_j(\mathbf{p})\delta f_p(\mathbf{k}, \omega), \tag{12}$$

where N is the number of unit cells in 1 cm^3, M' is the reduced mass of the unit cell, and Z is the effective ionic charge. The nonperturbed phonon frequency ω_k must be considered in the absence of the electric field and without any electron–phonon interactions. In the long-wave limit, we can expand it as $\omega_k^2 = \omega_0^2 \pm s^2k^2$ with a value of the dispersion parameter s on the order of the typical sound velocity in solids. We note that the optical phonons always have the so-called natural width $\Gamma^{\mathrm{nat}} \sim \omega_0\sqrt{m/M}$. The natural width results from decay processes into two or more acoustic and optical phonons.

In the final expressions, we will substitute $\omega_k^2 - \omega^2 \longrightarrow \omega_k^2 - i\omega\Gamma^{\mathrm{nat}} - \omega^2$.

Equation (12) is applied to both the longitudinal and transverse phonons. It follows from the Maxwell equations that the electric field is longitudinal, $\mathbf{E} \parallel \mathbf{k}$, in the optical region $k \gg \omega/c$. If the excited phonons propagate in the symmetric direction, the TO and LO phonons are separated. Therefore, the electric field is involved in only Eq. (12) for the LO phonon. In addition, the coupling $\zeta_j(\mathbf{p})$ depends on the phonon representation j.

Using solution (10), we rewrite Eq. (12) as

$$(\tilde{\omega}_j^2 - \omega^2) b_j(\mathbf{k}, \omega) - \frac{\tilde{Z}}{M'} E_j(\mathbf{k}, \omega) = -\frac{\tilde{g}_j U(\mathbf{k}, \omega)}{M'N}, \quad (13)$$

where the phonon frequency

$$\tilde{\omega}_j^2 = \omega_k^2 + \frac{\omega\nu_0}{M'N}\left(\left\langle \frac{\zeta_j^2(\mathbf{p})}{\Delta_p} \right\rangle + \frac{i\langle \zeta_j(\mathbf{p})/\Delta_p \rangle^2}{\tau - \langle i/\Delta_p \rangle}\right), \quad (14)$$

the effective ionic charge

$$\tilde{Z} = Z - \frac{ie\nu_0}{N} \times \left(\left\langle \frac{v_z \zeta_j(\mathbf{p})}{\Delta_p} \right\rangle + \frac{i\langle v_z/\Delta_p \rangle \langle \zeta_j(\mathbf{p})/\Delta_p \rangle}{\tau - \langle i/\Delta_p \rangle}\right), \quad (15)$$

and the deformation-optic coupling

$$\tilde{g}_j = g_j + \omega\nu_0\left(\left\langle \frac{\zeta_j(\mathbf{p})\gamma(\mathbf{p})}{\Delta_p} \right\rangle + \frac{i\langle \zeta_j(\mathbf{p})/\Delta_p \rangle \langle \gamma(\mathbf{p})/\Delta_p \rangle}{\tau - \langle i/\Delta_p \rangle}\right) \quad (16)$$

are renormalized because of the electron–phonon interaction $\zeta_j(\mathbf{p})$.

5. POISSON EQUATION FOR THE MACROSCOPIC FIELD

We consider the longitudinal electric induction D that accompanies lattice vibrations. There are several contributions to the field: (1) the polarization $\alpha E(\mathbf{r}, t)$ of the filled electron bands; (2) the lattice polarization $NZb_{\mathrm{LO}}(\mathbf{r}, t)$; (3) the contribution of the free carrier density $\rho = -\mathrm{div}\mathbf{P}_e$; and (4) the term $P = -\partial\mathcal{H}/\partial E = -g_E U$ that explicitly results from Hamiltonian (1), (2). Collecting all these terms into the Poisson equation $\mathrm{div}D = 0$, we find

$$\varepsilon_\infty E(\mathbf{k}, \omega) + 4\pi NZb_{\mathrm{LO}}(\mathbf{k}, \omega) + \frac{4\pi ie}{k}\int \frac{2d^3p}{(2\pi)^3}\delta f_p(\mathbf{k}, \omega) - 4\pi g_E U(\mathbf{k}, \omega) = 0, \quad (17)$$

where the high-frequency permittivity $\varepsilon_\infty = 1 + 4\pi\alpha$. Using the solution of the Boltzmann equation, we rewrite the Poisson equation in the form

$$\varepsilon_e(\mathbf{k}, \omega)E(\mathbf{k}, \omega) + 4\pi N\bar{Z}b_{\mathrm{LO}}(\mathbf{k}, \omega) = 4\pi\tilde{g}_E U(\mathbf{k}, \omega), \quad (18)$$

where the electronic dielectric function

$$\varepsilon_e(\mathbf{k}, \omega) = \varepsilon_\infty + \varepsilon_\infty \frac{k_0^2}{k^2}\left[1 - \frac{\langle \omega/\Delta_p(k) \rangle}{\langle i/\Delta_p(k) \rangle/\tau}\right] \quad (19)$$

contains the Thomas–Fermi parameter $k_0^2 = 4\pi e^2\nu_0/\varepsilon_\infty$.

Because of the electron–phonon interactions $\zeta_{LO}(\mathbf{p})$, the ionic charge obtains an additional term,

$$\bar{Z} = Z + \frac{ie\nu_0}{N} \times \left(\left\langle \frac{v_z \zeta_{LO}(\mathbf{p})}{\Delta_p} \right\rangle + \frac{i\langle v_z/\Delta_p \rangle \langle \zeta_{LO}(\mathbf{p})/\Delta_p \rangle}{\tau - \langle i/\Delta_p \rangle}\right) \quad (20)$$

of the opposite sign compared with that in Eq. (15). The electrooptic coupling in Eq. (17) also changes, but because of the light scattering by carriers $\gamma(\mathbf{p})$:

$$\tilde{g}_E = g_E - ie\nu_0\left(\left\langle \frac{v_z \gamma(\mathbf{p})}{\Delta_p} \right\rangle + \frac{i\langle v_z/\Delta_p \rangle \langle \gamma(\mathbf{p})/\Delta_p \rangle}{\tau - \langle i/\Delta_p \rangle}\right). \quad (21)$$

6. RAMAN SCATTERING BY ELECTRON–HOLE PAIRS, PHONONS, AND COUPLED MODES

We are now in a position to calculate susceptibility (4). Using Eqs. (3), (10), (11), (16), and (21), we obtain

$$\langle\langle \mathcal{N}(\mathbf{k}, \omega) \rangle\rangle = -\chi_e(\mathbf{k}, \omega)U(\mathbf{k}, \omega) + \tilde{g}_j b_j(\mathbf{k}, \omega) + \bar{g}_E E(\mathbf{k}, \omega), \quad (22)$$

where

$$\chi_e(\mathbf{k}, \omega) = -\omega\nu_0\left(\left\langle \frac{\gamma^2(\mathbf{p})}{\Delta_p} \right\rangle + \frac{i\langle \gamma(\mathbf{p})/\Delta_p \rangle^2}{\tau - \langle i/\Delta_p \rangle}\right)$$

gives the light scattering with the excitation of electron–hole pairs. We note that the renormalized coupling $\bar{g}_E$ entering here differs from $\tilde{g}_E$ in (21) by the sign of the second term:

$$\bar{g}_E = g_E + ie\nu_0\left(\left\langle \frac{v_z \gamma(\mathbf{p})}{\Delta_p} \right\rangle + \frac{i\langle v_z/\Delta_p \rangle \langle \gamma(\mathbf{p})/\Delta_p \rangle}{\tau - \langle i/\Delta_p \rangle}\right).$$

To find $E(\mathbf{k}, \omega)$ and $b_j(\mathbf{k}, \omega)$, we must solve the system of algebraic equations (13) and (18). Using Eq. (22), we then obtain the generalized susceptibility

$$\chi(\mathbf{k}, \omega) = \chi_e(\mathbf{k}, \omega) + \frac{\tilde{g}_j^2 \varepsilon_e(\mathbf{k}, \omega)/NM' - 4\pi \tilde{g}_E \bar{g}_E (\tilde{\omega}_j^2 - \omega^2) - 4\pi \tilde{g}_j (\tilde{g}_E \tilde{Z} + \bar{g}_E \bar{Z})/M}{(\tilde{\omega}_j^2 - \omega^2)\varepsilon_e(\mathbf{k}, \omega) + 4\pi N \tilde{Z} \bar{Z}/M'}. \tag{23}$$

Expression (23) is our main result. The poles of the second term give the spectrum of collective excitations of the electron–phonon system. We discuss Eq. (23) in various limiting cases.

6.1. The Electronic Scattering

We obtain the Raman electronic scattering from Eq. (23) if we set $\tilde{g}_j = g_E = \tilde{Z} = \bar{Z} = 0$. We then have

$$\chi(\mathbf{k}, \omega) = \chi_e(\mathbf{k}, \omega) + \frac{4\pi \tilde{g}_E^2}{\varepsilon_e(\mathbf{k}, \omega)}, \tag{24}$$

where $\tilde{g}_E$ is given by Eq. (21) with $g_E = 0$.

For the isotropic Fermi surface, we calculate the dielectric function in Eq. (19) by performing integration:

$$\left\langle \frac{1}{\Delta_p(k)} \right\rangle = \frac{1}{2kv_F} \ln\frac{1+\kappa}{1-\kappa}, \quad \kappa = \frac{kv_F}{\omega + i\tau^{-1}}, \tag{25}$$

where we must take the branch of ln$x$ that is real for positive real values of x.

For the anisotropic Fermi surface, the calculations can be performed in limiting cases. For $|\kappa| \gg 1$, we use the expansion for electronic dielectric function (19),

$$\varepsilon_e(\mathbf{k}, \omega) = \varepsilon_\infty \left\{ 1 + \left(\frac{k_0}{k}\right)^2 \left[1 + i\frac{\pi\nu_0\omega}{k} \left\langle \frac{1}{v}\delta(\mu) \right\rangle \right] \right\}, \tag{26}$$

where $\mu = \mathbf{v} \cdot \mathbf{k}/vk$ and $\delta(x)$ is the Dirac delta function. In this case, the Raman efficiency has a "tail" due to the Landau damping,

$$\mathrm{Im}\chi(\mathbf{k}, \omega) = \frac{\pi\nu_0\omega}{k} \left\langle \frac{\gamma^2(\mathbf{p})}{v} \delta(\mu) \right\rangle. \tag{27}$$

We see that the Raman cross section vanishes for the isotropic vertex $\gamma(\mathbf{p})$ because of Eq. (9). This is a result of Coulomb screening. It was first obtained in [16] for Raman scattering in semiconductors (see [17]).

In the opposite case, where $|\kappa| \ll 1$, the first term in Eq. (24) gives the result

$$\mathrm{Im}\chi_e(\mathbf{k}, \omega) = \nu_0 \langle \gamma^2(\mathbf{p}) \rangle \frac{\omega\tau}{(\omega\tau)^2 + 1}, \tag{28}$$

which was first found in [18] with the help of the Green-function technique. The second term in Eq. (24) reveals a plasmon pole at small values of k. The k-expansion of the dielectric function is

$$\varepsilon_e(k, \omega) = \varepsilon_\infty \left(1 - \frac{\omega_{pe}^2 + k^2 w}{\omega(\omega + i\tau^{-1})} \right), \tag{29}$$

where the k-independent term represents the Drude conductivity and the electron plasma frequency is given by the integral over the Fermi surface, $\omega_{pe}^2 = k_0^2 \langle v_z^2 \rangle$. The complex coefficient

$$w = \frac{k_0^2(\langle v_z^4 \rangle + i\langle v_z^2 \rangle^2/\omega\tau)}{(\omega + i\tau^{-1})^2}.$$

For the quadratic electron spectrum $\langle v_z^2 \rangle = \langle v_F^2 \rangle/3$ and $\langle v_z^4 \rangle = \langle v_F^4 \rangle/5$.

The k-expansion of $\tilde{g}_E$ gives

$$\tilde{g}_E = -\frac{ie\nu_0 k \langle \gamma(\mathbf{p}) v_z^2 \rangle}{(\omega + i\tau^{-1})^2},$$

because $g_E = 0$ and the zero-order term in the k-expansion vanishes due to the time invariance $\mathbf{v} \longrightarrow -\mathbf{v}$. The intensity of the plasmon peak is then proportional to k^2, in accordance with the known behavior of the dynamical structure factor.

6.2. Raman Scattering by TO Phonons

The second term in Eq. (23) gives the TO phonon scattering if we set $\tilde{Z} = \bar{Z} = \tilde{g}_E = 0$:

$$\chi(\mathbf{k}, \omega) = \frac{\tilde{g}_{TO}^2/NM'}{\tilde{\omega}_{TO}^2 - \omega^2 - i\omega\Gamma^{nat}}, \tag{30}$$

where $\tilde{\omega}_{TO}$ and $\tilde{g}_{TO}$ are defined in Eqs. (14) and (16) with $\zeta_j(\mathbf{p}) = \zeta_{TO}(\mathbf{p})$; we add the phonon width Γ^{nat} mentioned above.

Two points must be noted here. First, the TO-resonance occurs at the renormalized frequency $\tilde{\omega}_{TO}$. Taking the real and imaginary parts of (14), we obtain the TO phonon shift and width due to the deformation interaction $\zeta(\mathbf{p})$ with carriers:

$$\Delta\omega_k = \mathrm{Re}(\tilde{\omega}_{TO}^2 - \omega_k^2)/2\omega_k, \quad \Gamma = \Gamma^{nat} - \mathrm{Im}\tilde{\omega}_{TO}^2/\omega_k.$$

Second, because of the interaction with carriers, the coupling $\tilde{g}_{\rm TO}$ in (16) has an imaginary part. Therefore, the line shape of the resonance becomes asymmetric (the Fano resonance),

$$\mathrm{Im}\chi(\mathbf{k}, \omega) = \frac{1}{NM'} \times \frac{\omega\Gamma g_{\rm TO}^2 + [(\mathrm{Re}\tilde{\omega}_{\rm TO})^2 - \omega^2]\mathrm{Im}\tilde{g}_{\rm TO}^2}{[(\mathrm{Re}\tilde{\omega}_{\rm TO})^2 - \omega^2]^2 + (\omega\Gamma)^2}. \tag{31}$$

The line-shape asymmetry depends on the sign of Re$g_{\rm TO}$. For instance, if Re$g_{\rm TO} > 0$, the high-frequency wing of the resonance line drops more slowly than the low-frequency one. In the limiting case where $\kappa \gg 1$, we expand

$$\tilde{g}_{\rm TO} = g_{\rm TO} + \frac{\nu_0\omega}{k}\left\langle \frac{\gamma(\mathbf{p})\zeta_{\rm TO}(\mathbf{p})}{v}\left(-i\pi + \frac{2\omega}{kv}\right)\delta(\mu)\right\rangle, \tag{32}$$

and for $\kappa \ll 1$, we have

$$\tilde{g}_{\rm TO} = g_{\rm TO} + \frac{\omega\nu_0}{\omega + i\tau^{-1}} \times \left(\langle\gamma(\mathbf{p})\zeta_{\rm TO}(\mathbf{p})\rangle + \frac{k^2\langle v_z^2\gamma(\mathbf{p})\zeta_{\rm TO}(\mathbf{p})\rangle}{(\omega + i\tau^{-1})^2}\right). \tag{33}$$

Notice that the electron–phonon interaction $\zeta_{\rm TO}(\mathbf{p})$ and the light scattering $\gamma(\mathbf{p})$ by carriers jointly renormalize the coupling $g_{\rm TO}$ The frequency renormalization $\tilde{\omega}_{\rm TO}^2$ (see Eq. (14)) results only from the electron–phonon interaction $\zeta_{\rm TO}(\mathbf{p})$. The corresponding expressions can be obtained from Eqs. (32) and (33) by the substitution $\gamma(\mathbf{p}) \longrightarrow \zeta_{\rm TO}(\mathbf{p})$. We see that the TO phonons become broader and harder because of the interaction with carriers.

We emphasize that the phonon renormalizations depend on the carrier density ν_0 and the average coupling $\zeta_j(\mathbf{p}) - \langle\zeta_j(\mathbf{p})\rangle$. They vanish for the isotropic Fermi surface. The maximum value of the relative renormalization in on the order of $\lambda a p_{\rm F} m^*\omega/m|\omega + i\tau^{-1}|$ at $kv \sim |\omega + i\tau^{-1}|$, where λ is the dimensionless electron–phonon coupling and m^* is the effective electron mass.

6.3. Raman Scattering by LO-Phonon–plasmon Coupled Modes

In this case, the carriers interact with each other and the ion vibrations via both the macroscopic electric field $E(\mathbf{r}, t)$ and the deformation potential $\zeta_{\rm LO}(\mathbf{p})$. In the long-wave limit $k \longrightarrow 0$, Eqs. (15), (20), and (21) show no renormalization of the ionic charge, $\tilde{Z} = \bar{Z} = Z$, and of the electrooptic constant, $\tilde{g}_E = g_E$. Equation (23) then becomes

$$\chi(0, \omega) = \chi_e(0, \omega) + \frac{\tilde{g}_{LO}^2\varepsilon_e(0, \omega)/NM' - 4\pi g_E^2(\tilde{\omega}_0^2 - i\omega\Gamma^{\rm nat} - \omega^2) - 8\pi g_E\tilde{g}_{\rm LO}Z/M'}{(\tilde{\omega}_0^2 - i\omega\Gamma^{\rm nat} - \omega^2)\varepsilon_e(0, \omega) + 4\pi NZ^2/M'}, \tag{34}$$

where the first term is given in Eq. (28). The deformation potential $\zeta_{\rm TO}(\mathbf{p})$ renormalizes the phonon frequency $\tilde{\omega}_0$ in (14), as well as the deformation-optic constant $\tilde{g}_{\rm LO}$ in (16). The corresponding expansion in the limiting cases are similar to Eqs. (32) and (33). All of the above-mentioned about the TO line asymmetry also applies to the LO line.

Because the dielectric function of the electron–ion system is given by

$$\varepsilon(0, \omega) = \varepsilon_e(0, \omega) + \frac{4\pi NZ^2}{M'(\tilde{\omega}_0^2 - i\omega\Gamma^{\rm nat} - \omega^2)}, \tag{35}$$

the second term in the right-hand side of Eq. (34) has poles at the points where $\varepsilon(0, \omega) = 0$. This condition defines the frequency of coupled phonon–plasmon modes in the long-wave limit.

In the absence of electron and phonon collisions ($\tau^{-1} = \Gamma^{\rm nat} = 0$), and without electron–phonon interaction ($\zeta(\mathbf{p}) = 0$), we obtain a biquadratic equation from Eq. (29). It gives the frequencies of the coupled phonon–plasmon modes at $k = 0$,

$$\omega_\pm^2 = \frac{1}{2}(\omega_{pe}^2 + \omega_{\rm LO}^2) \pm \frac{1}{2}[(\omega_{pe}^2 + \omega_{\rm LO}^2)^2 - 4\omega_{pe}^2\omega_{\rm TO}^2]^{1/2}, \tag{36}$$

where $\omega_{\rm TO} = \omega_k$ is the TO-mode frequency at $k = 0$, $\omega_{\rm LO}^2 = \omega_{\rm TO}^2 + \omega_{pi}^2$, and $\omega_{pi}^2 = 4\pi NZ^2/\varepsilon_\infty M'$. These frequencies (related to $\omega_{\rm TO}$) are shown in Fig. 1 as functions of the electron concentration, namely, $\omega_{pe}/\omega_{\rm TO}$. The upper line begins at $\omega_{\rm LO}$ and tends to the electron plasma frequency ω_{pe}. The lower frequency starts as $\omega_{pe}\omega_{\rm TO}/\omega_{\rm LO}$ and then approaches $\omega_{\rm TO}$. In other words,

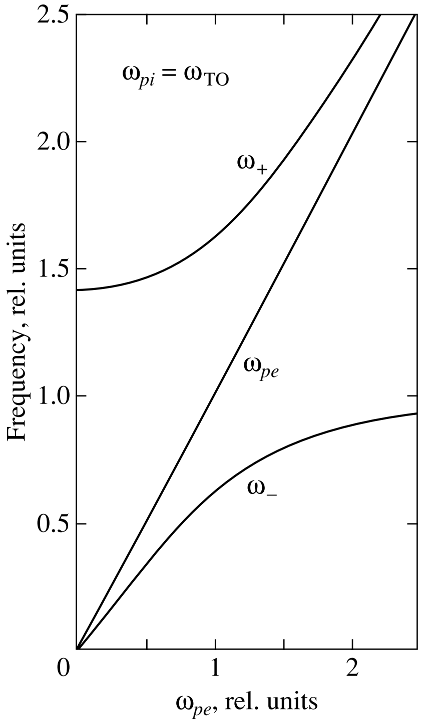

Fig. 1. Frequencies (in units of ω_{TO}) of the phonon–plasmon modes at $k = 0$ versus the free-carrier concentration, namely, the electron plasma frequency (in units of ω_{TO}). We set the ion plasma frequency $\omega_{pi} = \omega_{\text{TO}}$ in the absence of the free carriers. Then $\omega_{\text{LO}}/\omega_{\text{TO}} = \sqrt{2}$.

observing the longitudinal phonon mode in the optic range and adding electrons, we see a transition of the longitudinal phonon frequency from ω_{LO} to ω_{TO}. This is a result of the Coulomb screening.

We can compare Eq. (34) with the theory of Hon and Faust [11]. Because the electron–phonon interaction, $\zeta_{\text{LO}}(\mathbf{p})$, as well as the electronic scattering, $\gamma(\mathbf{p})$, were ignored in their theory, the phonon frequency and the deformation-optic constant were not renormalized. Equation (34) can then be rewritten as

$$\chi(0,\omega) = \frac{(4\pi g_E)^2}{\varepsilon_\infty\varepsilon(0,\omega)} \times\left[\frac{\varepsilon_e(0,\omega)A^2\chi_I}{\varepsilon_\infty} - \frac{\varepsilon_\infty}{4\pi} - 2A\chi_I\right], \tag{37}$$

where

$$\chi_I = NZ^2/M'(\omega_{\text{TO}}^2 - i\omega\Gamma^{\text{nat}} - \omega^2),$$

$$A = \frac{C\omega_{\text{TO}}^2 M'\varepsilon_\infty}{4\pi NZ^2}, \quad C = \frac{g_{\text{LO}}Z}{g_E M\omega_{\text{TO}}^2}$$

is the Faust–Henry coefficient. We now see that expression (37) coincides with the result of Hon and Faust (see, e.g. [13, Eq. (3.1)]).

For $k \neq 0$, Eq. (23) includes the dielectric function (19), which differs from the Lindhard–Mermin expression. The condition

$$(\tilde{\omega}_j^2 - i\omega\Gamma^{\text{nat}} - \omega^2)\varepsilon_e(\mathbf{k},\omega) + 4\pi N\tilde{Z}\bar{Z}/M' = 0 \tag{38}$$

determines the frequencies and damping of the phonon–plasmon coupled modes and contains the phonon frequency $\tilde{\omega}$ and the ionic charge renormalized by the electron–photon interaction $\zeta_{\text{LO}}(\mathbf{p})$. The electrooptic coupling g_E in Eq. (21) is modified because of the light scattering $\gamma(\mathbf{p})$ from the electron–hole pairs. This effect is not canceled in the product $\tilde{g}_E\bar{g}_E$ in Eq. (23) even in the absence of the electron–phonon interaction $\zeta_{\text{LO}}(\mathbf{p})$. The expansion of $\tilde{g}_E$ has the form

$$\tilde{g}_E = g_E - \frac{ie\nu_0 k\langle v_z^2\gamma(\mathbf{p})\rangle}{(\omega + i\tau^{-1})^2} \tag{39}$$

for $|\kappa| \ll 1$ and

$$\tilde{g}_E = g_E + e\nu_0 k^{-2}(\omega + i\tau^{-1}) \times\left\langle\frac{\gamma(\mathbf{p})}{v}\left(-\frac{\pi}{2} - \frac{i\omega}{kv}\right)\delta(\mu)\right\rangle \tag{40}$$

for $|\kappa| \gg 1$. We note that the term $\tilde{g}_E\bar{g}_E$ has the largest imaginary part for $\omega\tau \approx 1$ and then results most significantly in the line-shape asymmetry.

The dispersion of the phonon–plasmon modes is schematically shown in Fig. 2. There are two main peculiarities in this figure. First, the behavior of the upper mode near the line $\omega = kv_{\text{F}}$. Around this line $\tau^{-1} < \omega - kv_{\text{F}} \ll kv_{\text{F}}$, dielectric function (19) has a singularity,

$$\varepsilon_e(k,\omega) = \varepsilon_\infty + \varepsilon_\infty\frac{k_0^2}{k^2} \times\left\{1 - \frac{\omega}{2kv_{\text{F}}}\left[\frac{1}{2}\ln\frac{4k^2v_{\text{F}}^2}{(\omega - kv_{\text{F}})^2 + \tau^{-2}} - i\frac{\tau^{-1}}{\omega - kv_{\text{F}}}\right]\right\}. \tag{41}$$

Because of this singularity, the upper mode approaches the asymptote $\omega = kv_{\text{F}}$ as the wave vector k increases. Second, in the region $kv_{\text{F}} \gg \omega$, there is always one mode that has a predominantly phonon character. The

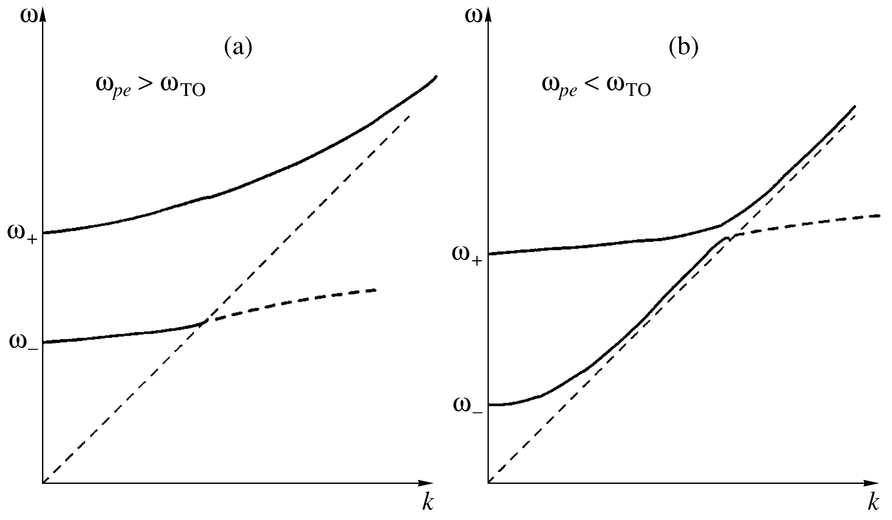

Fig. 2. Schematic representation of the dispersion of phonon–plasmon modes for (a) metallic ($\omega_{pe} > \omega_{TO}$) and (b) semiconducting ($\omega_{pe} < \omega_{TO}$) carrier concentrations. The dashed straight lines separate the domain $kv_F > \omega$ where the Landau damping exists the dashed curves represent damped modes there.

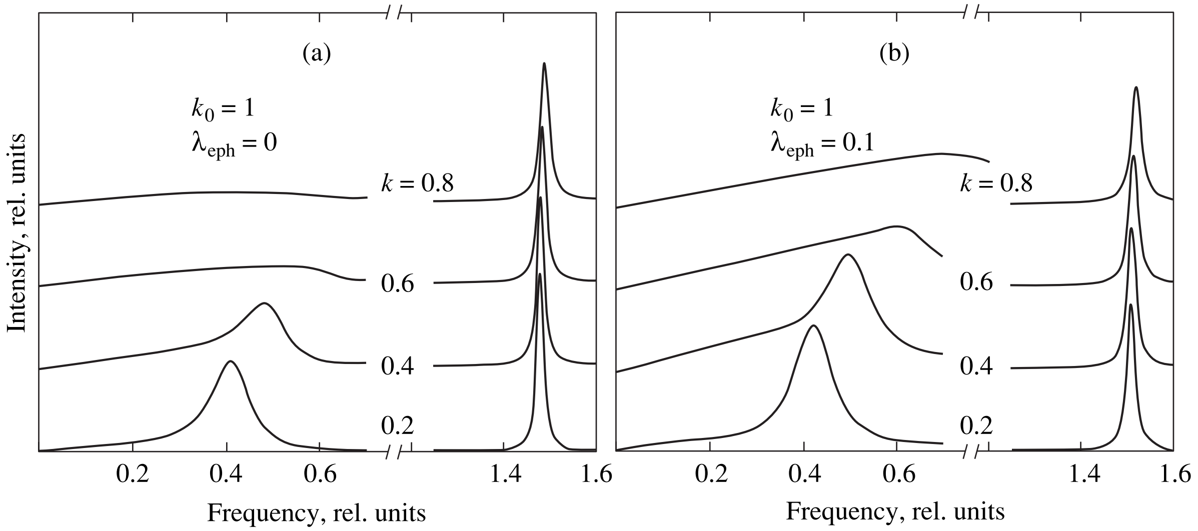

Fig. 3. Raman spectra from a semiconductor with low carrier concentration as a function of the frequency transfer ω for the indicated values of the momentum transfer *k*, the Thomas–Fermi parameter k_0 (in units of ω_{TO}/v_F), and the electron–phonon coupling constant $\lambda_{eph} = 0$ (a) and 0.1 (b). We set the ion plasma frequency $\omega_{pi} = \omega_{TO}$, the phonon natural width $\Gamma^{nat}/\omega_{TO} = 10^{-2}$, and the carrier relaxation rate $\tau^{-1}/\Gamma^{nat} = 10$.

reason is the decrease with *k* of the imaginary part of dielectric function (26).

7. DISCUSSION

We now consider the results in the simplest way. We assume that the electronic scattering is negligibly small, $\gamma(\mathbf{p}) = 0$. The second term in the parentheses in Eq. (14) is less than the first one in both limiting cases, $\kappa \ll 1$ and $\kappa \gg 1$. We neglect this term entirely. We also do not take the ion charge renormalization into account because it vanishes at small values of κ. We can then use expression (25) not only for the dielectric function $\varepsilon_e(\mathbf{k}, \omega)$, Eq. (19), but also for the renormalized phonon frequency $\tilde{\omega}$, Eq. (14).

In such an approximation, solving Eq. (38), we find the frequency and damping of the photonlike mode for the limiting cases of the parameter $\kappa = kv_F/(\omega + i/\tau)$ and for low and large carrier concentration.

1. Low carrier concentration, $\omega_{pe} < \omega_O$: $|\kappa| \ll 1$,

$$\omega^2 = \omega_{LO}^2 - i\omega\Gamma^{nat}$$

$$+ \frac{\lambda\omega\omega_{TO}^2}{\omega^*}\left(1 + \frac{\kappa^2}{3}\right) + \frac{(\omega_{pi}\omega_{pe})^2}{\omega\omega^*}\left(1 + \frac{i\kappa^2}{3\omega\tau}\right),$$

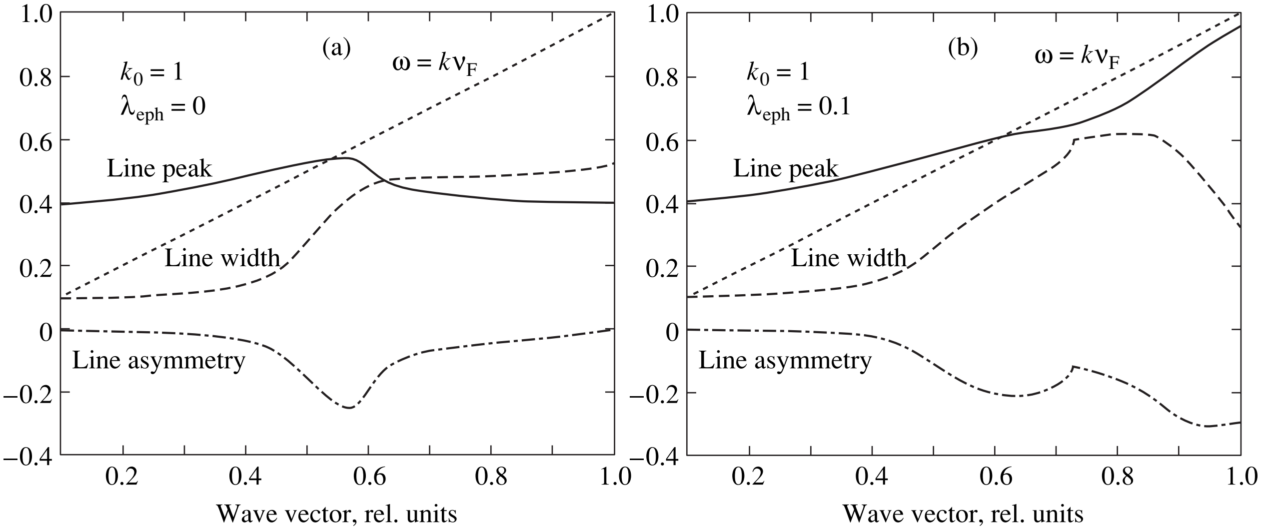

Fig. 4. The plasmon dispersion as a function of k in units of ω_{TO}/v_F (the position in units of ω_{TO} of the line peak of Raman spectra; upper part of the figure, solid line) for $k_0 = 1$ and $\lambda_{eph} = 0$ (a) and 0.1 (b). In the bottom, the line width (the full width at half-maximum, dashed line), and the line asymmetry (the difference between the right and left wings at half-maximum, dash–dotted line) in units of ω_{TO}. The Landau damping exists to the right of the dotted line $\omega = kv_F$.

$|\kappa| \gg 1$,

$$\omega^2 = \omega_{LO}^2 - i\omega\Gamma^{nat} + \frac{\lambda\omega\omega_{TO}^2}{kv_F}\left(-i\frac{\pi}{2}+\frac{1}{\kappa}\right) - \frac{3(\omega_{pi}\omega_{pe})^2}{(kv_F)^2}\left(1+\frac{i\pi\omega}{2kv_F}\right),$$

where $\omega^* = \omega + i/\tau$ and, instead of ω, we substitute $\omega \longrightarrow \omega_{LO}$.

2. Large carrier concentration, $\omega_{pe} \gg \omega_{TO}$: $|\kappa| \ll 1$,

$$\omega^2 = \omega_{TO}^2 - i\omega\Gamma^{nat} + \frac{\lambda\omega\omega_{TO}^2}{\omega^*}\left(1+\frac{\kappa^2}{3}\right) - \frac{\omega_{pi}^2\omega\omega^*}{\omega_{pe}^2}\left(1-\frac{i\kappa^2}{3\omega\tau}\right),$$

$|\kappa| \gg 1$,

$$\omega^2 = \omega_{TO}^2 - i\omega\Gamma^{nat} + \frac{\lambda\omega\omega_{TO}^2}{kv_F}\left(-i\frac{\pi}{2}+\frac{1}{\kappa}\right) + \frac{(\omega_{pi}kv_F)^2}{3\omega_{pe}^2}\left(1-\frac{i\pi\omega}{2kv_F}\right),$$

where we substitute $\omega \longrightarrow \omega_{TO}$. The definition of λ depends on κ,

$$\lambda = \frac{\nu_0\langle\zeta^2(\mathbf{p})\rangle}{M'N\omega_{TO}^2}, \quad \kappa \ll 1,$$

$$\lambda = \frac{\nu_0 v_F^2\langle\zeta^2(\mathbf{p})/v_z^2\rangle}{M'N\omega_{TO}^2}, \quad \kappa \gg 1,$$

but gives a value on the same order, $\lambda \sim p_F am^*/m$.

Results of the numerical calculations of the Raman spectra, Eq. (23), in this approximation are shown in Fig. 3 for two values of electron–phonon coupling:

$$\lambda_{eph} = \frac{\nu_0\langle\zeta^2(\mathbf{p})\rangle}{\omega_0^3 M'N} \approx \frac{p_F am^*}{m}.$$

We take the value of the Faust–Henry coefficient $C = -0.5$ and the phonon natural width $\Gamma^{nat} = 10^{-2}\omega_{TO}$. The electron collision rate is taken as $\tau^{-1} = 10^{-1}\omega_{TO}$, which is the usual value for heavily doped semiconductors [13, 19].

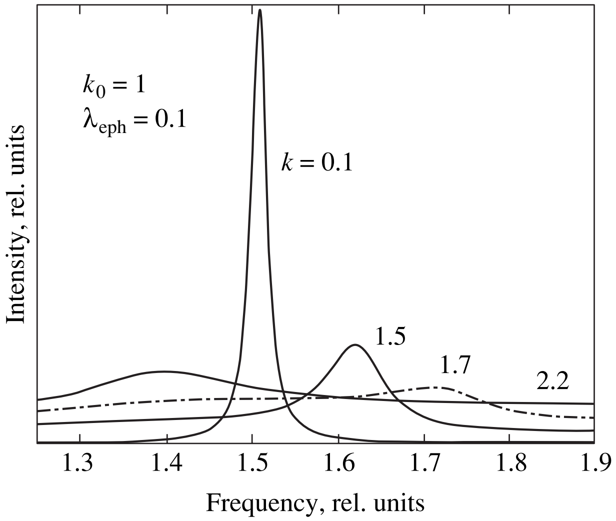

Fig. 5. The LO phonon Raman spectra for large momentum transfers k.

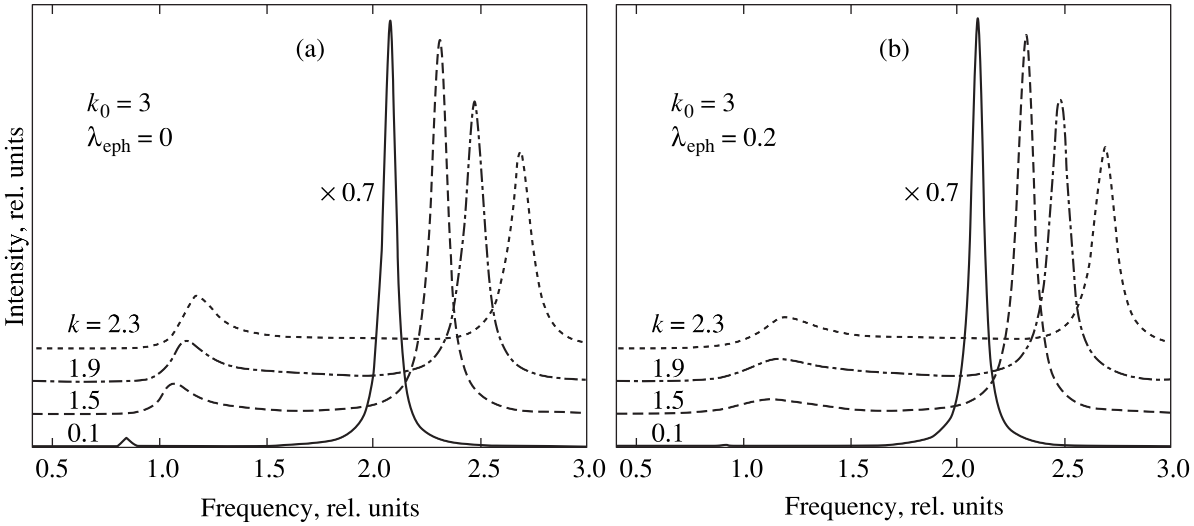

Fig. 6. Raman spectra from a heavily doped semiconductor for $k_0 = 3$ and $\lambda_{eph} = 0$ (a) and 0.2 (b); the notations as in Fig. 1.

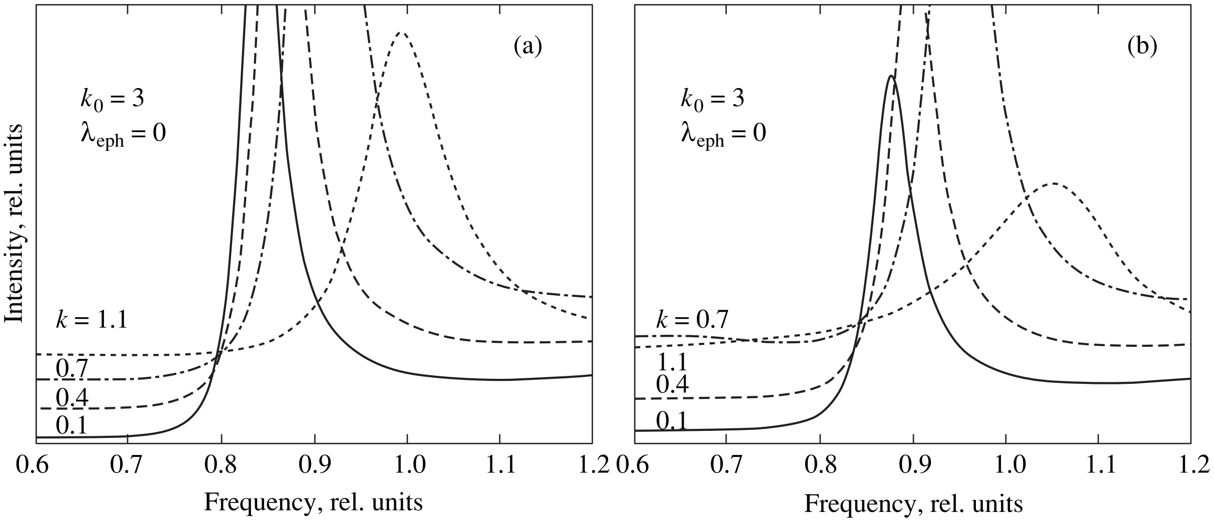

Fig. 7. The LO phonon part of the Raman spectra from heavily doped semiconductor for various momentum transfers without (a) and with (b) electron–phonon interaction taken into account.

In Fig. 3, the wave vector k and the Thomas–Fermi parameter k_0 are given in units of ω_{TO}/v_F, and the frequency ω, in units of ω_{TO}. Both these figures correspond to the case of small carrier numbers $\omega_{pe} < \omega_{TO}$ (see Fig. 2a; for the quadratic electron spectrum, $\omega_{pe} = k_0 v_F/\sqrt{3}$). The left peak mainly has a plasmon character and the right peak is mainly the LO phonon. We set the ion plasmon frequency $\omega_{pi} = \omega_{TO}$; therefore, $\omega_{LO} = \sqrt{2}\,\omega_{TO}$. As the wave vector k approaches the boundary of the Landau damping region $kv_F > \omega$, the plasmon peak becomes broader and almost disappears at $k = 0.8$. The broad continuum in the region $kv_F > \omega$ results from the excitation of electron–hole pairs. The intensity of the plasmon peak becomes larger in comparison to the phonon peak as the electron–phonon interaction λ_{eph} increases.

The k-dispersion of the plasmon (the peak position of the Raman spectra as a function of k), the line width (the full width at half-maximum), and the line asymmetry (the frequency difference between the right and left wings of the resonance line at half-maximum) are shown in Fig. 4, all in units of ω_{TO}. The width and asymmetry become much larger, while the plasmon peak is immersed in the electron–hole continuum. The maximum in this region of the spectra is essentially nothing but the electron–hole contribution. In Fig. 4b, we see how close this maximum is located on the line $\omega = kv_F$ for $\lambda_{eph} = 0.1$.

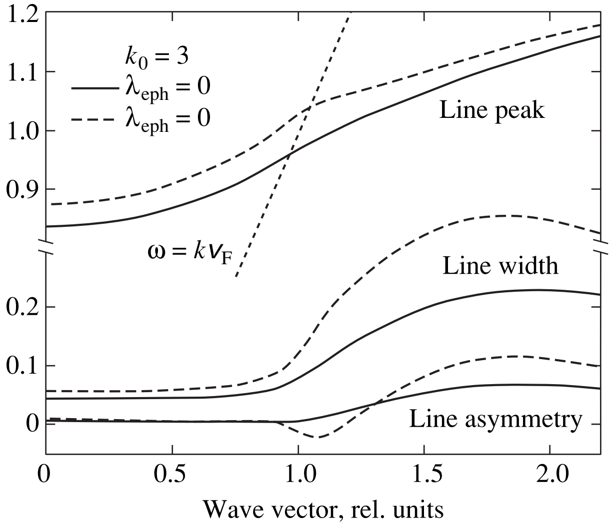

Fig. 8. Dispersion of the phonon peak (upper), the line width, and the line asymmetry (bottom). The boundary of the Landau damping region is shown with a dotted line.

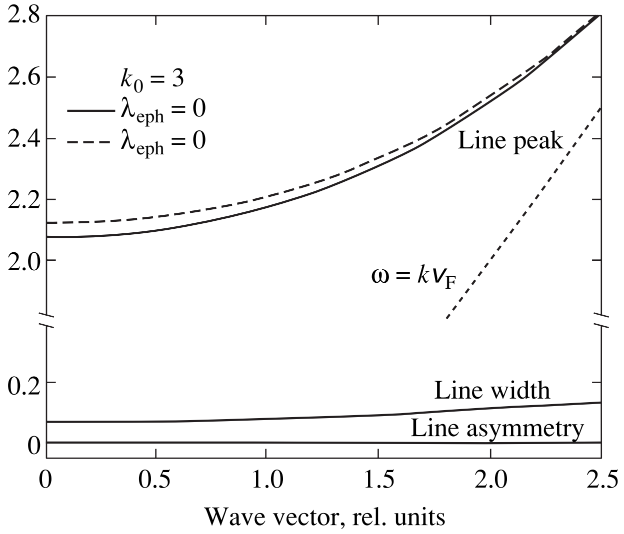

Fig. 9. Dispersion of the plasmon peak in heavily doped semiconductors.

The behavior of the phonon peak around $\omega = \omega_{LO}$ as k increases is shown in Fig. 5. As the wave vector increases from $k = 0$ to $k = 1.7$, the phonon peak is evidently shifted to a higher frequency and becomes broader. This is an effect of the Landau damping (see Fig. 2b). However, for $k > 1.75$, this peak appears at a lower frequency, $\omega \approx 1.4$, and becomes sharper for $k > 2.2$ because the Landau damping decreases with k (see Eq. (26)).

The Raman spectra for heavily doped semiconductors and metals are shown in Fig. 6 (see also Fig. 2a). The phonon peak is now located around $\omega \sim \omega_{TO}$ instead of $\omega \sim \omega_{LO}$. This is an effect of the Coulomb screening: carriers decrease the frequency of the LO mode from ω_{LO} to ω_{TO}. We also see that the electron–phonon interaction suppresses the phonon peak.

The effect of the Coulomb screening and electron–phonon interaction on the phonon mode is clearly seen in Fig. 7, where the phonon part of the spectra is shown in detail. The lines are very asymmetric. The phonon dispersion, the line width, and the line asymmetry as functions of k are shown in Fig. 8. We see a singularity at $k \approx \omega/v_F$. It is interesting to estimate the value of the phonon dispersion. With the help of Fig. 7b, we find $d\omega/dk \le 10^{-1} v_F$. On the other hand, using Eqs. (26) and (29), we find for the phonon dispersion $\omega^2 = \omega_{TO}^2 + \omega_{pi}^2 k^2/k_0^2$, which corresponds well with the previous estimate for our values of k_0 and ω_{pi}. We note that these estimates confirm the adiabatic approximation, because the value of dispersion $s = \omega_{pi}/k_0 \sim v_F\sqrt{m/M}$ contains the adiabatic parameter.

In Fig. 9, the dispersion, the line width, and the line asymmetry are shown for the plasmon peak in heavily doped semiconductors. Here, the effect of the electron–phonon interaction on the phonon dispersion is weak and no influence on the width and asymmetry of the line is seen.

8. CONCLUSIONS

In conclusions, we first emphasize that our result (23) describes the renormalization of the phonon frequencies, the effective ion charge, and the coupling constants due to the electron–phonon deformation interaction $\zeta_j(\mathbf{p})$. Second, this result involves the k-dependent semiclassical dielectric function instead of the Lindhard–Mermin expression. Finally, the light scattering vertex $\gamma(\mathbf{p})$ with excitations of the electron–hole pairs not only gives an additional contribution e in Eq. (23), but also modifies the electrooptic g_E and deformation-optic g_j coupling constants, which become dependent on the frequency and momentum transfers.

The author acknowledges the kind hospitality of the Max-Planck-Institut für Physik Komplexer Systeme (Dresden), were this work was completed. The work was partially supported by the RFBR (project no. 01-02-16211).

REFERENCES

1. A. B. Migdal, Zh. Éksp. Teor. Fiz. **34**, 1438 (1958) [Sov. Phys. JETP **7**, 996 (1958)].
2. A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Physics* (Fizmatgiz, Moscow, 1962; Prentice Hall, Englewood Cliffs, N.J., 1963).
3. A. S. Alexandrov and J. R. Schrieffer, Phys. Rev. B **56**, 13731 (1997).

4. M. Born and Kung Huang, *Dynamical Theory of Crystal Lattices* (Oxford Univ. Press, New York, 1954).
5. E. G. Brovman and Yu. Kagan, Zh. Éksp. Teor. Fiz. **52**, 557 (1967) [Sov. Phys. JETP **25**, 365 (1967)].
6. V. M. Kontorovich, Usp. Fiz. Nauk **142**, 265 (1984) [Sov. Phys. Usp. **27**, 134 (1984)].
7. M. Reizer, Phys. Rev. B **61**, 40 (2000).
8. V. L. Gurevich, A. I. Larkin, and Yu. A. Firsov, Fiz. Tverd. Tela (Leningrad) **4**, 185 (1962) [Sov. Phys. Solid State **4**, 131 (1962)].
9. L. A. Falkovsky, Phys. Rev. B **66**, 020302 (2002); Zh. Éksp. Teor. Fiz. **122**, 411 (2002) [JETP **95**, 354 (2002)].
10. G. Abstreiter, M. Cardona, and A. Pinczuk, *Light Scattering in Solids*, Ed. by M. Cardona and G. Güntherodt (Springer, Berlin, 1984), p. 5, Topics in Applied Physics, Vol. 54.
11. D. T. Hon and W. L. Faust, Appl. Phys. **1**, 241 (1973).
12. B. H. Bairamov, I. P. Ipatova, V. A. Milorava, *et al.*, Phys. Rev. B **38**, 5722 (1988).
13. L. Artús, R. Cuscó, and J. Ibánez, Phys. Rev. B **60**, 5456 (1999).
14. N. D. Mermin, Phys. Rev. B **1**, 2362 (1970).
15. L. A. Falkovsky and E. G. Mishchenko, Phys. Rev. B **51**, 7239 (1995).
16. A. A. Abrikosov and V. M. Genkin, Zh. Éksp. Teor. Fiz. **65**, 842 (1974) [Sov. Phys. JETP **38**, 417 (1974)].
17. A. A. Abrikosov and L. A. Falkovsky, Zh. Éksp. Teor. Fiz. **40**, 262 (1961) [Sov. Phys. JETP **13**, 179 (1961)].
18. A. Zawadowski and M. Cardona, Phys. Rev. B **42**, 10732 (1990).
19. R. Fukasawa and S. Perkowitz, Phys. Rev. B **50**, 14119 (1994).

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 806–809.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 898–902.
Original Russian Text Copyright © 2003 by Nikolaev, Dementiev.

SOLIDS
Electronic Properties

NMR Study of the Displacement of Vortices in the Superconductor $Tl_2Ba_2CuO_{6+\delta}$ under the Action of a Magnetic Field Pulse

E. G. Nikolaev* and A. E. Dementiev

Kapitsa Institute for Physical Problems, Russian Academy of Sciences, ul. Kosygina 2, Moscow, 119334 Russia
**e-mail: nikolaev@kapitza.ras.ru*

Received April 2, 2003

Abstract—A transition from irreversible to reversible displacement of vortices in the high-temperature superconductor $Tl_2Ba_2CuO_{6+\delta}$ (T_c = 78 K) is investigated by the NMR method combined with pulse modulation of the magnetic field in the temperature range from 20 to 35 K. The results also make it possible to estimate the distance between pinning centers in the sample, which lies in the interval from 4 to 10 nm at a temperature of 20 K. *© 2003 MAIK “Nauka/Interperiodica”.*

Nuclear magnetic resonance, as well as other local methods (such as various versions of probe microscopy [1] and muonic precession [2]), is widely used for studying the mixed state in high-T_c superconductors. This method provides information on the static magnetic field distribution in the bulk of a superconductor in the mixed state [3–5] and on thermal motion of vortices [6–8]. An analysis of the effect of external agencies on a vortex system (e.g., transport current or magnetic field variation) with the help of the NMR method is also of considerable interest for understanding the flux pinning and creep mechanisms in superconductors. The number of publications in this field is scarce and the results reported in these publications are not always convincing. In one of the first publications [9], a narrowing of the NMR line due to field fluctuations at thallium nuclei under the conditions of flux flow under the effect of a direct current was observed in the PbTl superconducting alloy. According to Carretta [10], a similar effect was also observed at the ^{89}Y line in a polycrystalline sample of the high-temperature superconductor $YBa_2Cu_3O_{7-\delta}$. Lefloch *et al.* [11] used the NMR method for studying the motion of vortices in a superconductor under the action of a deflecting transverse low-frequency ac field and determined the magnetic field penetration depth in the NbTi alloy. Other attempts at using the NMR method for observing the motion of vortices in HTSC, which is initiated by field [12] or current pulses [13], were also made (in the latter publication, the result was negative).

This work is devoted to an analysis of a system of vortices in the high-temperature superconductor $Tl_2Ba_2CuO_{6+\delta}$ by the NMR method combined with pulsed magnetic field modulation. This method was successfully used for observing displacement of vortices after the passage of a field pulse through oriented powder samples and for analyzing the temperature dependence of this effect. As a result, information was obtained on the spatial distribution of pinning centers and on thermally activated depinning in this system.

A sample of the compound $Tl_2Ba_2CuO_{6+\delta}$ (T_c = 78 K), which was used earlier in [5], was in the form of a powder in paraffin, which was oriented in a magnetic field of 110 kG. The powder was preliminarily separated so that the particle size did not exceed 20 μm. The NMR experiments were made for the magnetic field orientation along the preferred orientation of particles in the sample, which corresponded to the axis of the structure of the compound in question (perpendicular to CuO_2 planes). The displacement of vortices under the action of a field pulse was observed with the help of the nuclear spin echo, which made it possible to determine changes in the local fields at nuclei during the echo formation from the decrease in the signal intensity. For a conventional sequence of two rf pulses ($\pi/2 - \tau - \pi$), this time corresponds to the interval 2τ between the first rf pulse and the instant of the formation of the echo signal and is limited by a value on the order of the spin–spin relaxation time T_2. In our case, we used a sequence of induced echo ($\pi/2 - \tau_1 - \pi/2 - \tau_2 - \pi/2$) [14], for which an echo signal appears after time interval τ_1 following the third rf pulse. In this case, the echo attenuates as a result of relaxation over time interval τ_2 with a characteristic time considerably longer than T_2. For thallium nuclei in the given system, this time does not exceed 50 μs [15], which makes it possible to detect the variation of local fields at nuclei over a much longer time interval. A magnetic field pulse having the shape of an isosceles triangle, a duration of up to 3 ms, and an amplitude of B_m up to 130 G, which was superimposed on a static field of 9 kG, was supplied 1 ms after the second rf pulse.

Figure 1 shows the normalized dependences of the echo signal intensity on delay time τ_2 for different values of the modulating magnetic field pulse amplitude at a temperature of 18 K. In the absence of pulsed modulation, the echo signal decay for small values of τ_2 is associated with spin diffusion, while for large values of τ_2, it is determined by spin–lattice relaxation. It should be noted that constant magnetic field B_0 was applied to the sample after its zero-field cooling; as a result, a nonuniform distribution of vortices was formed in particles of the sample due to pinning even before a sequence of pulses was supplied. The sample particles were characterized by an irregular (nonplanar) shape; for this reason, it is impossible to describe exactly the field distribution in an individual particle. However, we assume qualitatively that, on the average, the shape of the particles is close to that of a short cylinder whose axis is parallel to the external field. In this case, for qualitatively describing the field distribution in superconducting particles, we will use the simple Bean model if the critical state [16] as we did earlier [5]. In this model, after the application of a field, the density of vortices and the magnetic induction averaged over the length exceeding the separation between vortices in a superconducting cylinder decrease linearly due to pinning in the direction from the lateral surface of the cylinder to its axis. In this case, in accordance with the Maxwell equation $\operatorname{curl} B = 4\pi j/c$, a circular current flows in the bulk of the superconductor, the current density being equal to the critical current density j_c. The applicability of this model in our case is confirmed by a hysteresis loop in the NMR line for thallium, which is typical of the critical state and was studied in detail in [5]. The corresponding radial distribution of the average induction in a superconducting particle after the application of field is shown schematically in the inset to Fig. 1 by a thin line starting at B_0. The field direction in a pulse was opposite to the direction of the constant magnetic field. For this reason, while the field decreased from the lateral surface of the superconducting particle to its axis before the application of the pulse, the magnetic field gradient at the peak of the pulse in a layer adjoining the lateral surface reverses its sign since vortices in this region partly emerge from the superconductor (the line beginning at $B_0 - B_m$). The absolute value of this gradient must be considerably larger than the value of the initial gradient, which is reduced due to flux creep. Accordingly, after the passage of a magnetic field pulse, part of the vortices return to the superconductor and the field distribution in a layer adjoining the lateral surface assumes the form shown by the bold line in the inset to Fig. 1. After the passage of several modulating pulses, a quasi-stationary distribution of magnetic induction, shown by the dashed curve, must set in near the particle axis due to flux creep. In this region vortices are stationary and, hence, this region makes zero contribution to the effect.

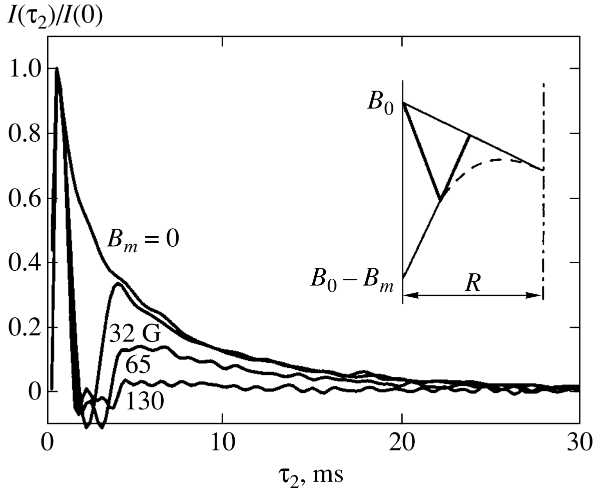

Fig. 1. Dependence of the normalized spin echo signal intensity for thallium nuclei on the delay time τ_2 in a powder sample of $Tl_2Ba_2CuO_{6+\delta}$ for different values of the modulating magnetic field pulse amplitude: $T = 18$ K, $B_0 = 9$ kG. The inset schematically shows the radial distribution of the magnetic induction in a superconducting particle under the conditions of pulsed field modulation (see text).

It can be seen from Fig. 1 that a current pulse additionally reduces the echo signal amplitude at subsequent instants as compared to the relaxation curve in the absence of pulse modulation, the effect increasing with the pulse amplitude at a constant temperature. The observed decrease in the echo signal intensity indicates that the local field distribution changes significantly after each field pulse, while the average field distribution in the sample particles must remain unchanged. This change in the local field distribution in the particles can be naturally attributed to the fact that, after the passage of each modulating pulse, vortices do not return exactly to their positions before the passage of the field pulse. The enhancement of the observed effect upon an increase in the modulation amplitude can be explained by an increase in the thickness of the layer in which the motion of vortices takes place together with a variation of the magnetic induction gradient during the action of the modulating pulse associated with this motion; the increased thickness may cover the entire sample volume.

In order to characterize the observed effect quantitatively, we will use the quantity $(1 - I_m/I_0)$, where I_m is the intensity of the echo signal for $\tau_2 = 10$ ms (by this time instant, perturbations induced by a field pulse attenuate) and I_0 is the same quantity in the absence of pulsed field modulation. Figure 2 shows the temperature dependences of quantity $(1 - I_m/I_0)$ for various values of the modulating field amplitude B_m. It can be seen from the figure that the maximal effect for which $(1 - I_m/I_0)$ attains values of about 0.8 for $B_m = 130$ G is observed in the vicinity of 20 K. This corresponds to complete penetration of the modulating field in the bulk

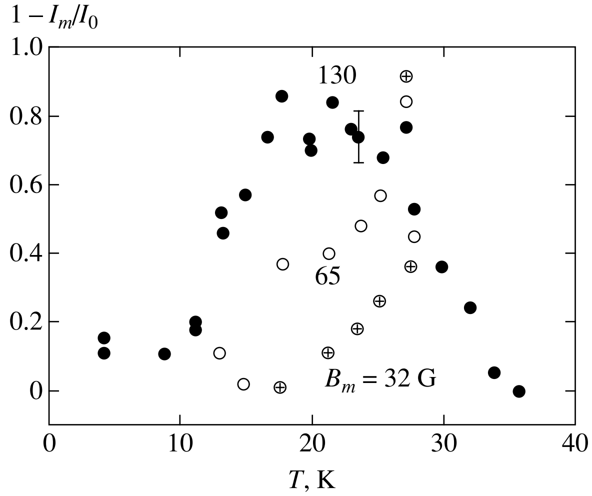

Fig. 2. Temperature dependence of the relative decrease in the echo signal intensity $(1 - I_m/I_0)$ for thallium nuclei in $Tl_2Ba_2CuO_{6+\delta}$ under the action of a modulating magnetic field pulse for different values of modulation amplitude.

of superconducting particles, indicating a displacement of vortices in the major part of the volume except the very central region during the passage of a pulse. At lower temperatures, modulation of such amplitude does not affect the entire volume of the particles in view of strong pinning and, hence, large values of the vortex concentration gradient and the magnetic induction gradient. The decrease in the value of $(1 - I_m/I_0)$ at B_m = 130 G above 20 K indicates that the motion of vortices under the action of a field pulse in this temperature range becomes irreversible, which is apparently due to a decrease in the number of possible effective pinning centers at the initial stage. A further increase in temperature may lead to complete depinning, which follows from completely reversible motion of vortices at 35 K. Thus, we can state that this curve characterizes the thermally activated depinning in a temperature range of 20–35 K. As the modulation amplitude decreases, the effect becomes weaker in general, and the peak of the quantity $(1 - I_m/I_0)$ is displaced towards higher temperatures. It was noted above that this peak is associated with complete penetration of the modulating field in a superconducting particle; for lower values of the modulation amplitude, this peak must be observed for a lower value of the magnetic induction gradient and, hence, at a higher temperature. Actually, above 20 K, the temperature dependence of the effect for $B_m <$ 130 G must be the result of two competing processes: (i) the enhancement of the observed effect due to an increase in the volume in which the motion of vortices takes place due to a decrease in the induction gradient with temperature and (ii) the suppression of the effect in view of the fact that part of the vortices start moving in reverse.

Using the results obtained, we can estimate the displacement of vortices under the action of a magnetic field pulse in the present case. The mean distance between vortices in a field of B_0 = 9 kG amounts to $a_0 = (2\Phi_0/\sqrt{3}B_0)^{1/2}$ = 50 nm (Φ_0 is the magnetic flux quantum). The existence of a gradient in the critical state changes this value insignificantly since the difference between the field at the surface and in the bulk of a particle is much smaller than B_0. It can be proved that a change in the external field by 130 G (pulse peak) leads to a displacement of vortices; for particles 20 μm in size, the average value of this displacement is on the order of 10 nm (provided that a pulse penetrates the particle completely). After the passage of the pulse peak, vortices move in the opposite direction. As noted above, the strong decrease in the echo signal intensity under the action of a field pulse near 20 K indicates that vortices do not return to their previous positions. The resultant minimal displacement of vortices can be estimated as follows. A spin echo signal decays completely when $\gamma(\mathrm{Tl})\tau_1\delta B = \pi/2$, where $\gamma(\mathrm{Tl}) \approx 2.5$ kHz/G is the gyromagnetic ratio for thallium nuclei, $\tau_1 = 12$ μs is the delay between the first and second pulses, and δB is the change of the field at a nucleus due to the displacement of vortices. In our case, the value of δB obtained using this formula amounts to 7 G. Using the estimate for the average field nonuniformity in a vortex lattice, $\langle \Delta B^2 \rangle^{1/2} = 0.064\Phi_0/\lambda_{ab}^2$ = 40 G (penetration depth λ_{ab} = 180 nm [3]), we find that the average value of the magnetic induction gradient between the cores of vortices is equal to $2\langle \Delta B^2 \rangle^{1/2}/a_0 \sim 1.6$ G/nm. Accordingly, the minimal displacement $\delta a_0 \sim (\delta B/2\langle \Delta B^2 \rangle^{1/2})a_0$ of vortices induced by a field pulse is 4 nm in the vicinity of 20 K.

It should be noted that the value of field B_0 = 9 kG in which our measurements were made is considerably higher than the value of the field corresponding to the second peak on the curve describing the dependence of the magnetic moment on the field (3 kG at 20 K), which was determined for $Tl_2Ba_2CuO_{6+\delta}$ single crystals with a close value of T_c [17, 18]. According to Giamarchi and Doussal [19], an increase in the field in the vicinity of this peak gives rise to a large number of dislocations in the vortex lattice due to the crystal structure defects and leads to a transition to the state of a vortex glass or, to be more precise, a supercooled vortex liquid. This is also facilitated by a crossover from the 3D to a 2D system of vortices with a characteristic field of $B_{3D-2D} = \Phi_0/\gamma^2 s^2 \sim 1$ kG for the given system, where $\gamma = 190$ is the anisotropy parameter and $s = 1.16$ nm is the distance between CuO_2 layers [17]. Consequently, we can assume that each vortex in our case (at 20 K) is located at its own pinning center and the number of such closely spaced centers is sufficiently large. As a result, after each field pulse, a different possible configuration of vortices is realized, which leads to a considerable decrease in the echo signal intensity. In this connection, the value of 4 nm obtained above for the minimal displacement of vortices as a result of the passage of a field pulse can serve as an estimate of the lower limit of the distance between pinning enters. The above estimate (10 nm) of the average displacement of vortices at the pulse peak gives the upper limit of the distance between pinning centers in the studied system.

Thus, we have applied the NMR method for studying the displacement of vortices in the high-temperature superconductor $Tl_2Ba_2CuO_{6+\delta}$ after the action of a short magnetic field pulse superimposed on a constant magnetic field with an induction of 9 kG. It is shown that the minimal displacement of vortices in sample particles after the application of a field pulse with an amplitude of 130 G at 20 K amounts to 4 nm. In our case, this value can be taken as the lower limit of the distance between pinning centers at this temperature. The upper limit of the distance between pinning centers, which is equal to 10 nm, is determined by an estimate of the average displacement of vortices at the pulse peak. It has been established that the fraction of vortices returning after the action of a field pulse to their initial positions as a result of depinning upon an increase in temperature from 20 to 35 K increases to 100%.

ACKNOWLEDGEMENTS

The authors are grateful to A.I. Khar'kovskiĭ, V.I. Tsebro, and A.V. Mitin for fruitful discussions and comments and to A.V. Mitin and D.Yu. Sharvin for their help in preparing samples. The orientation of samples in a magnetic field was carried out at the International Laboratory of Strong Magnetic Fields and Low Temperatures (Wroclaw, Poland).

This study was partly financed by the Ministry of Industry and Science of the Russian Federation (state contract no. 40.012.1.1.1357).

REFERENCES

1. A. de Lozanne, Supercond. Sci. Technol. **12**, R43 (1999).
2. J. E. Sonier, J. H. Brewer, and R. F. Kiefl, Rev. Mod. Phys. **72**, 769 (2000).
3. M. Mehring, F. Hentsch, H. Mattausch, *et al.*, Solid State Commun. **75**, 753 (1990).
4. H. B. Brom and H. Alloul, Physica C (Amsterdam) **177**, 297 (1991).
5. E. G. Nikolaev, Physica C (Amsterdam) **250**, 39 (1995).
6. B. J. Suh, D. R. Torgeson, and F. Borsa, Phys. Rev. Lett. **71**, 3011 (1993).
7. B. J. Suh, F. Borsa, J. Sok, *et al.*, Phys. Rev. Lett. **76**, 1928 (1996).
8. E. G. Nikolaev, J. Witteveen, M. de Kok, *et al.*, Phys. Rev. B **55**, R8717 (1997).
9. J.-M. Delrieu, J. Phys. (France) **3**, 893 (1973).
10. P. Carretta, Phys. Rev. B **48**, 528 (1993).
11. F. Lefloch, W. G. Clark, and W. H. Wong, Phys. Rev. B **59**, 7094 (1999).
12. F. Borsa, P. Caretta, F. Cintonesi, *et al.*, Physica C (Amsterdam) **235–240**, 2629 (1994).
13. C. H. Recchia, C. H. Pennington, H. Hauglin, *et al.*, Phys. Rev. B **52**, 9746 (1995).
14. J. Jeener and P. Broekaert, Phys. Rev. **157**, 232 (1967).
15. F. Hentsch, N. Winzek, M. Mehring, *et al.*, Physica C (Amsterdam) **165**, 485 (1990).
16. C. P. Bean, Rev. Mod. Phys. **36**, 31 (1964).
17. F. Zuo, S. Khizroev, G. C. Alexandrakis, *et al.*, Phys. Rev. B **52**, R755 (1995).
18. M. Xu, T. W. Li, D. G. Hinks, *et al.*, Phys. Rev. B **59**, 13632 (1999).
19. T. Giamarchi and P. L. Doussal, Phys. Rev. B **55**, 6577 (1997).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 810–825.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 903–920.
Original Russian Text Copyright © 2003 by Uchaikin.

NONLINEAR PHYSICS

Anomalous Diffusion and Fractional Stable Distributions

V. V. Uchaikin
Ul'yanovsk State University, ul. L. Tolstogo 42, Ul'yanovsk, 432970 Russia
e-mail: uchaikin@sv.uven.ru

Received October 2, 2002

Abstract—A new set of distributions, called fractional stable distributions, is described. A subset of this set is represented by stable laws, while its particular case is the Gaussian distribution. These distributions arise as solutions to fractional-order partial differential equations that represent a generalization of the ordinary diffusion equation to the case of anomalous diffusion. The properties of multidimensional fractional stable densities are described, their expressions in terms of special functions are presented, and physical problems that lead to these densities are discussed. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Since Richardson, who discovered that a diffusion packet spreads in turbulent medium by the law $t^{3/2}$ [1], a large volume of experimental data has been accumulated on various processes of anomalous diffusion in which not only the diffusion law but also the shape of the diffusion packet itself essentially differ from those in the normal case. These data include propagation of resonant radiation in a medium of large optical size [2, 3]; a charge transfer in amorphous semiconductors [4]; an NMR diffusometry of inhomogeneous structures [5]; the dynamics of polymer systems [6]; diffusion on fractal structures [7], diffusion of a scalar tracer in an array of convection rolls [8], on solid surfaces [9], in porous glasses [10], and in rock [11]; and processes in quantum optics [12], in plasma [13], etc. A survey of problems of this type can be found in [14–17].

Several approaches have been developed to describe anomalous diffusion; they use a space-dependent diffusivity [18], power-law correlations of fractional order [19], jump random walks [20–22], fractional derivatives [23–26], generalizations of the Fokker–Planck equations [27], a generalized thermodynamic approach [28], etc. (see the surveys in [14–16, 29–32]).

By now, it has been established that the mathematical basis of anomalous self-similar diffusion is formed by fractional differential equations. In simple one-dimensional cases, solutions to these equations have been found and expressed in terms of H functions—the Fox functions that generalize the well-known G functions of Meyer [33]. In [34, 46], we considered three-dimensional models of enhanced diffusion (superdiffusion) and retarded diffusion (subdiffusion), derived appropriate equations, found their solutions, and presented numerical results. However, further investigations have shown that the cases considered (superdiffusion as a stable Lévy motion and subdiffusion as a retarded Brownian motion) do not exhaust all possible regimes of anomalous diffusion. We succeeded in deriving the basic equations of anomalous diffusion from general principles without resorting to specific models; we also found solutions of these equations that form a new class of distributions, called fractional stable distributions, and analyzed the properties of these solutions. These and related problems are discussed in the present paper.

2. STABLE DISTRIBUTIONS

Normal isotropic diffusion is described by the equation

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = D\Delta p(\mathbf{x}, t), \quad \mathbf{x} \in R^m, \tag{1}$$

where D is the diffusion coefficient and $p(\mathbf{x}, t)$ is the spatial probability density of a diffusing particle at moment t. Under the initial condition $p(\mathbf{x}, 0) = \delta(\mathbf{x})$, this density is expressed as

$$p(\mathbf{x}, t) = (4\pi Dt)^{-m/2}\exp\left(-\frac{r^2}{4Dt}\right), \quad r = |\mathbf{x}|. \tag{2}$$

Notice the following three features of solution (2) to Eq. (1). First, this equation is self-similar,

$$p(\mathbf{x}, t) = t^{-m/\alpha}g_m(\mathbf{x}t^{-1/\alpha}), \quad \alpha > 0; \tag{3}$$

second, it is homogeneous in the sense that it describes a process that is homogeneous in time and space,

$$p(\mathbf{x}, t_1 + t_2) = \int p(\mathbf{x} - \mathbf{x}', t_1)p(\mathbf{x}', t_2)d\mathbf{x}'; \tag{4}$$

and, third, it has a finite variance,

$$\int |\mathbf{x}|^2 p(\mathbf{x}, t)d\mathbf{x} < \infty. \tag{5}$$

The converse is also true: if we define a process by properties (3)–(5), then we inevitably arrive at distribution (2), which satisfies Eq. (1).[1] In this case, the similarity exponent $\alpha = 2$ is also determined uniquely. Thus, expressions (3)–(5) can be considered as three postulates (three axioms) that determine the process under consideration. In this statement, the generalization procedure is reduced to the rejection of one or even two axioms that restrict the class of processes under consideration.

Let us pass from the probability densities $p(\mathbf{x}, t)$ to their Fourier images,

$$\tilde{p}(\mathbf{k}, t) = \int \exp(i\mathbf{k}\cdot\mathbf{x})p(\mathbf{x}, t)d\mathbf{x},$$

which are called characteristic functions in probability theory. Then, formulas (3) and (4) are rewritten as

$$\tilde{p}(\mathbf{k}, t) = \tilde{g}_m(\mathbf{k}t^{1/\alpha}), \tag{6}$$

$$\tilde{p}(\mathbf{k}, t_1 + t_2) = \tilde{p}(\mathbf{k}, t_1)\tilde{p}(\mathbf{k}, t_2), \tag{7}$$

respectively.

Combining these equalities, we obtain the equation

$$\tilde{g}_m((t_1 + t_2)^{1/\alpha}\mathbf{k}) = \tilde{g}_m(t_1^{1/\alpha}\mathbf{k})\tilde{g}_m(t_2^{1/\alpha}\mathbf{k}). \tag{8}$$

The density

$$g_m(\mathbf{x}) = \int \exp(i\mathbf{k}\cdot\mathbf{x})\tilde{g}_m(\mathbf{k})d\mathbf{k}$$

generated by this characteristic function is called the density of a stable distribution or, in short, a stable density with characteristic exponent α.

In the foreign literature, stable distributions are often called Lévy distributions, or, when one deals with multidimensional distributions, Lévy–Feldheim distributions [47].

A set of solutions to Eq. (8) that satisfies the conditions $\tilde{g}_m(0) = 1$ and $|\tilde{g}_m(\mathbf{k})| \le 1$ necessary for the characteristic functions is expressed as [48]

$$\tilde{g}_m(\mathbf{k}; \alpha, \Gamma(\cdot))$$
$$= \exp\left\{-\int_{U_m} |\mathbf{k}\cdot\mathbf{u}|^\alpha\left[1 - i\,\mathrm{sgn}(\mathbf{k}\cdot\mathbf{u})\tan\frac{\pi\alpha}{2}\right]\Gamma(d\mathbf{u})\right\}, \tag{9}$$

where $\Gamma(d\mathbf{u})$ is the spectral measure, i.e., a finite measure defined on an m-dimensional sphere U_m of unit radius in an m-dimensional space.

The value $\alpha = 2$ corresponds to the Gaussian distribution, while the values $\alpha < 2$ generate an infinite set of other stable distributions. The basic difference of these distributions from the Gaussian one is that their variance is infinite.

One-dimensional stable distributions were considered in [47]. The characteristic function of these distributions (more precisely, of strictly stable distributions) can be represented in the so-called C form:

$$\tilde{g}_1(k; \alpha, \theta) = \exp\left\{-|k|^\alpha \exp\left(-\frac{i\pi\alpha\theta}{2\,\mathrm{sgn}\,k}\right)\right\},$$

where

$$|\theta| \le \theta_\alpha = \min\left\{1, \frac{2}{\alpha} - 1\right\}$$

is the skewness parameter.

Among multidimensional distributions, the ones that have been more or less carefully studied are symmetric stable distributions that satisfy the condition $g(\mathbf{x}) = g(-\mathbf{x})$ [48]. They include isotropic stable distributions whose spectral measure $\Gamma_0(d\mathbf{u}) = Cd\mathbf{u}$ is uniformly distributed over sphere U_m and the characteristic function (under a special choice of constant C) is given by

$$\tilde{g}(\mathbf{k}; \alpha) = \exp\{-|\mathbf{k}|^\alpha\}.$$

When $\alpha = 2$, we have the characteristic function of isotropic Gaussian distribution (2), while the case $\alpha = 1$ corresponds to the m-dimensional Cauchy distribution

$$g(\mathbf{x}; 1) = \Gamma\left(\frac{m+1}{2}\right)[\pi(1 + r^2)]^{-(m+1)/2}, \quad r \equiv |\mathbf{x}|.$$

Other isotropic densities cannot be expressed in terms of elementary functions.

As we have already mentioned above, the variance and, hence, all the higher order moments of stable distributions are infinite. This fact is associated with the behavior of densities at large distances. In contrast to the normal case, for $\alpha \ne 2$, we have

$$g_m(\mathbf{x}; \alpha) \propto r^{-\alpha - m}, \quad r \longrightarrow \infty.$$

Another important feature of non-Gaussian isotropic distributions is that the projections $X_1, \ldots, X_m$ of random vectors $\mathbf{X}$ described by these distributions correlate; the correlation vanishes only for $\alpha = 2$. To obtain a symmetric stable distribution with independent parameters for $\alpha < 2$, one has to concentrate the spectral measure at the points where sphere U_m intersects the

---

[1] These requirements are satisfied by equations with any integer powers of the Laplacian; however, the condition that the probability density should be nonnegative eliminates all powers higher than first.

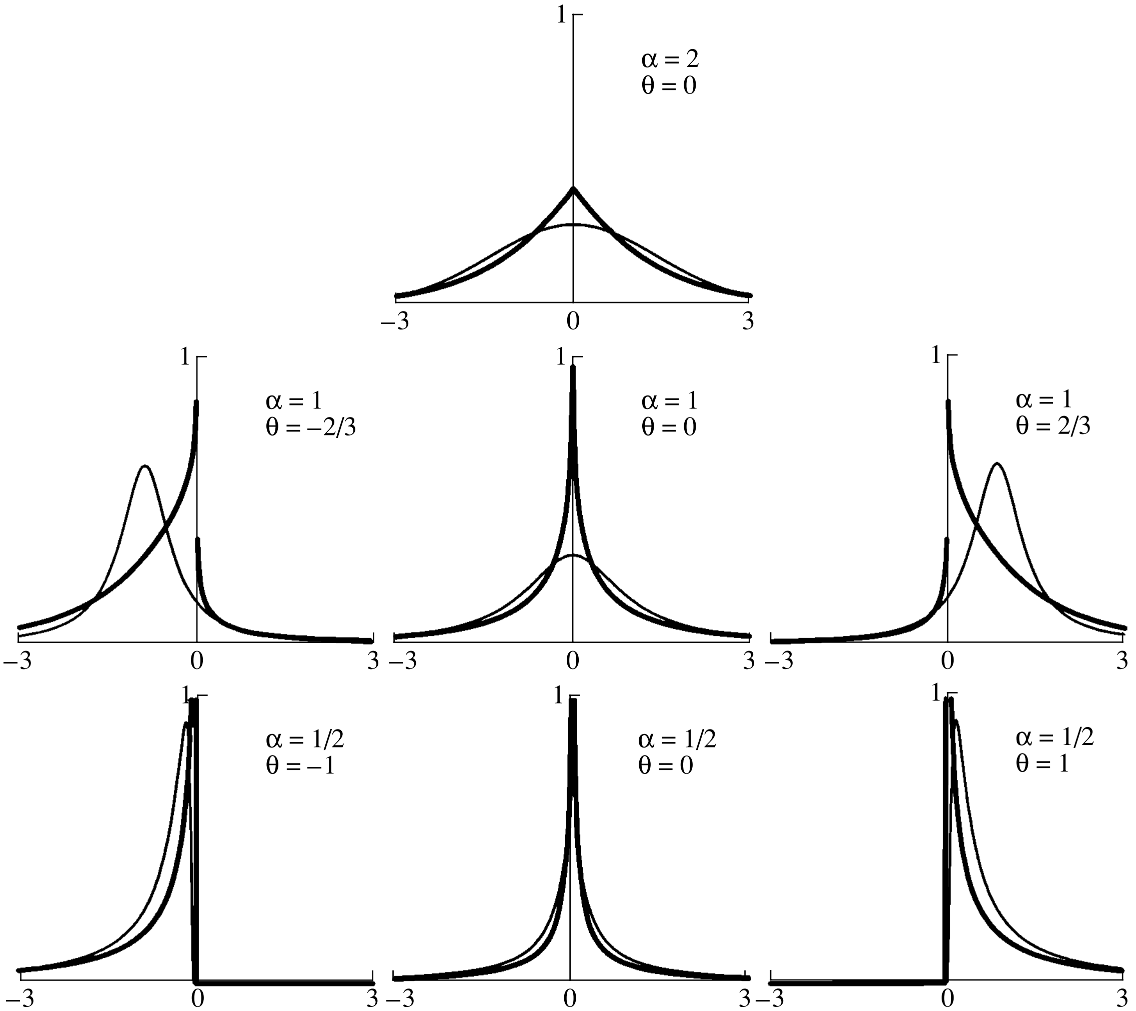

Fig. 1. Diagrams of one-dimensional stable (thin lines) and fractional stable (heavy lines) densities $q(x;\alpha, 1/2, \theta)$.

coordinate axes. Such a distribution is not isotropic; its characteristic function has the form

$$\tilde g(\mathbf{k};\alpha) = \exp\{-|k_1|^\alpha - \ldots - |k_m|^\alpha\},$$
$$\mathbf{k} = \{k_1, \ldots, k_m\}.$$

When the characteristic exponent α is different from 2 or 1, the densities of symmetric stable distributions are determined by numerical methods [45, 46]. Figures 1 (thin lines) and 2 represent the graphs of certain one-dimensional stable densities. Figures 3 shows two-dimensional stable distributions (Fig. 3a represents a distribution of the first, isotropic, type, and Fig. 3b, a distribution of the second type, with independent components).

The discovery of the class of stable distributions by Lévy is a milestone in the history of probability theory in the 20th century. The point is that these and only these distributions may serve as the limit distributions in a summation scheme for independent random variables. In this case, the central limit theorem, which is based on the requirement that the variances of random variables should be finite, is generalized to the case of variables with infinite variances. This generalized limit theorem underlies the mathematics of anomalous diffusion.

3. EQUATIONS OF ANOMALOUS DIFFUSION

Substituting expression (9) into (6), differentiating with respect to time, and taking into account the condition $\tilde p(\mathbf{k}, 0) = 1$, we arrive at the following equation for the characteristic function of the desired density:

$$\frac{\partial \tilde p(\mathbf{k}, l)}{\partial t} = \tilde L(\alpha, \Gamma(\cdot))\tilde p(\mathbf{k}, t) + \delta(t), \qquad (10)$$

where

$$\tilde L(\alpha, \Gamma(\cdot)) = -\int_{U_m} |\mathbf{k}\cdot\mathbf{s}|^\alpha \left[1 - i\,\mathrm{sgn}(\mathbf{k}\cdot\mathbf{u})\tan\frac{\pi\alpha}{2}\right]\Gamma(d\mathbf{u})$$

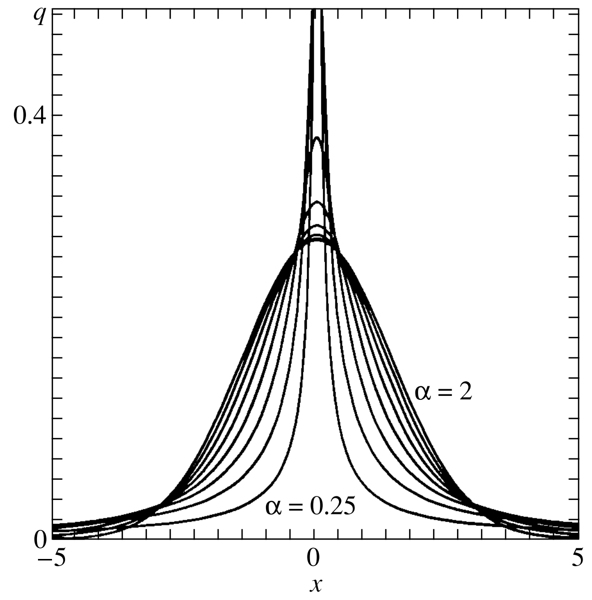

Fig. 2. One-dimensional symmetric stable distributions $q(x)$.

is the Fourier image of a new, generalized, diffusion operator $L(\alpha, \Gamma(\cdot))$. This operator, or, more precisely, its particular form for isotropic diffusion in a turbulent medium ($-(-\Delta)^{\alpha/2}$ in modern notation) was first introduced by Monin in [23]. Applying formally the inverse Fourier transformation to Eq. (10), we obtain

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = L(\alpha, \Gamma(\cdot))p(\mathbf{x}, t) + \delta(\mathbf{x})\delta(t). \qquad (11)$$

Equation (11) with $\alpha \in (0, 2]$ for $\alpha \neq 2$ describes a superdiffusion regime (which is also called Lévy flights). This is clearly seen from expression (3): the width of a diffusion packet increases with time as $t^{1/\alpha}$, which, for $\alpha < 2$, corresponds to a faster increase than in the normal ($t^{1/2}$) case.

Suppose that the self-similarity condition (3) holds for an arbitrary positive H,

$$p(\mathbf{x}, t) = t^{-mH} f(\mathbf{x}t^{-H}), \quad H \equiv \frac{\omega}{\alpha},$$

and the equation includes the operator $L(\alpha, \Gamma(\cdot))$, while homogeneity condition (4) is omitted. Then we arrive at the equation

$$\frac{\partial^{\omega} p(\mathbf{x}, t)}{\partial t^{\omega}} = L(\alpha, \Gamma(\cdot))p(\mathbf{x}, t) + \frac{t^{-\omega}}{\Gamma(1-\omega)}\delta(\mathbf{x}), \qquad (12)$$

$$\omega \le 1$$

with the Riemann–Liouville fractional part on the left-hand side. Note that, by the relation

$$\frac{\partial^{\omega-1}\delta(t)}{\partial t^{\omega-1}} = \frac{t^{-\omega}}{\Gamma(1-\omega)},$$

which can easily be proved by the Laplace transformation, we can represent Eq. (12) as

$$\frac{\partial p(\mathbf{x}, t)}{\partial t} = \frac{\partial^{1-\omega}}{\partial t^{1-\omega}} L(\alpha, \Gamma(\cdot))p(\mathbf{x}, t) + \delta(t)\delta(\mathbf{x}).$$

In the isotropic case, this equation is reduced to

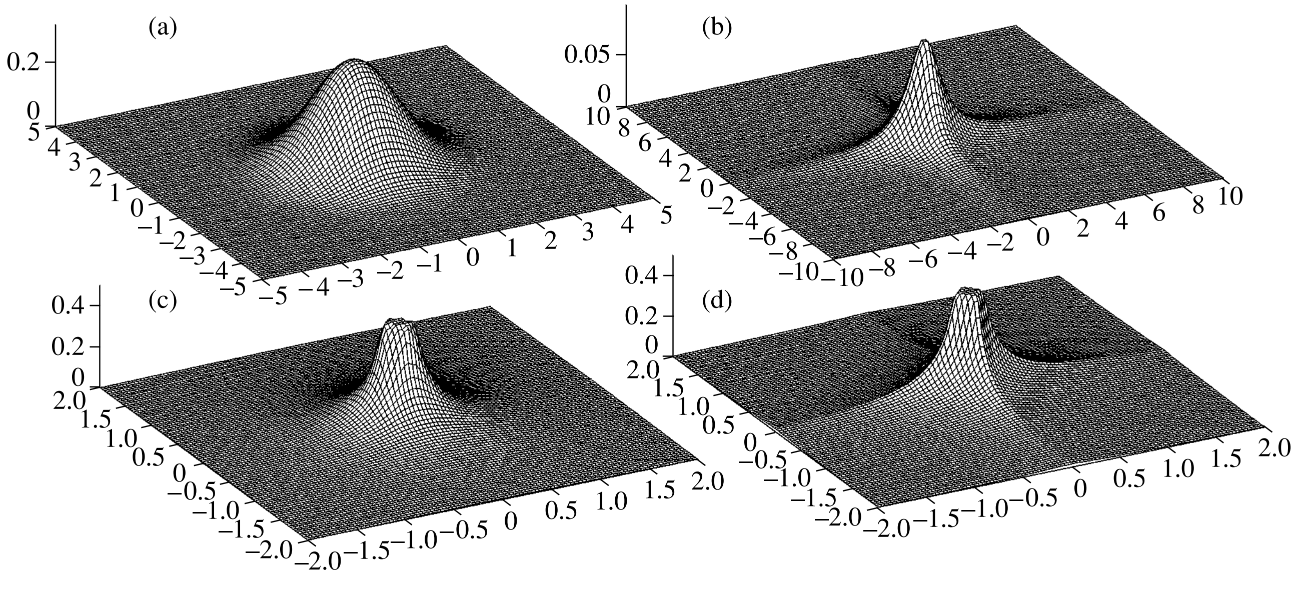

Fig. 3. (a, b) Two-dimensional symmetric stable distributions and (c, d) fractional stable densities; (a, c) isotropic densities and (b, d) densities with independent projections of a stable vector.

$$\frac{\partial^{\omega} p(\mathbf{x}, t)}{\partial t^{\omega}} = -D(-\Delta)^{\alpha/2} p(\mathbf{x}, t) + \frac{t^{-\omega}}{\Gamma(1-\omega)}\delta(\mathbf{x}), \tag{13}$$

$$\omega \le 1.$$

Now, exponent H may take values smaller than 1/2 (the subdiffusion regime) or be equal to 1/2 for $\alpha \neq 2$ (quasinormal diffusion; a non-Gaussian packet expanding at normal velocity).

In diffusion processes, a fractional time derivative appeared for the first time in [24]. A relation of this derivative to the macroscopic characteristics of a medium was discussed in [39–41]. A detailed account of fractional integrodifferentiation can be found in [49].

4. FRACTIONAL STABLE DISTRIBUTIONS

For $\omega \longrightarrow 1$, we have

$$\frac{t^{-\omega}}{\Gamma(1-\omega)} \longrightarrow \delta(t)$$

and Eq. (12) reduces to Eq. (11). Denote a solution to Eq. (12) by $p(\mathbf{x}, t; \alpha, \omega, \Gamma(\cdot))$; the limit of this solution as $\omega \longrightarrow 1$ is expressed in terms of the stable density $g_m(\mathbf{x}; \alpha, \Gamma(\cdot))$ as

$$p(\mathbf{x}, t; \alpha, 1, \Gamma(\cdot)) = t^{-m/\alpha} g_m(\mathbf{x}t^{-1/\alpha}; \alpha, \Gamma(\cdot)).$$

Equation (12) can be solved by applying the Laplace transformation in the time variable, as in [41] in the particular case of $\alpha = 2$. As a result, we obtain

$$p(\mathbf{x}, t; \alpha, \omega, \Gamma(\cdot)) = t^{-m\omega/\alpha} q_m(\mathbf{x}t^{-\omega/\alpha}; \alpha, \omega, \Gamma(\cdot)),$$

where

$$q_m(\mathbf{x}; \alpha, \omega, \Gamma(\cdot)) = \int_0^{\infty} g_m(\mathbf{x}\tau^{\omega/\alpha}; \alpha, \Gamma(\cdot)) g_1^{+}(\tau; \omega) \tau^{m\omega/\alpha} d\tau \tag{14}$$

is an m-dimensional density of the fractional stable distribution. The function g_1^{+} appearing in this definition is one-sided (vanishing on the negative half-line) stable density whose Laplace transform is

$$\int_0^{\infty} e^{-\lambda\tau} g_1^{+}(\tau; \omega) d\tau = e^{-\lambda^{\omega}}, \quad 0 < \omega \le 1.$$

As $\omega \longrightarrow 1$, the function $g_1^{+}(\tau, \omega) \longrightarrow \delta(\tau - 1)$, and the fractional stable distribution (14) reduces to the stable distribution

$$q_m(\mathbf{x}; \alpha, 1, \Gamma(\cdot)) = g_m(\mathbf{x}; \alpha, \Gamma(\cdot)).$$

A one-dimensional variant of the fractional stable distribution (14) was first obtained, as the limit of a random walk scheme, in [50] and, as a solution to fractional differential equations, in [51]; the term "fractional stable" was proposed in our paper [44]. Let us explain its origin. Let $S^{+}(\omega)$ be a random variable and $\mathbf{S}(\alpha, \Gamma(\cdot))$ be an n-dimensional random vector with distribution densities $g_1^{+}(t, \omega)$ and $g_m(\mathbf{x}; \alpha, \Gamma(\cdot))$, respectively. Consider a random vector

$$\mathbf{Z}(\alpha, \omega, \Gamma(\cdot), \mu) = \frac{\mathbf{S}(\alpha, \Gamma(\cdot))}{[S^{+}(\omega)]^{\mu}}, \tag{15}$$

which is naturally called a fractional stable vector. Its density is expressed in terms of stable densities as

$$p_{\mathbf{Z}}(\mathbf{z}; \alpha, \omega, \Gamma(\cdot), \mu) = \int_0^{\infty} dt \int_{R^m} d\mathbf{x} g_m(\mathbf{x}; \alpha, \Gamma(\cdot)) g_1^{+}(\tau; \omega) \delta\left(\mathbf{z} - \frac{\mathbf{x}}{\tau^{\mu}}\right).$$

After integrating with respect to $\mathbf{x}$, we obtain

$$p_{\mathbf{Z}}(\mathbf{z}; \alpha, \omega, \Gamma(\cdot), \mu) = \int_0^{\infty} g_m(\mathbf{z}\tau^{\mu}; \alpha, \Gamma(\cdot)) g_1^{+}(\tau; \omega) d\tau. \tag{16}$$

Comparing (16) and (14), we obtain

$$q_m(\mathbf{x}; \alpha, \omega, \Gamma(\cdot)) = p_{\mathbf{Z}}\left(z; \alpha, \omega, \Gamma(\cdot), \frac{\omega}{\alpha}\right).$$

This relation justifies the term introduced above. Formula (15), which can now be rewritten as

$$\mathbf{Z}\left(\alpha, \omega, \Gamma(\cdot), \frac{\omega}{\alpha}\right) \equiv \mathbf{Y}(\alpha, \omega; \Gamma(\cdot)) = \frac{\mathbf{S}(\alpha, \Gamma(\cdot))}{[S^{+}(\omega)]^{\omega/\alpha}},$$

yields and algorithm for simulating random vectors $\mathbf{Y}(\alpha, \omega; \Gamma(\cdot))$ with fractional stable distribution (algorithms for simulating isotropically distributed stable vectors and random variables with one-sided stable distribution are described in [48, 52, 53]).

Isotropic fractional stable distributions can conveniently be characterized by radial functions,

$$\Psi_m^{(\alpha, \omega)}(r) \equiv q_m(\mathbf{x}; \alpha, \omega, \Gamma_0(\cdot)), \quad r = |\mathbf{x}|.$$

The main properties of these functions and their representations in terms of special functions [43, 46] are given in the Appendix. One-dimensional fractional stable densities are shown in Fig. 1 (heavy lines) against stable densities (thin lines). Examples of isotropic and anisotropic two-dimensional distributions are given in Figs. 3c and 3d; the diagrams of the radial functions of multidimensional isotropic fractional stable distributions are shown in Fig. 4.

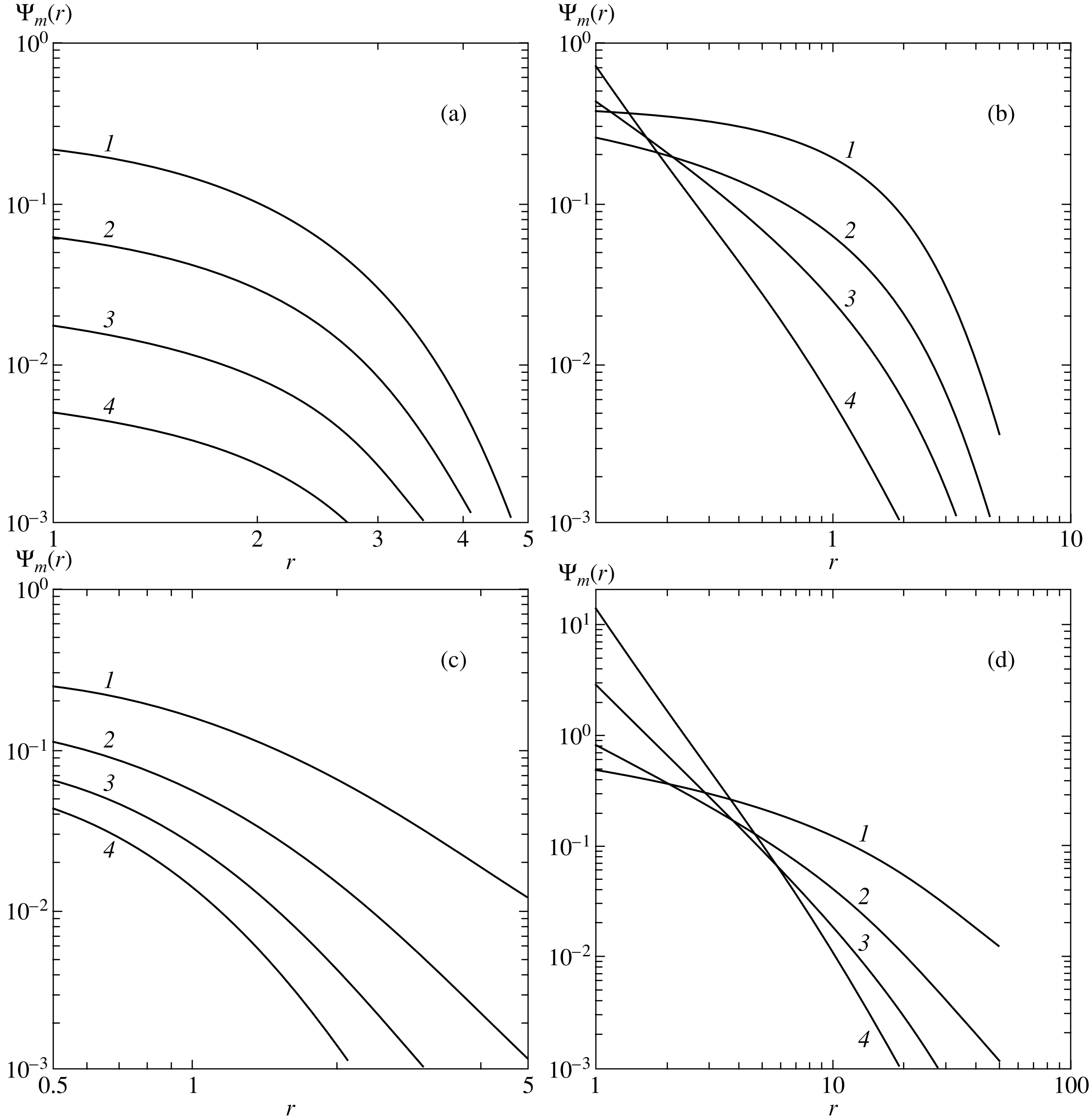

Fig. 4. Radial functions of isotropic fractional stable distributions; (a) $\Psi_m^{(2,\,1)}(r)$, (b) $\Psi_m^{(2,\,1/2)}(r)$, (c) $\Psi_m^{(1,\,1)}(r)$, and (d) $\Psi_m^{(1,\,1/2)}(r)$, $m = 1, 2, 3, 4$.

Below, we will consider specific problems that lead to fractional stable distributions.

5. APPLICATIONS OF FRACTIONAL STABLE DISTRIBUTIONS

5.1. Superdiffusion and Turbulence

The first model of superdiffusion (turbulent diffusion) appeared in the mid-1920s [1]. Turbulent diffusion is explained as due to the effect of vortices of various sizes that exist in a turbulent medium on test particles. The distance between any two such particles is substantially changed only under the action of vortices whose size is commensurable to this distance. From the viewpoint of classical diffusion, this means that the diffusion coefficient D depends on the distance $r = |\mathbf{x}|$ between these particles. This concept was realized by Richardson, who proposed the diffusion equation

$$\frac{\partial p}{\partial t} = \nabla[D(r)\nabla p(\mathbf{x}, n)]$$

with the coefficient $D(r) = ar^{\gamma}$, where $\gamma = 4/3$. This fourthirds Richardson law was later obtained theoretically by Kolmogorov and Obukhov [54, 55] as a corollary to the hypothesis about the self-similarity of locally isotropic turbulence defined by the only dimensional characteristic, the dissipation rate of turbulent energy. A solution to the equation with such a diffusion coefficient is given by

$$p(\mathbf{r}, t) = A(at)^{9/2}\exp\left\{-\frac{9r^{2/3}}{4at}\right\},$$

where A is a normalization constant. Thus, in the semiempirical Richardson theory, the width of a packet increases as $t^{3/2}$ and the distribution of a random vector proves to be non-Gaussian; certain components are noncorrelated but not independent, and the moments of any order are finite.

Monin [23] developed an alternative model. Assuming that a process is homogeneous (the diffusion coefficient is constant), he changed the diffusion operator so that the characteristic function of a centrally symmetric distribution satisfies the equation

$$\frac{\partial\tilde{p}(\mathbf{k}, t)}{\partial t} = -k^{2/3}\tilde{p}(\mathbf{k}, t).$$

Application of the inverse Fourier transformation leads to the following equation with a fractional Laplacian for the density $p(\mathbf{x}; t)$:

$$\frac{\partial p(\mathbf{x}; t)}{\partial t} = -(-\Delta)^{\alpha/2}p(\mathbf{x}; t),$$

where $\alpha = 2/3$. The solution to this equation subject to the initial condition $p(\mathbf{x}, 0) = \delta(\mathbf{x})$ is expressed in terms of the isotropic stable density $\Psi_3^{(2/3,\,1)}$ as follows:

$$p(\mathbf{x}; t) = t^{-3}\Psi_3^{(2/3,\,1)}(rt^{-3/2})$$
$$= Br^{-3}\exp\left\{\frac{2}{27}\frac{(bt)^3}{r^2}\right\}W_{-3/2,\,1/6}\left(\frac{4}{27}\frac{(bt)^3}{r^2}\right),$$

where B is a normalization constant and $W_{\mu,\,\nu}(x)$ is a Whittaker function [56].

These calculations show that, at the center of a three-dimensional diffusion cloud, the concentration decreases with time as $t^{-9/2}$, the linear dimensions of the cloud increase as $t^{3/2}$ (just as in the Richardson theory), the cloud appears more spread out than in the classical diffusion, its density is characterized by the asymptotics $r^{-11/3}$, and all the even-order moments are infinite.

5.2. Percolation and Subdiffusion

The concept of percolation was introduced in [57] to denote a certain class of phenomena (permeability of porous materials, and so on), which were supposed to be, in a certain sense, opposite to the diffusion phenomena. Percolation denoted a regular flow of a liquid (electric current, etc.) through a random medium, whereas diffusion was associated with random walks of particles in a homogeneous (in the mean) medium. The presently developed model of anomalous diffusion allows one to describe ordinary diffusion, diffusion in a turbulent medium (enhanced diffusion, or superdiffusion), and percolation (retarded diffusion, or subdiffusion) within a single formalism.

The simplest model of percolation is based on a two-dimensional square lattice with $L \times L$ nodes. Each node may be either conducting with probability p or nonconducting with probability $q = 1 - p$, independently of other nodes. This algorithm is applied to all the nodes of the lattice. For small p, one can observe only separate clusters that do not connect the opposite sides of the lattice. As p increases, the probability of percolation increases and there appear large clusters that connect the opposite sides of the lattice (contracting clusters). Without going too deeply into details, we note that, along with the contracting clusters, there exist dangling ends (dead ends), such that one cannot reach the opposite side when moving along these dead-end branches. If we take a conducting cluster by its ends, (mentally) stretch it into a straight line along the x axis, and place the dangling ends perpendicular to this axis, we obtain a comb structure. This structure has a remarkable property: an ordinary diffusion motion of a particle along its lines leads to the retarded diffusion of the x coordinate of the particle. Indeed, if we neglect that a dangling end is bounded, then the return time of the particle that diffuses along this line, i.e., the waiting time of the x coordinate in a trap, proves to be distributed by the law $g_1^+(t, 1/2)$. The diffusion equation for the x coordinate,

$$\frac{\partial^{1/2}p(x, t)}{\partial x^{1/2}} = D\frac{\partial^2 p(x, t)}{\partial x^2} + \frac{1}{\sqrt{\pi t}}\delta(x),$$

was first derived in [58]; the solution to this equation [59] is expressed via the fractional stable distribution

$$p(x, t) = (Dt^{1/2})^{-1/2}\Psi_1^{(2,\,1/2)}(x(Dt^{1/2})^{1/2}).$$

In [15], a large list of phenomena is presented that can be interpreted within the percolation model: hopping in semiconductors, gelation in polymers, localization of electrons in disordered potentials, the quantum Hall effect, Josephson junctions, gas–liquid transitions in colloids, plasma transfer in stochastic magnetic fields, and so on, up to forest fires and epidemic spreading.

5.3. Fractal Random Walks in Media with Memory

The so-called Montroll–Weiss problem—determining the distribution of the coordinates $p(\mathbf{x}, t)$, $\mathbf{x} \in R^m$, of a particle that performs instantaneous random jumps $\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_j, \ldots \in R^m$, separated by random time intervals T_1, $T_1 + T_2$, …, $T_1 + T_2 + T_j$, …, $T_i \in R_+^1$—is very important from the viewpoint of applications. The random variables $\mathbf{R}_i$ and T_i are independent of each other and among themselves, and their densities $p(\mathbf{x})$ and $q(t)$ are independent of time and coordinates, respectively. Numerous examples of applications of this problem to specific physical and biological systems are given in [14–16, 30–32, 39, 60]. Here, we only consider a special variant of this model when the densities $p(\mathbf{x})$ and $q(t)$ are characterized by the power-law asymptotics

$$\int\limits_{|\mathbf{x}|>r} p(\mathbf{x})d\mathbf{x} \sim Ar^{-\alpha}, \quad r \longrightarrow \infty, \quad 0 < \alpha < 2$$

and

$$\int\limits_t^\infty q(\tau)d\tau \sim Bt^{-\omega}, \quad t \longrightarrow \infty, \quad 0 < \omega < 1.$$

Both these conditions are interpreted in terms of fractal kinetics: the first condition describes self-similar inhomogeneity of a medium, and the second describes a special kind of memory (the power-law distribution implies the absence of characteristic scales, which suggests self-similarity).

The Fourier–Laplace transformation of the required distribution,

$$p(\mathbf{k}, \lambda) = \int\limits_{R^m} d\mathbf{x} \int\limits_0^\infty dt \exp(i\mathbf{k}\cdot\mathbf{x} - \lambda t)p(\mathbf{x}, t),$$

$$\mathbf{k} \in R^m,$$

leads to the following transform for a particle whose history begins with its waiting at the origin of coordinates ($\mathbf{x} = 0$, $t = 0$):

$$p(\mathbf{k}, \lambda) = \frac{1 - q(\lambda)}{\lambda[1 - p(\mathbf{k})q(\lambda)]}.$$

Henceforth, we assume that the angular distribution of jumps is isotropic, so that $p(\mathbf{k})$ depends only on the modulus $|\mathbf{k}| = k$.

For large times, when a particles executes many jumps and the spatial distribution of probabilities becomes wide, the density

$$p(\mathbf{x}, t) = i^{-1}(2\pi)^{-m-1}\int\limits_{R^m} d\mathbf{k}\int\limits_L d\lambda \exp(i\mathbf{k}\cdot\mathbf{x} + \lambda t)p(\mathbf{k}, \lambda)$$

is determined by the behavior of the transform $p(\mathbf{k}, \lambda)$ in the region of small k and λ, where [42]

$$1 - p(k) \sim A'k^\alpha, \quad k \longrightarrow 0,$$

$$A' = 2^{-\alpha}A\frac{\Gamma(m/2)\Gamma(1-\alpha/2)}{\Gamma(m/2+\alpha/2)},$$

$$1 - q(\lambda) \sim B'\lambda^\omega, \quad \lambda \longrightarrow 0, \quad B' = \Gamma(1-\omega)B.$$

This yields the asymptotic expression

$$p^{\mathrm{as}}(\mathbf{k}, \lambda) = \frac{\lambda^{\omega-1}}{\lambda^\omega + Dk^\alpha}, \quad D = \frac{A'}{B'},$$

which can be represented as

$$p^{\mathrm{as}}(\mathbf{k}, \lambda) = \lambda^{\omega-1}\int\limits_0^\infty \exp\{-(\lambda^\omega + Dk^\alpha)y\}dy.$$

Inverting the Laplace transform,

$$p^{\mathrm{as}}(\mathbf{k}, t) = \int\limits_0^\infty dy \exp(-Dk^\alpha y)(2\pi i)^{-1}$$

$$\times\int\limits_L d\lambda \lambda^{\omega-1}\exp(\lambda t - \lambda^\omega y),$$

integrating by parts,

$$\int\limits_L d\lambda\lambda^{\omega-1}\exp(\lambda t - \lambda^\omega y)$$

$$= -(\omega y)^{-1}\int\limits_L \exp(\lambda t)d\exp(-\lambda^\omega y)$$

$$= t(\omega y)^{-1}\int\limits_L \exp(-\lambda^\omega y + \lambda t)d\lambda,$$

and changing the variable,

$$s = y^{1/\omega}\lambda,$$

we obtain

$$p^{\mathrm{as}}(\mathbf{k}, t) = \omega^{-1}t\int\limits_0^\infty dy \exp(-Dk^\alpha y)y^{-1-1/\omega}$$

$$\times\left[(2\pi i)^{-1}\int\limits_S \exp(sy^{-1/\omega}t - s^\omega)ds\right].$$

The square brackets contain a one-sided stable density $g_1^+(ty^{-1/\omega}, \omega)$:

$$p^{\text{as}}(\mathbf{k}, t) = \int_0^\infty \exp\left(-\frac{D|k|^\alpha t^\omega}{\tau^\omega}\right) g_1^{(\omega,1)}(\tau)d\tau.$$

Application of the inverse Fourier transformation yields

$$p^{\text{as}}(\mathbf{x}, t) = (Dt^\omega)^{-m/\alpha}\Psi_m^{(\alpha,\omega)}(|\mathbf{x}|(Dt^\omega)^{-1/\alpha}), \qquad (17)$$

where

$$\Psi_m^{(\alpha,\omega)}(r) = \int_0^\infty g_m^{(\alpha)}(r\tau^{\omega/\alpha}) g_1^{(\omega,1)}(\tau)\tau^{m\omega/\alpha}d\tau. \qquad (18)$$

We can see that the asymptotics of the distribution of a particle that performs fractal random walks is expressed in terms of fractional stable distributions (these distributions have been obtained in nearly the same way in the one-dimensional problem [50]).

5.4. Diffusion on a One-Dimensional Fractal Gas

In [61], a random distribution $\{X_j\} = \ldots, X_{-2}, X_{-1}, X_0, X_1, X_2, \ldots$ of point atoms on a straight line with the properties that

(1) $X_0 = 0$,

(2) $X_i < X_j$ if $i < j$, and

(3) $X_j - X_{j-1} = R_j$ are mutually independent, identically distributed random variables with the common distribution function $F(x)$

was called a one-dimensional Lorentz gas.

Choosing different distribution functions $F(x)$, one can obtain various models of a random medium. For instance, the use of the Heaviside step function

$$F(x) = \begin{cases} 0, & x < a \\ 1, & x \ge a \end{cases}$$

yields a one-dimensional deterministic lattice, and the exponential distribution

$$F(x) = 1 - \exp(-\mu x)$$

leads to independently distributed atoms (a Poisson model).

In any case, if the mean value of R is finite, in the asymptotics of large x, we obtain $\langle N(x)\rangle \propto x$ and the relative fluctuations $\Delta(x)/\langle N(x)\rangle$ tend to 0. This means that, if $f(N(x), x)$ is a smooth function of a random variable N, then $f(N(x), x) \longrightarrow f(\langle N(x)\rangle, x)$ as $x \longrightarrow \infty$; i.e., an increase in layer thickness x leads to the self-averaging property:

$$\langle f(N(x), x)\rangle \longrightarrow f(\langle N(x)\rangle, x).$$

Now, let

$$1 - F(x) \sim \frac{A}{\Gamma(1-\alpha)}x^{-\alpha}, \quad x \longrightarrow \infty, \quad \alpha < 1.$$

Then, the mean distance between atoms is infinite, whereas the real distances are finite for any realization of the random medium. The infinity of the mean value of R leads to the situation when, in all scales, voids will alternate with clusters, i.e., to intermittent behavior. The application of the generalized limit theorem leads to the following result:

$$\sum_{i=1}^n W(i, x) \sim \int_0^z w_\alpha(z)dz, \quad x \longrightarrow \infty,$$

where $z = n/\langle N(x)\rangle$, and

$$w_\alpha(z) = \frac{z^{-1-1/\alpha}}{\alpha\Gamma(1+\alpha)} g_1^+\left(\frac{z^{-1/\alpha}}{\Gamma(1+\alpha)}; \alpha\right).$$

One can easily see that this Lorentz-gas model possesses the following properties:

1. All the atoms are equivalent, and all processes $N(x)$ with different initial conditions are statistically equivalent: $N \overset{d}{=} N(x)$.

2. The averaged (over an ensemble) number of particles increases with the layer thickness, measured from one of the atoms, by the power law

$$\langle N(x)\rangle \sim N_1 x^\alpha, \quad 0 < \alpha < 1.$$

3. Relative fluctuations of the number of atoms in this layer do not decrease as the layer thickness increases but tend to a constant value.

These properties give grounds for calling the structure obtained a stochastic fractal or a fractal gas—a self-similar (in the probabilistic sense) set of fractal dimension α. The fractal gas satisfies the relation

$$\langle f(N(x), x)\rangle \sim \int_0^\infty f(N_1 x^\alpha(z, x)) w_\alpha(z)dz$$

as $x \longrightarrow \infty$. It indicates the absence of the self-averaging property over fractal structures, which is the main difference between random walks on fractals and random walks in a regular medium.

Let us construct in a similar way a random set of points $\{T_j\}$ on the positive time half-axis. This set characterizes random moments of instantaneous jumps of a randomly walking particle from one atom to another, neighboring, atom. The random time intervals between these moments are assumed to be independent and identically distributed with the distribution function $Q(t)$. If

$$Q(t) = 1 - \exp(-\mu t), \quad \mu > 0,$$

the set $\{T_i\}$ forms a homogeneous Poisson flow. This means that the probability that a particle performs a jump in the interval $(t, t + dt)$ is independent of the moment of its preceding jump; in other words, a particle does not have memory. In all the other cases, one speaks of a particle with memory, and, if

$$1 - Q(t) \sim \frac{B}{\Gamma(1-\omega)} t^{-\omega}, \quad t \longrightarrow \infty, \quad \omega < 1,$$

one speaks of a particle with fractal memory. All of what has been said above concerning ensemble $\{X_i\}$ is also valid for ensemble $\{T_i\}$, including the averaging rule.

Suppose that a particle performs symmetric random walks: it reaches the left or the right atom with the same probability. When the number of jumps is fixed and large, the probability distribution with respect to the numbers of atoms is close to the normal distribution. Averaging (by the relation given above) this distribution over the coordinates of atoms and the jump moments, we can arrive at the expression

$$p(x, t) = (ct)^{-\omega/2\alpha} \Xi^{(\alpha, \omega)}((ct)^{-\omega/2\alpha} x),$$

where

$$\Xi^{(\alpha, \omega)}(x) = \int_0^\infty \Psi_1^{(2, \omega)}(xy^{-\alpha}) \overset{+}{g}_1(y; \alpha) y^{-\alpha} dy.$$

For $\alpha \longrightarrow 1$ and $\omega \longrightarrow 1$, this distribution takes the form of a Gaussian density, which corresponds to the normal diffusion in a regular medium.

A comparison of the spatial distribution of a particle performing random walks on a fractal with a solution to the equation of fractal random walks (Fig. 5) shows that, in the general case, the latter equation cannot be interpreted as an equation describing random walks on fractals: in the former case, a diffusion packet spreads as $t^{\omega/2\alpha}$, whereas, in the latter case, it spreads as $t^{\omega/\alpha}$, i.e., much faster. In the case of random walks on fractals, the exponent $H = \omega/2\alpha$ belongs to the interval (0, 1/2), and there is no superdiffusion regime. The reason for this difference is illustrated in Fig. 6: a particle that performs fractal random walks may reach large distances after leaving an atom, whereas, in the case of random walks, it may be trapped between neighboring clusters and perform a large number of transitions between them.

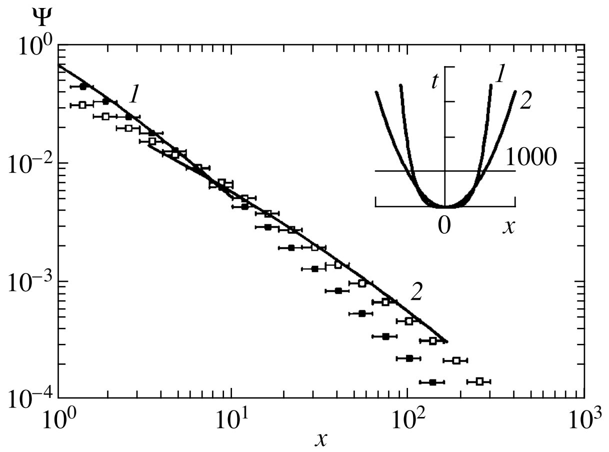

Fig. 5. Distribution density for a particle that performs (*1*) one-dimensional random walks and (*2*) fractal random walks. The exponents are $\alpha = 0.5$ and $\omega = 0.25$; the dots indicate the results obtained by the Monte Carlo method.

We should stress, however, that, in both cases, the results are based on the fractional stable distribution.

5.5. Multiple Small-Angle Scattering on Fractals

The situation changes when a particle moves along a given direction. Such a problem is faced, for example, in the case of multiple small-angle scattering of particles. In the small-angle approximation, the coordinate x plays the role of time; the direction of a particle is characterized by a small two-dimensional vector $\boldsymbol{\theta}$; and the angular distribution of particles after a large fixed number of collisions is given either by a two-dimensional normal distribution, if the second moment of the single-scattering angle is finite, or, otherwise, by a two-dimensional stable distribution. The averaging over the configurations of atoms in a fractal with fractal dimension ω leads to the angular distribution (see (17) and (18))

$$p(\boldsymbol{\theta}, x) = (Dt^{\omega})^{-2/\alpha} \Psi_2^{(\alpha, \omega)}(\theta (Dt^{\omega})^{-1/\alpha}).$$

When $\alpha = 2$ and $\omega = 1$, this formula reduces to the Fermi distribution, which is well known in the theory of multiple scattering. The Molière and Landau theories for the angular and energy distributions of particles are generalized similarly to the case of fractal media [62, 63]. This method was also used for calculating the transmission of a light beam through a randomly inhomogeneous gravitational field created by a fractal distribution of point galaxies [39].

5.6. Anomalous Transfer in Amorphous Semiconductors

The problem of multiple scattering of particles with a power-law distribution of their mean-free paths is similar to the model of another process, propagation of

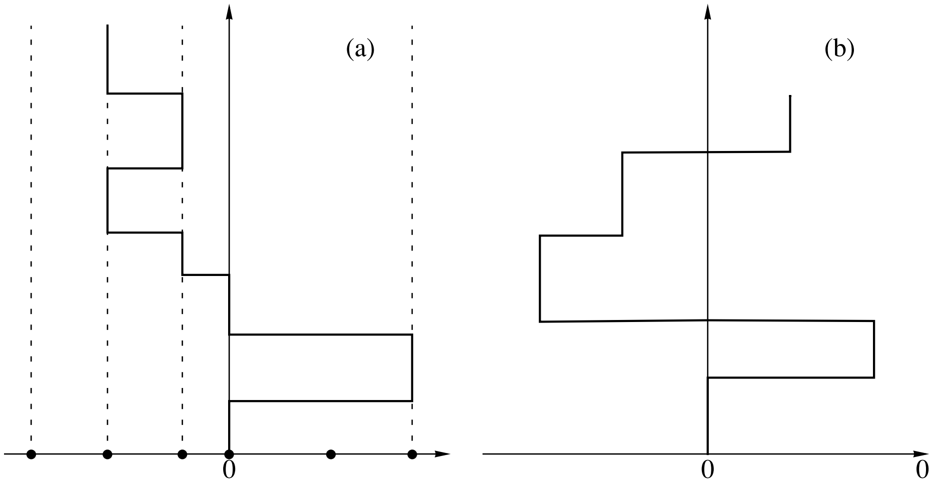

Fig. 6. (a) Trajectories of a particle that performs random walks on a fractal and (b) fractal random walks. The exponents are $\alpha = 0.5$ and $\omega = 1$.

a charge-carrier packet through materials with traps, i.e., with energy levels with a power-law distribution of the capture time.

A power-law distribution of the waiting time of particles in traps can be derived within the following model, which was first proposed in [64] (see also [20, 65]). Charge carriers move in the positive direction of the x axis, which is specified by the direction of the applied field. There are randomly spaced metal inclusions (islands) on the axis, and the charge carriers pass from one island to another by tunneling. In the quasiclassical approximation, the distribution of the random waiting time τ of a charge carrier at a given node is given by the distribution function

$$P\{\tau > t\} = 1 - \exp\left(-\frac{t}{\tau_0}\right),$$

where the mean time τ_0 is related to the distance d to the nearest (in the direction of the applied field) node by the Sommerfeld–Bethe formula

$$\tau_0 = \beta[\exp(\gamma d) - 1],$$

where γ is a positive constant and β is inversely proportional to the strength of the applied field. Since the position of the next node is random, time τ_0 is random. Under the assumption that the distribution of inclusions over the axis follows the Poisson law,

$$P\{d < x\} = 1 - \exp\frac{x}{d_0},$$

the distribution density of the expectation time of a carrier at a node, averaged over τ_0, i.e., the unconditional distribution density, is given by

$$p_\tau(t) = \int_0^\infty p_\tau(t|t')p_{\tau_0}(t')dt'$$

$$= \omega\left(\frac{\beta}{t}\right)^{\omega+1}\int_0^\infty \frac{\xi^\omega \exp(-\xi)d\xi}{(1+\xi\beta/t)^{\omega+1}} \sim \omega\Gamma(1+\omega)\beta^{1+\omega}t^{-\omega-1},$$

$$t \longrightarrow \infty.$$

When $x \gg d_0$, one can neglect the fluctuations of N; this results in an expression for a linear concentration that contains a one-sided stable density

$$n(x, t) \sim \frac{n_0}{\omega}C(\omega)t^{-1/\omega-1}g_1^+(C(\omega)tx^{-1/x}; \omega), \quad t \longrightarrow \infty,$$

where

$$C(\omega) = \left[\frac{B\Gamma(1-\omega)}{d_0}\right]^{-1/\omega}.$$

5.7. Diffusion of Cosmic Rays

The propagation of cosmic rays through the interstellar medium represents a very important problem for understanding the origin and acceleration of these rays. The stochastic character of this process is mainly determined by the scattering of charged particles by random inhomogeneities of the magnetic field. Naturally, it is impossible to directly measure the inhomogeneous time-evolving interstellar magnetic field. At the same time, one cannot neglect the influence of this field on cosmic rays. A standard approach that presumes the small-scale character of inhomegeneities leads to a normal diffusion model with a linear increase in the mean-square displacement of a diffusing particle and to a non-Gaussian form of the diffusion packet [66, 67]. A characteristic feature of this model is the presence of certain scales above which the field can be assumed homogeneous. However, there are certain indications that the magnetic field may have a multiscale hierarchical structure [68]. In this case, a more appropriate model is that of anomalous diffusion in which regions with high magnetic field serve as traps, while the fractality reflects its shaggy structure, i.e., the spatial intermittency.

The fractional stable distributions that arise in this model were successfully applied in [69, 70] to describe an experimentally observed kink in the primary spectrum of cosmic rays and to explain the difference in spectral indices for protons and other nuclei for $E \sim 10^2$–10^5 GeV per nucleon. In this case, the following equation for the propagator (Green's function) $G(\mathbf{r}, t, E, E_0)$ of cosmic rays with energy E was used:

$$\frac{\partial G}{\partial t} = -D(E, \alpha, \beta)\left(\frac{\partial}{\partial t}\right)^{1-\omega}$$
$$\times (-\Delta)^{\alpha/2} G(\mathbf{r}, t, E, E_0) + \delta(\mathbf{r})\delta(t)\delta(E - E_0).$$

The diffusion coefficient D is represented as

$$D = D_0(\alpha, \omega) R^{\delta},$$

where R is the geomagnetic rigidity of particles and $\delta > 0$. Using the solution

$$G(\mathbf{r}, E, t; E_0) = \delta(E - E_0)$$
$$\times (D(E_0, \alpha, \omega)t^{\omega})^{-3/\alpha} \Psi_3^{(\alpha,\,\omega)}(r(D(E_0, \alpha, \omega)t^{\omega})^{-1/\alpha})$$

of this equation for a point magnetic source with a density of

$$S(\mathbf{r}, t, E) = S_0 E^{-p} \delta(\mathbf{r})\delta(t),$$

we established in [69, 70] that the density of particles of cosmic radiation,

$$N(\mathbf{r}, t, E) = S_0 E^{-p} (D(E, \alpha, \omega)t^{\beta})^{-3/\alpha}$$
$$\times \Psi_3^{(\alpha,\,\omega)}(r(D(E, \alpha, \omega)t^{\omega})^{-1/\alpha}),$$

at low and high energies is characterized by the following power-law asymptotics:

$$N(\mathbf{r}, t, E) \propto E^{-p+\delta}, \quad E \longrightarrow 0,$$
$$N(\mathbf{r}, t, E) \propto E^{-p-\delta}, \quad E \longrightarrow \infty.$$

In other words, in the case of an instantaneous source, the spectral exponent decreases by 2δ when passing from low to high energies both in the subdiffusion ($\omega < \alpha/2$) and superdiffusion ($\omega > \alpha/2$) regimes; i.e., the spectrum of cosmic rays exhibits a kink.

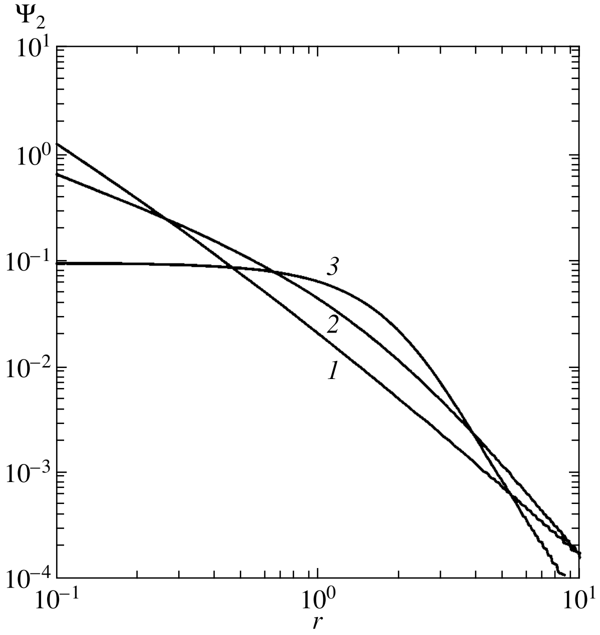

Fig. 7. Radial function of a two-dimensional fractional stable isotropic distribution with $H = \beta/\alpha = 2/3$ and $\alpha =$ (*1*) 1/2, (*2*) 1, and (*3*) 3/2.

5.8. Chaotic Dynamics

One of directions in the study of stochastic properties of the chaotic dynamics of Hamiltonian systems that arises under nonlinear resonance between periodic variations of the Hamilton function and the motion of phase points along close isoenergetic orbits is based on a fractional version of the Kolmogorov–Fokker–Planck equation [71–74]. A particular type of this equation was considered in [75]. Following [72], the author of [75] used the relation

$$\langle |X^2| \rangle = 2Dt^{2\omega/\alpha}$$

to describe the asymptotics of its solution; this relation is associated with the Hurst exponent $H = \omega/\alpha$ and the Hausdorff fractal dimension $1/H$ of chaotic trajectories.

However, in fact, $\langle |X^2| \rangle = \infty$ for any t (except the case $\alpha = 2$), and a correct analysis requires the use of another measure for the distribution width $\Delta(t)$, for example, the size of a domain containing fixed probability. The substitution of $\Delta(t)$ for $\langle |X^2| \rangle$ gives meaning to the above relations (instead of "infinity is proportional to $t^{2\omega/\alpha}$," we have a quite meaningful assertion) and allows one to apply fractional stable laws to studying the form of distributions using experimental data about the Hurst exponent. In [76, 77], it was shown that $H = 0.6$–$0.75 \approx 2/3$ for currents in tokamaks in the

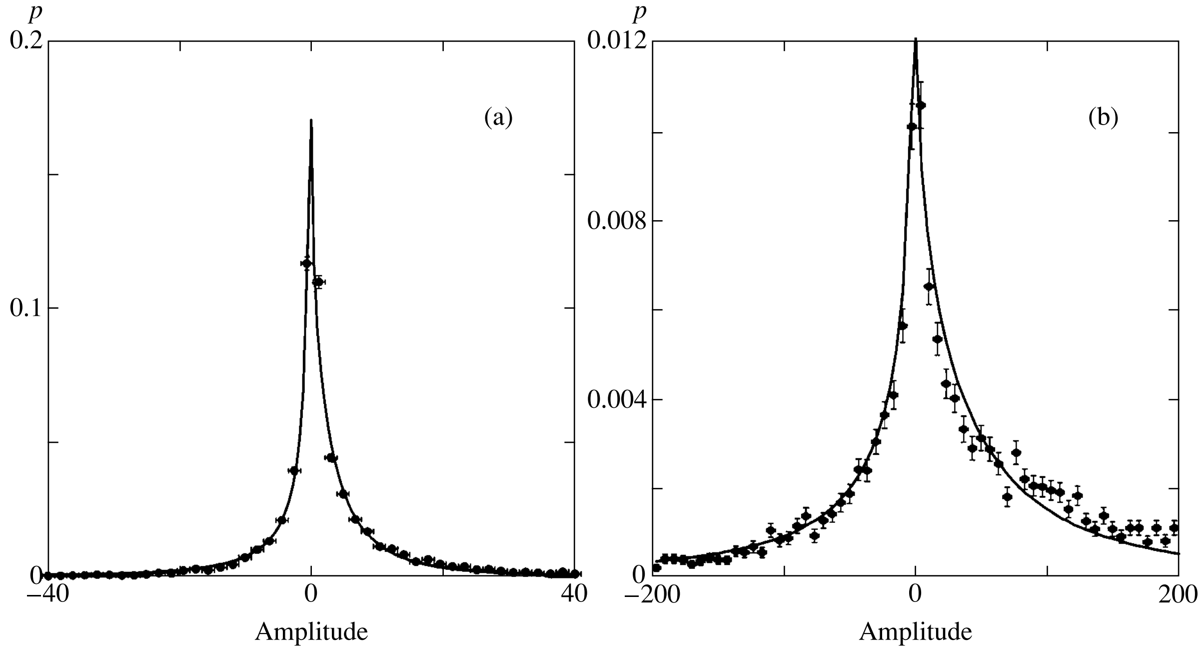

Fig. 8. Distribution of fluctuations of a stream of particles in plasma. The dots correspond to the experimental results of [79–81] and the curves represent their approximation by a fractional stable distribution with the parameters (a) $\alpha = 1.2$, $\omega = 0.05$, $\theta = 0.06$, $a = 70$, and $b = 3$ and (b) $\alpha = 1.1$, $\omega = 0.35$, $\theta = 0.05$, $a = 4$, and $b = 0.5$.

regime of self-organized criticality. Obviously, this parameter does not uniquely define the form of a distribution (Fig. 7). To specify a distribution that corresponds to experiment, one has to introduce one more independent parameter.

5.9. Fluctuations of Turbulent Flows in a Plasma

Fractional stable distributions proved appropriate for describing local fluctuations of a stream of particles in a magnetoactive plasma [78–80]. Experimental investigations of the characteristics of turbulent flows have been carried out on an L-2M stellarator. This is a toroidal thermonuclear installation with the major radius of the torus 100 cm; the mean radius of plasma is 11.5 cm, the mean density of plasma is $(1.3–1.8) \times 10^{13}$ cm^{-3}, and the central temperature of electrons is 0.6–1.0 keV; the working gas is hydrogen [80]. Turbulent flows are measured in the edge of the plasma, where the electron density is $(1–2) \times 10^{12}$ cm^{-3}, the electron temperature is 30–40 eV, and the relative level of density fluctuations in the outer regions of the plasma is $(\delta n/n)_{\text{out}} = 0.2–0.25$. A local turbulent stream of particles arising due to the development of the whole set of instabilities existing in the edge plasma of the L-2M stellarator is calculated by the formula $\Gamma = \delta n_e \delta v_\gamma$, where δn_e are the fluctuations of the plasma density,

$$\delta v_\gamma = \frac{c\delta E_\Theta}{B}$$

are the fluctuations of the radial velocity, and δE_Θ are the fluctuations of the poloidal electric field [79, 80].

To approximate the distributions obtained, we used fractional stable distributions. In Fig. 8, the densities of the amplitude of a local fluctuating stream of particles in a magnetoactive plasma are approximated by asymmetric fractional stable distributions in the C form:

$$p(x) = a^{-1} q_1(ax - b; \alpha, \omega, \theta).$$

Both a visual juxtaposition and the application of the χ^2-criterion show quite satisfactory agreement, which can be considered as a motivation for further investigations in this direction. However, for symmetric distributions, there may be other statistical interpretations [81].

6. CONCLUSIONS

An extension of the class of stable random variables by introducing their ratios would seem to be a mere formality that has no practical interest due to the lack of simple analytic expressions for their densities, if it were not the two essential circumstances. First, fractional stable distributions arise in a random walk model, or, more precisely, in a model of a jump Markov process in which the lengths of instantaneous jumps and the time intervals between them are characterized by distributions with heavy power-law tails. The second circumstance is associated with the fact that, while a stable distribution arises when solving an equation of the type

$$\frac{\partial p(x,t)}{\partial t} = L(\alpha, \Gamma(\cdot))p(x,t) + s(x,t),$$

where $L(\alpha, \Gamma(\cdot))$ is an operator containing a fractional derivative of order α with respect to the coordinates, a fractional stable distribution is associated with the solution of equations that have fractional time derivatives:

$$\frac{\partial^{\omega} p(x,t)}{\partial t^{\omega}} = L(\alpha, \Gamma(\cdot))p(x,t) + s(x,t).$$

The presence of differential equations allows one to approximate not only the distributions themselves but also the corresponding processes and give a certain physical meaning, having its roots in the random-walk model, to the parameters α and ω. Examples of such applications are given in the last section.

Let us stress once more that the fractional stable distributions represent an extension of the class of stable laws. Their common origin is the Gauss law, which determines a special position of these distributions among all probabilistic distributions.

ACKNOWLEDGMENTS

I am grateful to N.N. Skvortsova and her colleagues from the Institute of General Physics, Russian Academy of Sciences, whose results have been used in Section 5.9; I also thank V.V. Saenko for carrying out numerical calculations and E.V. Kozhemyakina for preparing the manuscript for publication.

This work was supported by the Russian Foundation for Basic Research (project no. 03-01-00163) and the Royal Society (grant gt/fSU/JP).

APPENDIX

ISOTROPIC FRACTIONAL STABLE DENSITIES

1. The densities $\Psi_{m+2}^{(\alpha,\omega)}(r)$ and $\Psi_{m}^{(\alpha,\omega)}(r)$ are related by the formula

$$\Psi_{m+2}^{(\alpha,\omega)}(r) = -\frac{1}{2\pi r}\frac{d\Psi_{m}^{(\alpha,\omega)}(r)}{dr}.$$

2. Just as in the case of the normal distribution, the projection of an m-dimensional vector $X(t)$ onto an m'-dimensional space ($m' < m$) diffuses according to the m'-dimensional law with the same parameters α and ω.

3. In contrast to the normal case, different coordinates $X_1(t), \ldots, X_m(t)$ of an anomalously diffusing particle are not independent of each other.

4. The distribution $\Psi_{m}^{(\alpha,\omega)}(r)$ decreases as r increases; its maximal value is finite for $m < \alpha$:

$$\Psi_{m}^{(\alpha,\omega)}(0) = \frac{\Gamma(1+m/\alpha)\Gamma(1-m/\alpha)}{(4\pi)^{m/2}\Gamma(1+m/2)\Gamma(1-m\omega/\alpha)}.$$

In particular,

$$\Psi_{1}^{(\alpha,\omega)}(0) = \frac{\operatorname{cosec}(\pi/\alpha)}{\alpha\Gamma(1-\omega/\alpha)}.$$

5. If $\alpha = 2$ and $\omega < 1$, then

$$\Psi_{1}^{(2,\omega)}(0) = \left[2\Gamma\left(1-\frac{\omega}{2}\right)\right]^{-1},$$

$$\Psi_{2}^{(2,\omega)}(r) \sim [2\pi\Gamma(1-\omega)]^{-1}|\ln r|, \quad r \longrightarrow 0,$$

and, for $m \geq 3$, we have

$$\Psi_{m}^{(2,\omega)}(r) \sim (4\pi)^{-m/2}\frac{\Gamma(m/2-1)}{\Gamma(1-\omega)}\left(\frac{r}{2}\right)^{-(m-2)}, \quad r \longrightarrow 0.$$

6. At large distances, we have

$$\Psi_{m}^{(2,\omega)}(r) \sim (4\pi)^{-m/2}(2-\omega)^{-1/2}\omega^{[(m+1)\omega/2-1]/(2-\omega)}$$
$$\times\left(\frac{r}{2}\right)^{-m(1-\omega)/(2-\omega)}\exp\left\{-(2-\omega)\omega^{\omega/(2-\omega)}\left(\frac{r}{2}\right)^{2/(2-\omega)}\right\}.$$

When $\omega = 1$, this formula is exact.

7. In the one-dimensional case with $\alpha = 2$, we have

$$\Psi_{1}^{(2,\omega)}(r) = \omega^{-1}r^{-1-2/\omega}g_1^{+}(r^{-2/\omega}; \omega/2).$$

In particular,

$$\Psi_{1}^{(2,1)}(r) = \frac{1}{\sqrt{4\pi}}\exp\left(-\frac{r^2}{4}\right)$$

for $\omega = 1$, and

$$\Psi_{1}^{(2,2/3)}(r) = \frac{1}{2\pi}\sqrt{r}K_{1/3}\left(\frac{2r^{3/2}}{\sqrt{27}}\right)$$

for $\omega = 2/3$.

8. For $\alpha = 1$ and $\omega = 1/2$, a fractional stable distribution of any dimension is expressed in terms of the incomplete gamma function:

$$\Psi_{m}^{(1,1/2)}(r) = \frac{2}{\sqrt{\pi}}\frac{\Gamma((m+1)/2)}{(4\pi)^{(m+1)/2}}$$
$$\times\exp\frac{r^2}{4}\Gamma\left(1-\frac{m+1}{2}, \frac{r^2}{4}\right).$$

For odd dimensions m, we have

$$\Psi_m^{(1,\,1/2)}(r) = \frac{2}{\sqrt{\pi}}\frac{\Gamma((m+1)/2)}{(4\pi)^{(m+1)/2}}\left(\frac{r^2}{4}\right)^{\mu}\exp\frac{r^2}{4}E_{(m+1)/2}\left(\frac{r^2}{4}\right),$$

where

$$\mu = 1 - \frac{m+1}{2}.$$

9.

$$\Psi_m^{(2/3,\,1)}(r) = \frac{\Gamma(m/2+1/3)\Gamma(m/2+2/3)}{2\sqrt{3\pi}^{m}\Gamma(5/6)\Gamma(7/6)}r^{-m}$$

$$\times\exp\left(\frac{2}{27}r^{-2}\right)W_{-m/2,\,1/6}\left(\frac{4}{27}r^{-2}\right).$$

10. An asymptotic series expansion is given by

$$\Psi_m^{(\alpha,\,\omega)}(r) = \frac{\alpha}{2(4\pi)^{m/2}}$$

$$\times\sum_{n=1}^{\infty}(-1)^{n-1}\frac{\Gamma((m+\alpha n)/2)}{\Gamma(1-\alpha n/2)\Gamma(1+n\omega)}\left(\frac{r}{2}\right)^{-n\alpha-m}.$$

11. The Mellin transform of $\Psi^{(\alpha,\,\omega)}(r)$ is expressed in terms of a gamma function:

$$\bar{\Psi}_m^{(\alpha,\,\omega)}(s) \equiv \int_0^{\infty}\Psi_m^{(\alpha,\,\omega)}(r)r^{s-1}dr$$

$$= \frac{2^s\Gamma(1-(m-s)/\alpha)\Gamma(s/2)\Gamma((m-s)/\alpha)}{\alpha(4\pi)^{m/2}\Gamma(1-(m-s)\omega/\alpha)\Gamma((m-s)/2)}.$$

12. For $\alpha = 2$, there exist even moments

$$\int_{-\infty}^{\infty}\Psi_m^{(\alpha,\,\omega)}(r^{2n})d\mathbf{x} = \frac{\Gamma(n+m/2)\Gamma(n+1)}{\Gamma(m/2)\Gamma(n\omega+1)}(4Dt^{\omega})^n.$$

13. A representation in terms of the H functions (the Fox functions) [33] is given by

$$\Psi_m^{(\alpha,\,\omega)}(r) = \frac{\omega}{2}(4\pi)^{-m/2}\left(\frac{2}{r}\right)^{m+\alpha}H_{32}^{12}$$

$$\times\left(\left(\frac{2}{r}\right)^{\omega}\middle|\begin{matrix}(-1,\,1/\alpha)(1-(\alpha+m)/2,\,\omega/2)(1-\alpha/2,\,1/2)\\(0,\,1/\alpha)(-1,\,1/\alpha)\end{matrix}\right),$$

$$\alpha < 1,$$

$$\Psi_m^{(\alpha,\,\omega)}(r) = (\alpha r\sqrt{\pi})^{-m}$$

$$\times H_{23}^{21}\left(\frac{r}{2}\middle|\begin{matrix}(1,\,1/\alpha)(1,\,\omega/\alpha)\\(1,\,1/\alpha)(m/2,\,1/2)(1,\,1/2)\end{matrix}\right),$$

$$1 \le \alpha < 2,$$

$$\Psi_m^{(2,\,\omega)}(r) = (2r\sqrt{\pi})^{-m}H_{12}^{20}\left(\frac{r}{2}\middle|\begin{matrix}(1,\,\omega/2)\\(1,\,1/2)(m/2,\,1/2)\end{matrix}\right).$$

REFERENCES

1. L. F. Richardson, Proc. R. Soc. London **110**, 709 (1926).
2. L. M. Biberman, Zh. Éksp. Teor. Fiz. **17**, 416 (1947).
3. Yu. Yu. Abramov, A. M. Dykhne, and A. P. Naportovich, Zh. Éksp. Teor. Fiz. **56**, 654 (1969) [Sov. Phys. JETP **29**, 358 (1969)].
4. B. Berkowith and H. Scher, Phys. Rev. E **57**, 5858 (1998).
5. R. Kimmich, *Tomography, Diffusometry, Relaxometry* (Springer, Berlin, 1997).
6. P. G. de Gennes, *Scaling Concepts in the Physics of Polymers* (Cornell Univ. Press, Ithaca, N.Y., 1979; Mir, Moscow, 1982).
7. M. Porto, A. Bunde, S. Havlin, and H. E. Roman, Phys. Rev. E **56**, 1667 (1997).
8. W. Young, A. Pumir, and Y. Pomeau, Phys. Fluids A **1**, 462 (1989).
9. W. D. Luedtke and U. Landmann, Phys. Rev. Lett. **82**, 3835 (1999).
10. J. Bodurka, R.-O. Seitter, R. Kimmich, and A. Gutsze, J. Chem. Phys. **107**, 5621 (1997).
11. M. Sahimi, Phys. Rep. **306**, 214 (1998).
12. S. Scaufler, W. P. Schleich, and V. P. Yakovlev, Phys. Rev. Lett. **83**, 3162 (1999).
13. R. Balescu, Phys. Rev. E **51**, 4807 (1995).
14. J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
15. M. B. Isichenko, Rev. Mod. Phys. **64**, 961 (1992).
16. M. F. Shlesinger, G. M. Zaslavsky, and J. Klafter, Nature **363**, 31 (1993).
17. J. K. Jonscher, *Universal Relaxation Law* (Chelsea Dielectric Press, London, 1996).
18. B. O'Shaugnessy and I. Procaccia, Phys. Rev. Lett. **54**, 455 (1985).
19. B. B. Mandelbrot, *The Fractal Geometry of Nature* (Freeman, New York, 1983).
20. H. Scher and E. W. Montroll, Phys. Rev. B **12**, 2455 (1975).
21. E. W. Montroll and G. H. Weiss, J. Math. Phys. **6**, 167 (1965).
22. E. W. Montroll and M. Schlesinger, in *Studies in Statistical Mechanics*, Ed. by J. Leibowitz and E. W. Montroll (North-Holland, Amsterdam, 1984), Vol. 2.
23. A. S. Monin, Dokl. Akad. Nauk SSSR **105**, 256 (1955).
24. R. R. Nigmatullin, Phys. Status Solidi B **123**, 739 (1989).
25. F. Mainardi, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **38**, 20 (1995).
26. V. L. Kobelev, E. P. Romanov, Ya. L. Kobelev, and L. Ya. Kobelev, Dokl. Akad. Nauk **361**, 755 (1998) [Dokl. Phys. **43**, 484 (1998)].
27. G. M. Zaslavsky, Physica D (Amsterdam) **76**, 110 (1994).
28. C. Tsallis, J. Stat. Phys. **52**, 479 (1988).
29. K. V. Chukbar, Zh. Éksp. Teor. Fiz. **108**, 1875 (1995) [JETP **81**, 1025 (1995)].
30. *Anomalous Diffusion: From Basics to Applications*, Ed. by R. Kutner, A. Pekalski, and K. Sznajd-Weron (Springer, Berlin, 1999).

31. R. Metzler and J. Klafter, Phys. Rep. **339**, 16 (2000).
32. D. ben-Avraham and S. Havlin, *Diffusion and Reactions in Fractals and Disordered Systems* (Cambridge Univ. Press, Cambridge, 2000).
33. A. M. Mathai and R. K. Saxena, *The H-function with Applications in Statistics and Other Disciplines* (Wiley, New Delhi, 1978).
34. V. V. Uchaĭkin, Teor. Mat. Fiz. **115**, 154 (1998).
35. V. V. Uchaĭkin, Zh. Tekh. Fiz. **68** (1), 138 (1998) [Tech. Phys. **43**, 124 (1998)].
36. V. V. Uchaikin, Physica A (Amsterdam) **255**, 65 (1998).
37. V. M. Zolotarev and V. V. Uchaĭkin, Teor. Veroyatn. Primen. **44**, 176 (1999).
38. V. V. Uchaĭkin, Teor. Veroyatn. Primen. **44**, 194 (1999).
39. V. V. Uchaikin and V. M. Zolotarev, *Chance and Stability. Stable Distributions and Their Applications* (VSP, Utrecht, Netherlands, 1999).
40. V. M. Zolotarev, V. V. Uchaĭkin, and V. V. Saenko, Zh. Éksp. Teor. Fiz. **115**, 1411 (1999) [JETP **88**, 780 (1999)].
41. V. V. Uchaĭkin, Zh. Éksp. Teor. Fiz. **115**, 2113 (1999) [JETP **88**, 1155 (1999)].
42. V. V. Uchaikin, Int. J. Theor. Phys. **38**, 2377 (1999).
43. V. V. Uchaikin, Int. J. Theor. Phys. **39**, 2087 (2000).
44. V. Kolokoltsov, V. Korolev, and V. Uchaikin, Preprint No. 23/00 (Nottingham Trent Univ., 2000).
45. V. V. Uchaikin, Physica A (Amsterdam) **305**, 205 (2002).
46. V. V. Uchaikin, J. Chem. Phys. **88**, 1141 (2002).
47. V. M. Zolotarev, *One-Dimension Stable Distributions* (Nauka, Moscow, 1983).
48. G. Samorodnitzky and M. S. Taqqu, *Stable Non-Gaussian Random Processes* (Chapman and Hall, New York, 1994).
49. S. G. Samko, A. A. Kilbas, and O. I. Marichev, *Integrals and Derivatives of a Fractional Order and Some of Their Applications* (Nauka i Tekhnika, Minsk, 1987).
50. M. Kotulski, J. Stat. Phys. **81**, 777 (1995).
51. A. I. Saichev and G. M. Zaslavsky, Chaos **7**, 753 (1997).
52. M. Kanter, Ann. Prob. **3**, 697 (1975).
53. V. V. Uchaikin and G. G. Gusarov, in *Proceedings of 3rd St. Petersburg Workshop on Simulation* (St. Petersburg Univ. Press, St. Petersburg, 1999), p. 306.
54. A. N. Kolmogorov, Dokl. Akad. Nauk SSSR **30**, 299 (1941).
55. A. M. Obukhov, Izv. Akad. Nauk SSSR, Ser. Geogr. Geofiz. **4–5**, 453 (1941).
56. A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics* (Nauka, Moscow, 1967; MIT Press, Cambridge, Mass., 1975), Vol. 2.
57. S. R. Broadbent and J. M. Hammersley, Proc. Cambridge Philos. Soc. **53**, 629 (1957).
58. R. R. Nigmatullin, Phys. Status Solidi B **123**, 739 (1984).
59. V. E. Arkhincheev and É. M. Baskin, Zh. Éksp. Teor. Fiz. **100**, 292 (1991) [Sov. Phys. JETP **73**, 161 (1991)].
60. B. J. West and W. Deering, Phys. Rep. **246**, 1 (1994).
61. E. Barkai, V. Fleurov, and J. Klafter, Phys. Rev. E **61**, 1164 (2000).
62. V. V. Uchaĭkin and D. A. Korobko, Pis'ma Zh. Tekh. Fiz. **25** (11), 34 (1999) [Tech. Phys. Lett. **25**, 435 (1999)].
63. D. A. Korobko and V. V. Uchaĭkin, Uch. Zap. Ul'yanovsk. Gos. Univ., Ser. Fiz. **6**, 15 (1999).
64. J. K. E. Tunaley, J. Appl. Phys. **11**, 4783 (1972).
65. I. P. Zvyagin, *Kinetic Phenomena in Disordered Semiconductors* (Mosk. Gos. Univ., Moscow, 1984).
66. V. L. Ginzburg and S. I. Syrovatskiĭ, *The Origin of Cosmic Rays* (Nauka, Moscow, 1964; Gordon and Breach, New York, 1969).
67. V. S. Berezinsky, S. V. Bulanov, V. L. Ginzburg, *et al.*, *Astrophysics of Cosmic Rays,* 2nd ed., Ed. by V. L. Ginzburg (Nauka, Moscow, 1990; North Holland, Amsterdam, 1990).
68. A. A. Ruzmaĭkin, D. D. Sokolov, and A. M. Shukurov, *Magnetic Fields of Galaxies* (Kluwer, Dordrecht, 1988).
69. A. A. Lagutin, Yu. A. Nikulin, and V. V. Uchaikin, Nucl. Phys. B **97**, 267 (2001).
70. A. A. Lagutin and V. V. Uchaikin, Nucl. Instrum. Methods Phys. Res. B **201**, 212 (2003).
71. G. M. Zaslavsky, in *Topological Aspects and the Dynamics of Fluids and Plasmas*, Ed. by H. K. Moffatt, G. M. Zaslavsky, P. Comte, and M. Tabor (Kluwer, Boston, 1992).
72. G. M. Zaslavsky, Physica D (Amsterdam) **76**, 110 (1994).
73. G. M. Zaslavskiĭ, *Chaos in Dynamical Systems* (Nauka, Moscow, 1984; Harwood, Chur, 1985).
74. G. M. Zaslavsky, M. Edelman, H. Weiltznez, *et al.*, Phys. Plasmas **7**, 3691 (2000).
75. A. E. Milovanov, Phys. Rev. E **63**, 047301 (2001).
76. B. A. Carreras, B. van Milligen, M. A. Pedrosa, *et al.*, Phys. Rev. Lett. **80**, 4438 (1998).
77. B. A. Carreras, B. van Milligen, C. Hidalgo, *et al.*, Phys. Rev. Lett. **83**, 3653 (1999).
78. J. W. Connor, P. Burraffi, J. G. Cordey, *et al.*, Plasma Phys. Controlled Fusion **41**, 693 (1999).
79. M. A. Pedrosa, M. A. Ochando, J. A. Jimenez, *et al.*, Plasma Phys. Controlled Fusion **38**, 365 (1996).
80. G. M. Batanov, O. I. Fedyanin, N. K. Kharchev, *et al.*, Plasma Phys. Controlled Fusion **40**, 1241 (1998).
81. G. M. Batanov, V. E. Bening, V. Yu. Korolev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 143 (2001) [JETP Lett. **73**, 126 (2001)].

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 826–835.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 921–931.
Original Russian Text Copyright © 2003 by Ginzburg, Gerashchenko.

NONLINEAR PHYSICS

A Study of Stochastic Resonance in a Bistable System with Double Stochasticity

S. L. Ginzburg and O. V. Gerashchenko*
Konstantinov Institute of Nuclear Physics, Russian Academy of Sciences,
Gatchina, St. Petersburg, 188350 Russia
**e-mail: gerashch@pnpi.spb.ru*

Received December 6, 2002

Abstract—In a simple stochastic system—an overdamped Kramers oscillator with two noise sources (sources of white and dichotomic noise)—stochastic resonance is investigated theoretically and by means of analog simulation as a function of the asymmetry of the potential and the amplitude and the correlation time of the dichotomic noise. It is found that stochastic resonance is observed under slow (compared with the Kramers switching time) dichotomic noise and the signal-to-noise ratio (SNR) has a maximum for a noise amplitude equal to static bias. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In recent years, the idea of the constructive role of noise in nature has attracted the special interest of researchers. Many experimental facts from various fields of science have shown that the effect of noise, which is an inevitable concomitant of all real systems, is not always reduced to destructive and thermodynamic phenomena. It has been established that noise sources in nonlinear dynamical systems may give rise to fundamentally new operation regimes, such as noise-induced oscillations [1], that are impossible in the absence of noise. These phenomena have been called noise-induced phase transitions [2].

Investigations in recent years have clearly shown that, in nonlinear systems, noise may give rise to new, more ordered regimes and more regular structures; increase the degree of coherence and the amplification factor; etc. In particular, when useful signals pass through nonlinear systems, thermal or external noise of optimal (nonzero) intensity may improve the signal transmission. This phenomenon is called stochastic resonance. The stochastic resonance involves a group of phenomena in which the response of a nonlinear system to a weak signal appreciably increases with the intensity of noise in the system. In this case, integral characteristics such as the amplification factor and the signal-to-noise ratio (SNR) exhibit clear-cut maxima at a certain optimal noise level. The phenomenon of stochastic resonance was observed in many bistable systems: in a ring laser [3], in passive optical bistable systems [4], in electron-paramagnetic-resonance systems [5], in experiments with Brownian particles [6], in a tunnel diode [7], in SQUIDs [8], in ferromagnets and ferroelectrics [9–11], in analog electron experiments [12–20], in chemical systems [21–23], and in sociological models [24]. Investigations have shown that stochastic resonance represents an essentially generic physical phenomenon, typical of nonlinear systems, in which one can control one of the characteristic time scales of a system by means of noise. The results obtained in the study of the stochastic resonance are summarized in [25, 26].

The main body of theoretical works on the problem of stochastic resonance have been devoted to the study of signal transmission through a system with a single source of noise, usually white noise; this phenomenon is identified as the appearance of a peak on the diagram of the signal-to-noise ratio (SNR) versus the noise intensity. However, some authors investigated systems with two sources of noise, the first without and the second with a periodic signal [27, 28]. This statement of the problem is physically natural and fruitful because any real system contains internal thermal noise whose intensity is determined by temperature. Control noise, just as a signal, is applied to the system from the outside. However, to date, the signal transmission has been studied as a function of the intensity of internal noise only. On the other hand, in many cases, especially in biology, temperature (and consequently the power of internal noise) is fixed within a narrow interval. However, the amplitudes of the external noise and signal can be varied within a wide range of values. It is this kind of approach that we apply in the present paper; here, the variable parameters are the ratios of the amplitudes of the control noise and signal to the intensity of the internal noise.

We study theoretically and experimentally a two-well dynamical system with a periodic square-wave signal and two noise sources, white (internal) noise and colored dichotomic (external) noise. We will solve our problem in two limiting cases, with slow and fast colored noise for arbitrary values of the signal. In principle, one can solve the problem in an intermediate region for small signals, applying the linear-response theory [29–33]. However, it is the choice of a square-wave signal and colored dichotomic noise that allows one to solve the problem for arbitrary amplitudes of signal and noise, without restricting oneself to perturbation theory.

The main result of this paper is the discovery of a resonant or strongly nonmonotonic behavior of the signal and noise powers and the SNR in the case when the transition probabilities from one state to another become equal. This result differs from the conventional stochastic resonance and is associated with the appearance of additional symmetry in the system when the above-mentioned probabilities coincide. In the domain of existence of such symmetry, all the characteristics behave nonmonotonically.

2. THEORY

Consider the equation of an overdamped Kramers oscillator with two noise sources:

$$\begin{gathered}\frac{dx}{dt} = f(x) + F + \Delta S(t) + \sigma\xi(t) + AR(t),\\ f(x) = ax - bx^3, \quad S(t) = \pm 1,\\ \langle S(t)S(t')\rangle = e^{-\gamma|t-t'|},\\ \langle \xi(t)\xi(t')\rangle = \delta(t-t'),\\ R(t+T) = R(t) = \begin{cases}1, & 0 < t \le T/2\\ -1, & T/2 < t \le T.\end{cases}\end{gathered} \tag{1}$$

Here, $\xi(t)$ is white noise; $S(t)$ is dichotomic (telegraphic) noise; $R(t)$ is a square-wave signal; and a, b, F, σ, Δ, and A are constants.

Note that function $f(x)$ should not necessarily be chosen in the form of (1). Below, we will show that it is important that $f(x)$ is an odd function that has only one unstable fixed point and two stable points at $\pm x_0$. In the experiments, $f(x)$ will have a different form satisfying our conditions.

It is important that the characteristic times in Eq. (1) form a hierarchy. Denoting by τ the correlation time of the (real) white noise and assuming that $a \sim b$, we adopt the following time hierarchy in our problem:

$$T \gg \gamma^{-1}, \quad w_0^{-1} \gg a^{-1} \gg \tau \longrightarrow 0, \tag{2}$$

where w_0 is the characteristic rate of switchings induced by the white noise (see (6)). Note that, in certain formulas, we will retain the notation τ for the noise correlation time, assuming that the white noise corresponds to the limit $\tau \longrightarrow 0$, because this notation, first, emphasizes that the necessary condition $a^{-1} \gg \tau$ should be satisfied and, second, is important for comparing theoretical results with experiment, where τ is always finite.

Let us pass from (1) to a model of two states between which jump transitions occur. It is convenient to define energy $U_0(x)$ as

$$\begin{gathered}U_0(x) = -\frac{a}{2}x^2 + \frac{b}{4}x^4,\\ \delta U_0 = U_0(0) - U_0(\pm x_0) = \frac{a^2}{4b}, \quad x_0 = \sqrt{\frac{a}{b}},\end{gathered} \tag{3}$$

where $\pm x_0$ are stable fixed points of Eq. (1) for $F = \Delta = A = \sigma = 0$ and δU_0 is the height of the potential barrier.

Suppose that the following a priori relations hold:

$$\begin{gathered}\sigma^2 \ll \delta U_0, \quad \Delta x_0, Fx_0 \ll \delta U_0,\\ \frac{(\Delta x_0)^2}{\sigma^2 \delta U_0}, \frac{(Fx_0)^2}{\sigma^2 \delta U_0} \ll 1.\end{gathered} \tag{4}$$

The first relation implies that the system mostly stays at one of the stable fixed points (potential wells). The second and third conditions imply that the fixed points do not move when the system is subject to external actions.

Thus, we can approximate the dynamic variable $x(t)$ by a dichotomic signal,

$$x(t) = x_0 d(t), \quad d(t) = \pm 1. \tag{5}$$

In other words, $x(t)$ assumes only two values, and we have a two-state model. The same conditions (4) determine the applicability domain of the Kramers approximation for the transition probabilities between these two states, i.e., the exponential dependence of these probabilities on the ratio of the time-dependent potential energy of the system to the noise intensity:

$$\begin{gathered}\alpha(t) = w_{-1\to 1} = w_0 \exp[p + uS(t) + \nu R(t)],\\ \beta(t) = w_{1\to -1} = w_0 \exp[-p - (uS(t) - \nu R(t))],\\ p = \frac{2x_0F}{\sigma^2}, \quad u = \frac{2x_0\Delta}{\sigma^2}, \quad \nu = \frac{2x_0A}{\sigma^2},\\ w_0 = \frac{a}{\pi\sqrt{2}}\exp\left(-\frac{a^2}{2b\sigma^2}\right).\end{gathered} \tag{6}$$

The governing equation for the distribution density of a nonstationary process $d(t)$ reads

$$\begin{aligned}\frac{dP(1,t)}{dt} &= \alpha(t)P(-1,t)-\beta(t)P(1,t),\\ \frac{dP(-1,t)}{dt} &= \beta(t)P(1,t)-\alpha(t)P(-1,t).\end{aligned}\tag{7}$$

Suppose first that $\Delta = A = 0$ (there is no external action). Then, α and β in (6) are independent of time, and the nonzero eigenvalues of Eq. (7) are given by

$$\lambda_0 = \alpha+\beta = 2w_0\cosh(p).\tag{8}$$

It is clear that equilibrium is established in the system within time $1/\lambda_0$.

2.1. Slow External Noise

Now, let us switch on an external action ($A, \Delta \neq 0$) and suppose that the adiabaticity condition is fulfilled not only with respect to the signal period T but also with respect to the correlation time $1/\gamma$ of the dichotomic noise:

$$T, \gamma^{-1} \gg w_0^{-1} \gg a^{-1} \gg \tau \longrightarrow 0.\tag{9}$$

Since $w_0 \sim \lambda_0$ and equilibrium is established within time $1/\lambda_0$, $\alpha(t)$ and $\beta(t)$ slowly vary over times on the order of T and $1/\gamma$; therefore, we can speak of local equilibrium and solve Eqs. (7) in the adiabatic approximation. The solution is given by

$$\begin{aligned}P(1,t) &= \frac{\alpha(t)}{\alpha(t)+\beta(t)},\\ P(-1,t) &= \frac{\beta(t)}{\alpha(t)+\beta(t)}.\end{aligned}\tag{10}$$

One can also easily calculate the transition probability for the process $d(t)$ from state m at moment t_2 to state n at moment $t_1 > t_2$:

$$\begin{aligned}P(n,t_1|m,t_2) &= P(n,t_1)+\frac{\phi(n,t_1)\phi(m,t_1)}{P(m,t_1)}e^{-(t_1-t_2)\lambda_0(t_1)},\\ \lambda_0(t) &= \alpha(t)+\beta(t),\\ \phi(1,t) &= -\phi(-1,t) = \frac{\sqrt{\alpha(t)\beta(t)}}{\alpha(t)+\beta(t)}.\end{aligned}\tag{11}$$

$\lambda_0(t)$ and $\phi(n, t)$ in the last equation are the adiabatically slowly varying eigenvalue and the corresponding eigenfunction of Eq. (7).

Now, let us calculate a two-particle distribution function $P_2(n, t_1; m, t_2)$. According to the general theory of Markov processes, this function is given by

$$\begin{aligned}P_2(n,t_1;m,t_2) &= P(n,t_1|m,t_2)P(m,t_2)\\ &= P(n,t_1)P(m,t_2)\\ &+\frac{\phi(n,t_1)\phi(m,t_1)}{P(m,t_1)}P(m,t_2)\exp\{-(t_1-t_2)\lambda_0(t_1)\}.\end{aligned}\tag{12}$$

We are interested in the time scales

$$\frac{1}{\lambda_0} \ll T, \quad \frac{1}{\gamma} \sim |t_1-t_2|.\tag{13}$$

Then, we can replace the exponential function in (11) by the δ function to obtain the following expression for the two-particle distribution function:

$$\begin{aligned}P_2(n,t_1|m,t_2) &= P(n,t_1)P(m,t_2)\\ &+\frac{2\delta(t_1-t_2)}{\lambda_0(t_1)}\phi(n,t_1)\phi(m,t_1),\end{aligned}\tag{14}$$

where $P(n, t)$, $\lambda_0(t)$, and $\phi(n, t)$ are defined in (10) and (11). Taking into account these definitions, we can easily obtain the correlation function

$$\begin{aligned}D(t_1,t_2) &= \langle x(t_1)x(t_2)\rangle_\sigma = y(t_1)y(t_2)\\ &+\frac{8\alpha(t_1)\beta(t_1)x_0^2}{[\alpha(t_1)+\beta(t_1)]^3}\delta(t_1-t_2),\\ y(t) &= \langle x(t)\rangle_\sigma = x_0\frac{\alpha(t)-\beta(t)}{\alpha(t)+\beta(t)}.\end{aligned}\tag{15}$$

The subscript σ at the expressions in angular brackets denotes averaging over the white noise.

Now, let us average (15) over the stochastic process $S(t)$ and over the phase of the periodic signal $R(t)$. We calculate the irreducible correlator,

$$\begin{aligned}K(t_1-t_2) &= \langle D(t_1,t_2)\rangle_{S,R}\\ &-\langle y(t_1)\rangle_{S,R}\langle y(t_2)\rangle_{S,R}.\end{aligned}\tag{16}$$

From (15) and (16), we obtain

$$\begin{aligned}K(t_1-t_2) &= \langle y(t_1)y(t_2)\rangle_{S,R}-\langle y(t)\rangle_{S,R}^2\\ &+\frac{x_0^2}{w_0}\langle\cosh^{-3}[p+uS(t)+vR(t)]\rangle_{S,R}\delta(t_1-t_2),\\ y(t) &= x_0\tanh[p+uS(t)+vR(t)].\end{aligned}\tag{17}$$

The correlator (17) is easily calculated in the general case. Taking into account that $S(t) = \pm 1$ and $R(t) = \pm 1$, we obtain the following relations for an arbitrary function $f(x)$:

$$\begin{aligned} f(\Delta S(t) + AR(t)) &= \frac{1+S(t)}{2}\frac{1+R(t)}{2} f(\Delta + A) \\ &+ \frac{1-S(t)}{2}\frac{1-R(t)}{2} f(-\Delta - A) \\ &+ \frac{1+S(t)}{2}\frac{1-R(t)}{2} f(\Delta - A) \\ &+ \frac{1-S(t)}{2}\frac{1+R(t)}{2} f(-\Delta + A). \end{aligned} \tag{18}$$

Then, we easily obtain the following expression for the correlator:

$$\begin{gathered} K(t) = K^{(0)}(t) + C(p, u, \nu)\delta(t), \\ C(p, u, \nu) = \frac{x_0^2}{4w_0} \\ \times[\cosh^{-3}(p+u+\nu) + \cosh^{-3}(p-u-\nu) \\ + \cosh^{-3}(p+u-\nu) + \cosh^{-3}(p-u+\nu)], \\ K^{(0)}(t) = B_1(p, u, \nu)e^{-\gamma|t|} + B_2(p, u, \nu)\phi_0(t) \\ + B_3(p, u, \nu)e^{-\gamma|t|}\phi_0(t), \\ \phi_0(t) = \langle R(t)R(0)\rangle_R, \end{gathered} \tag{19}$$

$$\begin{aligned} B_1(p, u, \nu) &= \frac{x_0^2}{16}[\tanh(p+u+\nu) - \tanh(p-u-\nu) \\ &+ \tanh(p+u-\nu) - \tanh(p-u+\nu)]^2, \end{aligned}$$

$$\begin{aligned} B_2(p, u, \nu) &= \frac{x_0^2}{16}[\tanh(p+u+\nu) - \tanh(p-u-\nu) \\ &- \tanh(p+u-\nu) + \tanh(p-u+\nu)]^2, \end{aligned}$$

$$\begin{aligned} B_3(p, u, \nu) &= \frac{x_0^2}{16}[\tanh((p+u+\nu) + \tanh(p-u-\nu) \\ &- \tanh(p+u-\nu) - \tanh(p-u+\nu)]^2. \end{aligned}$$

The coefficients $B_i(p, u, \nu)$ defined in (19) possess the following symmetry:

$$\begin{aligned} B_1(p, 0, \nu) &= B_2(p, u, 0) \\ &= B_3(p, 0, \nu) = B_3(p, u, 0) = 0. \end{aligned} \tag{20}$$

Now, we calculate the spectral density

$$S(\omega) = \int_{-\infty}^{\infty} K(t)e^{i\omega t}dt. \tag{21}$$

Since

$$\phi_0(t) = \frac{4}{\pi^2}\sum_{k=-\infty}^{\infty}\frac{e^{-i(2k+1)\Omega t}}{(2k+1)^2}, \quad \Omega = \frac{2\pi}{T}, \tag{22}$$

from (19) we obtain

$$\begin{aligned} S(\omega) &= C(p, u, \nu) + \frac{2\gamma}{\omega^2+\gamma^2}B_1(p, u, \nu) \\ &+ \frac{8}{\pi}B_2(p, u, \nu)\sum_{k=-\infty}^{\infty}\frac{\delta(\omega-(2k+1)\Omega)}{(2k+1)^2} \\ &+ \frac{8\gamma}{\pi}B_3(p, u, \nu) \\ &\times\sum_{k=-\infty}^{\infty}[(2k+1)^2((\omega-(2k+1)\Omega)^2+\gamma^2)]^{-1}. \end{aligned} \tag{23}$$

If $\omega, \Omega \ll \gamma$, expression (23) simplifies to

$$\begin{aligned} S(\omega) &= C(p, u, \nu) + \frac{2}{\gamma}(B_1(p, u, \nu) + B_3(p, u, \nu)) \\ &+ \frac{8}{\pi}B_2(p, u, \nu)\sum_{k=-\infty}^{\infty}\frac{\delta(\omega-(2k+1)\Omega)}{(2k+1)^2}. \end{aligned} \tag{24}$$

Hence, we obtain an expression for the signal-to-noise ratio, defined as the ratio of the power of the first harmonic of the process $x(t)$ to the noise background:

$$\begin{aligned} \mathrm{SNR} &= \frac{8}{\pi} \\ &\times\frac{B_2(p, u, \nu)}{C(p, u, \nu) + \frac{2}{\gamma}(B_1(p, u, \nu) + B_3(p, u, \nu))}. \end{aligned} \tag{25}$$

2.2. Fast External Noise

Now, consider a case when the external noise is fast compared with the transition time w_0^{-1}; i.e., instead of (2), we apply the relation

$$T \gg w_0^{-1} \gg \gamma^{-1} \gg a^{-1} \gg \tau \longrightarrow 0 \tag{26}$$

rather than (9). In this case, we can average $\alpha(t)$ and $\beta(t)$ immediately over $S(r)$, replacing them by the averaged values

$$\begin{aligned} \alpha_1(t) &= \langle \alpha(t) \rangle_S = w_0 e^{p + vR(t)} \cosh(u), \\ \beta_1(t) &= \langle \beta(t) \rangle_S = w_0 e^{-p - vR(t)} \cosh(u) \end{aligned} \tag{27}$$

in Eq. (7).

Then, taking into account that the signal is still adiabatic, instead of (19), (24), and (25), we obtain

$$\begin{aligned} K(t) &= C\delta(t) + B_2\phi(t), \\ C &= \frac{x_0^2}{2w_0\cosh(u)}(\cosh^{-3}(p+v) + \cosh^{-3}(p-v)), \\ B_2 &= \frac{x_0^2}{4}(\tanh(p+v) - \tanh(p-v))^2, \\ S(\omega) &= C + \frac{8}{\pi}B_2\sum_{k=-\infty}^{\infty}\frac{\delta(\omega-(2k+1)\Omega)}{(2k+1)^2}, \\ \mathrm{SNR} &= \frac{8}{\pi}\frac{B_2}{C} \\ &= \frac{4w_0}{\pi}\cosh(u)\frac{(\tanh(p+v) - \tanh(p-v))^2}{\cosh^{-3}(p+v) + \cosh^{-3}(p-v)}. \end{aligned} \tag{28}$$

2.3. The Case of a Small Signal $v \ll 1$

Let us analyze the results obtained in greater detail. Consider the domain

$$v \ll 1; \quad p, u \gg 1. \tag{29}$$

Then, using formulas (19), (24), and (25) and the fact that $B_3 \ll B_1$ for $v \ll 1$, we obtain the following expressions for the total noise $S(0)$, signal power B_2, and the SNR in the adiabatic case:

$$\begin{aligned} S(0) &= C(p, u, 0) + \frac{2}{\gamma}B_1(p, u, 0) \\ &= \frac{x_0^2}{2w_0}(\cosh^{-3}(p+u) + \cosh^{-3}(p-u)) \\ &+ \frac{x_0^2}{2\gamma}(\tanh(p+u) - \tanh(p-u))^2, \\ B_2&(p, u, v \ll 1) \\ &= \frac{x_0^2v^2}{4}(\cosh^{-3}(p+u) + \cosh^{-3}(p-u))^2, \\ \mathrm{SNR} &= \frac{8}{\pi}\frac{B_2}{C + 2B_1/\gamma}. \end{aligned} \tag{30}$$

Figure 1 represents the functions C, B_1, B_2, and SNR calculated by (30). This figure clearly manifests resonance peaks at $p = \pm u$. Thus, when $p = \pm u$, all the quantities of interest are strongly nonmonotonic. On the other hand, it is clear from (6) that points $p = \pm u$ for $v \ll 1$ correspond to the situation when the equality

$$\alpha(t) = \beta(t) \tag{31}$$

holds for $S = \pm 1$. One can readily show that, under condition (31), all the quantities exhibit singularities for arbitrary (not only exponential, as in (6)) dependences of α and β on the external signal. This fact requires special consideration.

Let us write out the asymptotics of expression (30) in the domain $p > 0$, p, u, $p - u \gg 1$. From (30), we obtain

$$\begin{aligned} B_1 &= x_0^2\exp(-4(p-u)), \\ B_2 &= 4x_0^2v^2\exp(-4(p-u)), \\ C &= \frac{4x_0^2}{w_0}\exp(-3(p-u)), \\ \mathrm{SNR} &= \frac{8v^2w_0}{\pi}\exp(-(p-u)). \end{aligned} \tag{32}$$

For fast noise for $v \ll 1$, formulas (28) yield

$$\begin{aligned} S(0) &= C = \frac{x_0^2}{w_0\cosh(u)\cosh^3(p)}, \\ B_2 &= \frac{x_0^2v^2}{\cosh^4(p)}, \quad \mathrm{SNR} = \frac{8w_0v^2}{\pi}\frac{\cosh(u)}{\cosh(p)}. \end{aligned} \tag{33}$$

Expressions (33) show suppression of noise by noise when an addition of fast external noise u leads to

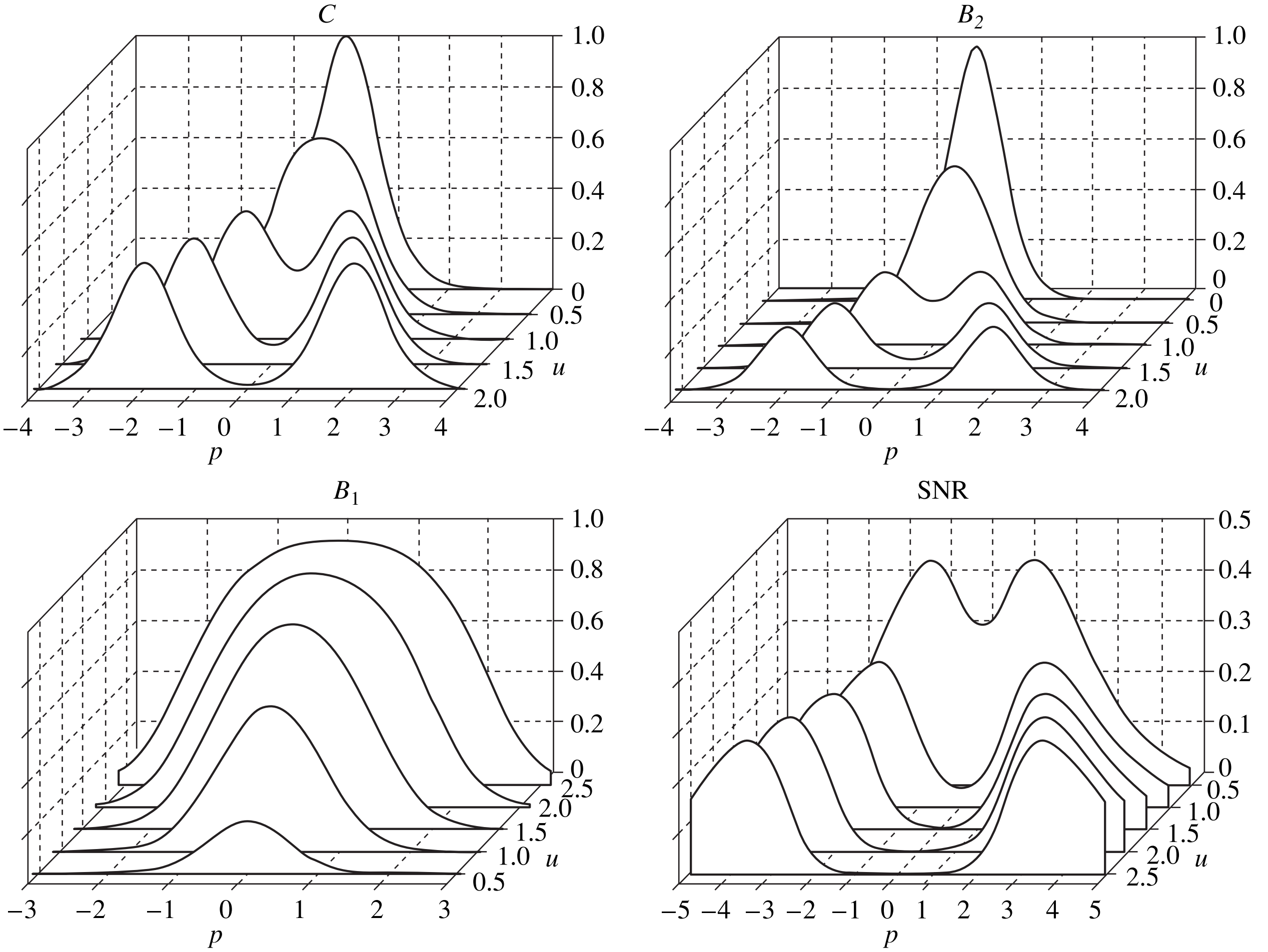

Fig. 1. Computed functions (30) for adiabatically slow noise; $x_0 = 1$, $v^2 = 0.2$, and $\gamma = 0.1 \ll w_0 = 1$.

a decrease in the total noise $S(0)$ [34, 35].

3. ANALOG SIMULATION

To verify the results obtained, we carried out an analog simulation of our stochastic system using an electron model of a two-level nonlinear system—an overdamped oscillator with a piecewise linear current–voltage characteristic $I(x)$ (see Fig. 2), which is similar to the oscillator used in [36, 37]. The equation of our circuit is as follows:

$$\frac{V(t)-x(t)}{R} = C\frac{dx}{dt} + I_{\mathrm{N}}(x),$$
$$RC\frac{dx}{dt} = -RI(x) + V(t), \quad I(x) = I_{\mathrm{N}}(x) + \frac{x}{R}. \tag{34}$$

Here, $V(t)$ is the applied voltage; $I_{\mathrm{N}}(x)$ and $I(x)$ are the current–voltage characteristics of the nonlinear element and the entire circuit, respectively; and $x(t)$ is the voltage to be analyzed. Let us introduce dimensionless time by changing $t \longrightarrow t/RC$ and obtain the following equation, which is completely similar to Eq. (1):

$$\frac{dx}{dt} = f(x) + V(t),$$
$$V(t) = F + \Delta S(t) + AR(t) + \sigma\xi(T). \tag{35}$$

$$f(x) = \begin{cases} -a(x+x_0), & x < -x_1 \\ \lambda x, & (|x| < x_1) = ax_0/(a+\lambda) \\ -a(x-x_0), & x > x_1. \end{cases} \tag{36}$$

From the current–voltage characteristic of our circuit, we determine $x_1 \approx 2.4$ V, $x_0 \approx 2x_1 \approx 4.7$ V, and the potential barrier height in volts, $\delta U_0 \approx 11\ V^2$. The voltage $V(t)$ applied to the circuit represents the sum of Gaussian noise $\xi(t)$ with a correlation function of the form $(\sigma^2/2\tau)\exp(-|t|/\tau)$ (see (1) and (2)), where the characteristic cutoff frequency $1/2\pi\tau \approx 30$ kHz is much greater than all the characteristic frequencies in the system

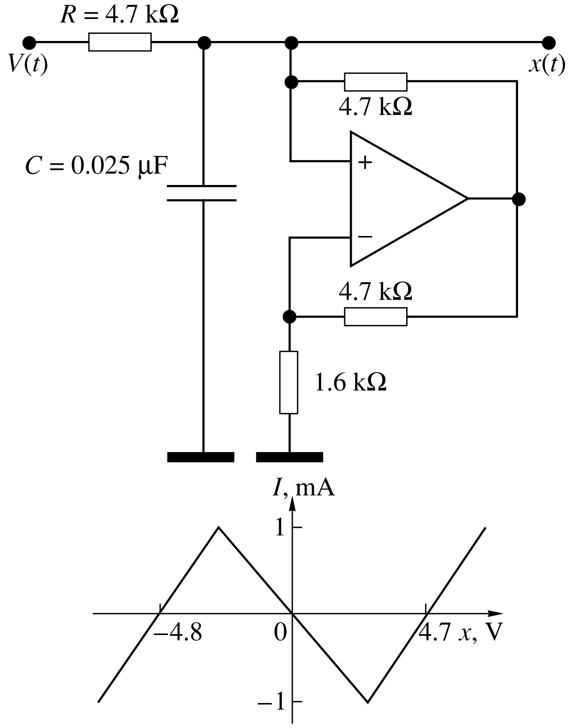

Fig. 2. Schematic diagram and the current–voltage characteristic of an analog model of an overdamped oscillator; $V(t)$ is applied voltage (a sum of a square-wave signal and white and dichotomic noise) and $x(t)$ is the signal to be analyzed. The operational amplifier is of the KP544UD1 type.

(physical white noise), dichotomic noise with an amplitude of $0 \le \Delta \le 10$ V and a characteristic frequency of $0 \le \gamma/2\pi \le 2000$ Hz, and a variable square-wave signal with amplitude A and a frequency of $f_s = \Omega/2\pi$. The characteristic frequency (inverse relaxation time) of the nonlinear system is $1/2\pi RC = 1400$ Hz. The switching rate w_0 between two states in the system can be changed by varying the intensity of the Gaussian noise.

In the experiment, we measured the spectral density of the output-signal power $x(t)$ at zero frequency, $S(0)$, and at the frequency of the square-wave signal, $S(\Omega)$, as well as the SNR (25):

$$S(\Omega) = \frac{4}{\pi^2}\frac{B_2}{\Delta f}, \quad \text{SNR} = \frac{2\pi\Delta f S(\Omega)}{S(0)}, \qquad (37)$$

where $\Delta f \approx 0.16$ Hz is the frequency interval in which the power spectral density $S(\omega)$ is measured for two limiting correlation times of the control dichotomic noise—fast noise (Figs. 3–5) and adiabatic slow noise (Figs. 6–8)—as a function of constant bias F. The curves are plotted by formula (33) for the fast noise and by (30) for the slow noise. Note that, since expression (30) correctly describes the SNR only for small values of A, we used exact formula (19) for B_2 in Fig. 8. One can see that the experimental and theoretical data are in good agreement.

The behavior of the quantities considered in the transient region where the hopping rate is on the order of the correlation time of the control noise, i.e., $w_0 \sim \gamma$, attracts interest. Since this domain cannot be studied theoretically, we carried out a simulation, which has shown that the results obtained in the adiabatic limit $\gamma \ll w_0$ can be used even for $\gamma \sim w_0$. Figure 9 shows the SNR for this case.

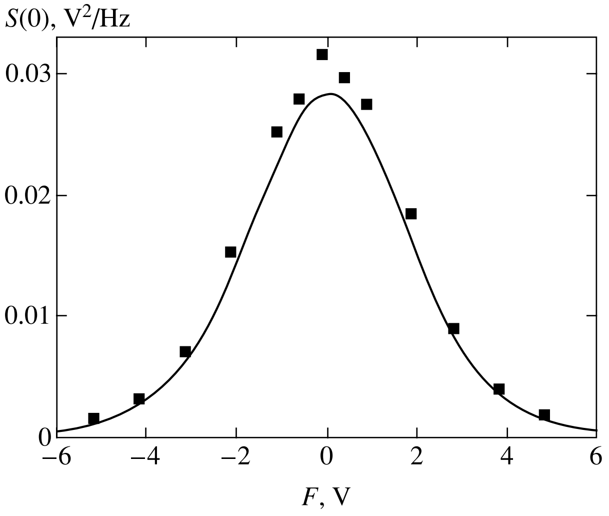

Fig. 3. Spectral density of the background noise $S(0)$ versus the bias voltage F for slow dichotomic noise $w_0/2\pi = 40$ Hz $\ll \gamma/2\pi = 560$ Hz. The solid curve represents the diagram of $C(p, u, \nu)$ from (33) plotted using the experimental parameters $x_0 = 4.7$ V, $2x_0/\sigma^2 = 0.34$ V^{-1}, $\Delta = 2.5$ V, and $\nu \approx 0.7$ V.

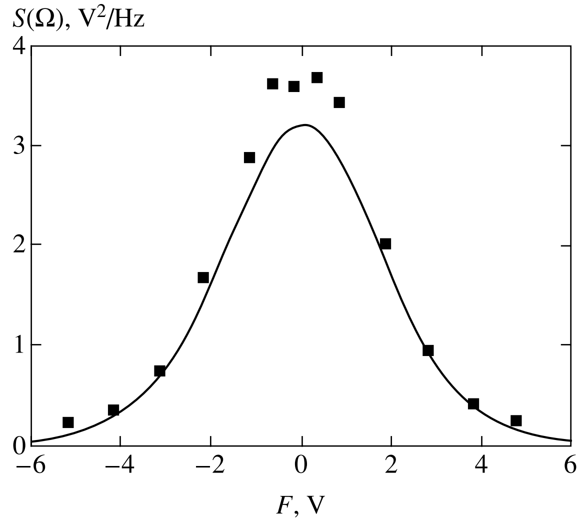

Fig. 4. Spectral density of the first harmonic of the output signal $S(\Omega)$, where $\Omega = 2\pi f_s$, $f_s = 0.5$ Hz, for fast dichotomic noise (see Fig. 3) and the function $S(\Omega) = (4/\pi^2)B_2(p, u, \nu)/\Delta f$ (see (33) and (37)).

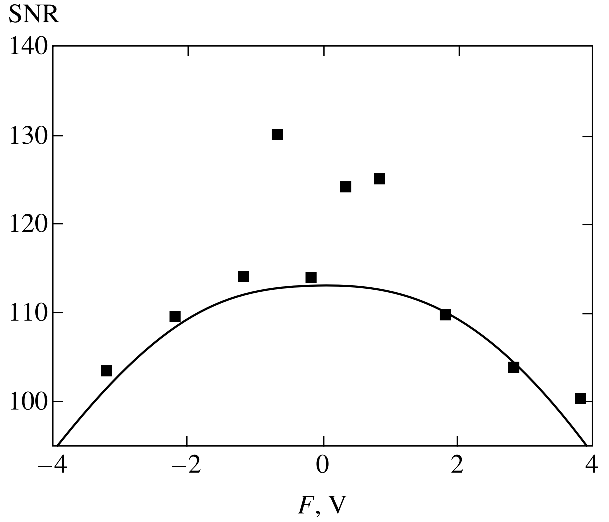

Fig. 5. SNR versus bias voltage F for fast dichotomic noise. The solid curve is plotted with the use of experimental values of the parameters; $\Delta f = 0.16$ Hz (see (30) and (37)).

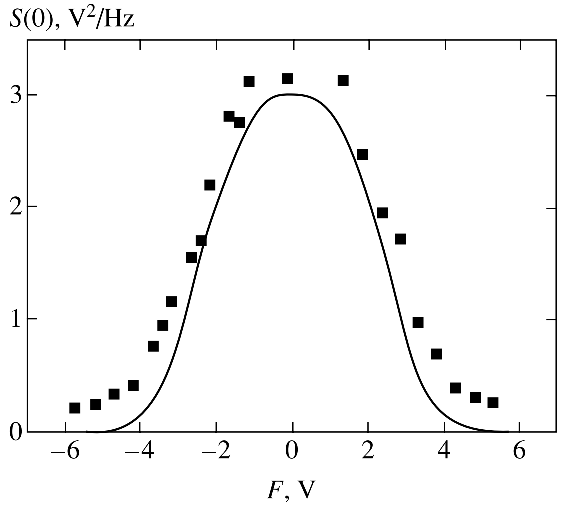

Fig. 6. Noise background $S(0)$ for slow dichotomic noise $\gamma/2\pi = 2.2$ Hz $\ll \gamma/2\pi = 40$ Hz. The solid curve represents $S(0)(p, u, \nu)$ from (30); $\Delta \approx 2.6$ V, $2x_0/\sigma^2 = 0.74$ V^{-1}, and $A = 0.6$ V.

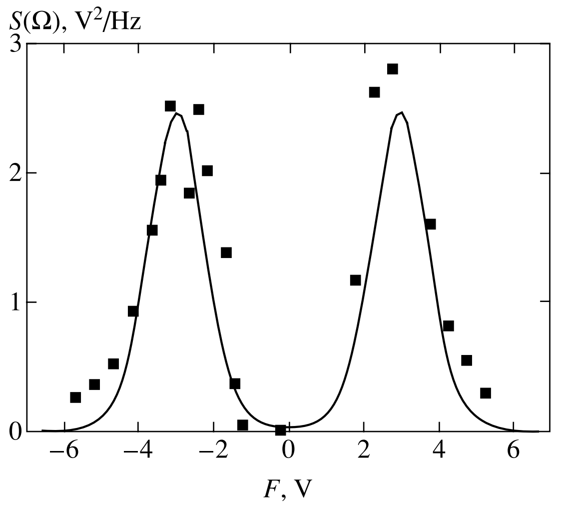

Fig. 7. Function $S(\Omega)$ for slow dichotomic noise. The solid curve represents the function $S(\Omega) = (4/\pi^2)B_2(p, u, \nu)/\Delta f$ (see (33) and (37)) with the same parameters as in Fig. 6.

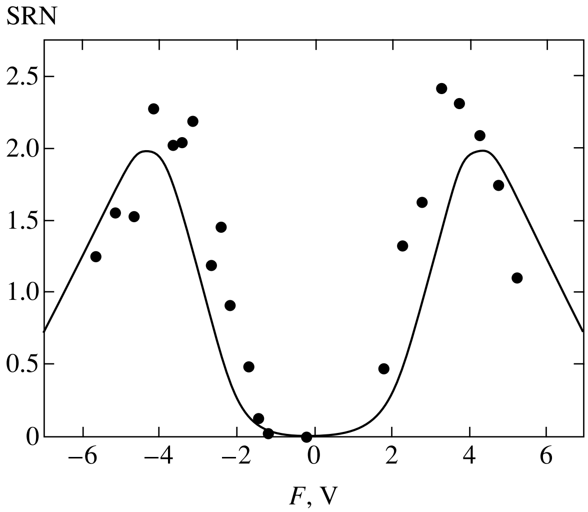

Fig. 8. SNR for slow dichotomic noise (see (19) and (30)).

It follows from the preceding section that the control dichotomic noise with amplitude Δ equal to the constant bias F (which guarantees the static asymmetry of the potential) restores the symmetry of the potential well for $\Delta S(t) = -F$; i.e., we can observe a dynamical restoration of symmetry. On the other hand, it is well known that the SNR is also maximal in the case of a symmetric potential; this fact leads to two maxima on the diagrams of SNR in Figs. 8 and 9. It is natural to assume that a similar phenomenon occurs in the case when there are two square-wave signals rather than dichotomic noise and a signal. Therefore, we carried out an investigation when dichotomic noise in (1) was replaced by a second square-wave signal with a period of $T_1 = 2\pi/\Omega_1$. Thus, we obtain

$$\frac{dx}{dt} = f(x) + F + \Delta R_1(t) + \sigma\xi(t) + AR(t), \qquad (38)$$

$$R_1(t + T_1) = R_1(t) = \begin{cases} 1, & 0 < t \le T_1/2 \\ -1, & T_1/2 < t \le T_1. \end{cases} \qquad (39)$$

The adiabaticity condition, similar to (9), is given by

$$T, T_1 \gg w_0^{-1} \gg a^{-1} \gg \tau \longrightarrow 0, \qquad (40)$$

while an analogue of the fast-noise condition (26) is

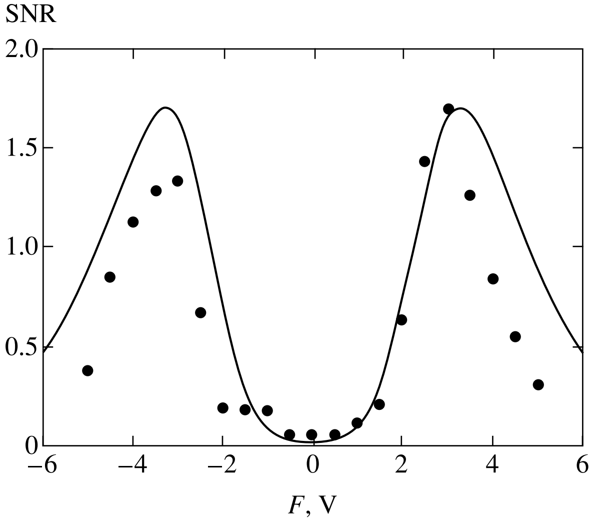

Fig. 9. SNR for an intermediate case when the correlation time of dichotomic noise is on the order of the Kramers switching rate $w_0/2\pi \sim \gamma/2\pi \approx 10$ Hz.

given by the fast signal

$$T \gg w_0^{-1} \gg T_1 \gg a^{-1} \gg \tau \longrightarrow 0. \qquad (41)$$

In the case of (40), instead of (19), we obtain

$$\begin{aligned} K(t) &= B_1(p, u, v)\phi_1(t) + B_2(p, u, v)\phi_0(t) \\ &\quad + B_3(p, u, v)\phi_0(t)\phi_1(t), \\ \phi_1(t) &= \langle R_1(t)R_1(0)\rangle_R = \frac{4}{\pi^2}\sum_{k=-\infty}^{\infty}\frac{e^{-i(2k+1)\Omega_1}}{(2k+1)^2}. \end{aligned} \qquad (42)$$

Other quantities are the same as in (19). Then, instead of (24), we obtain

$$\begin{aligned} S(\omega) &= C(p, u, v) \\ &+ \frac{8}{\pi}B_1(p, u, v)\sum_{k=-\infty}^{\infty}\frac{\delta(\omega-(2k+1)\Omega_1)}{(2k+1)^2} \\ &+ \frac{8}{\pi}B_2(p, u, v)\sum_{k=-\infty}^{\infty}\frac{\delta(\omega-(2k+1)\Omega)}{(2k+1)^2} \\ &+ \frac{32}{\pi^3}B_3(p, u, v)\sum_{k, n=-\infty}^{\infty}\frac{1}{(2k+1)^2(2n+1)^2} \\ &\times\delta(\omega-(2k+1)\Omega-(2n+1)\Omega_1). \end{aligned} \qquad (43)$$

Comparing (24) and (43), we can see that, although the coefficients C and B_i in these formulas are identical, the spectra are completely different. In particular, selecting different harmonics in (43), one can measure all the quantities C and B_i separately, whereas, in (24), we just have the combination $C + 2(B_1 + B_3)/\gamma$.

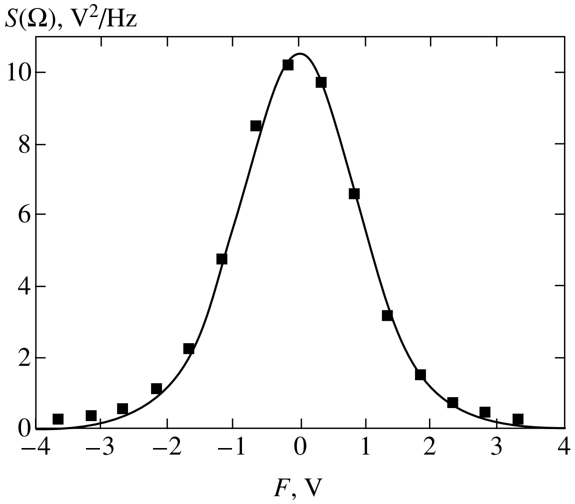

Fig. 10. Spectral density $S(\Omega)$ of the output signal $x(t)$ at frequency $f_s = \Omega/2\pi = 3.05$ Hz (the first low-frequency square-wave signal) for the case of two square-wave signals when the frequency of the second square-wave signal is $f_1 = \Omega_1/2\pi = 180$ Hz $\gg w_0/2\pi = 18$ Hz.

Now, let us consider the case when the second signal $\Delta R_1(t)$ is fast; i.e., condition (41) is satisfied. It is clear that, in this case, just as in the case of fast noise, one can carry out averaging over the fast signal. Then, all formulas (27) and (28) remain valid. Thus, the cases of fast noise and a fast signal are completely equivalent in the zeroth order with respect to the rapidity parameter.

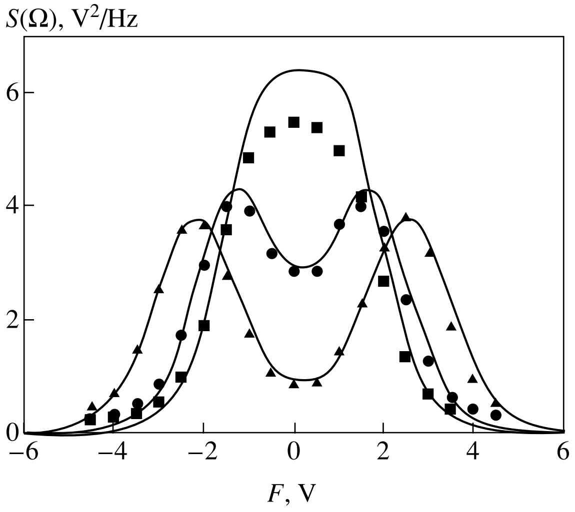

Fig. 11. Spectral density $S(\Omega)$ of the output signal $x(t)$ at frequency $f_s = \Omega/2\pi = 3.05$ Hz (the first low-frequency square-wave signal with amplitude 0.6 V) for the case of two square-wave signals when the frequency of the second square-wave signal is $f_1 = \Omega_1/2\pi = 18$ Hz $< w_0/2\pi = 50$ Hz and the amplitude is 1.06, 1.45, and 2.27 V (from top to bottom).

The experimental simulation carried out for two square-wave signals for the cases of fast ($\Omega_1/2\pi$ = 180 Hz $\gg w_0/2\pi \approx$ 18 Hz) and slow ($\Omega_1/2\pi$ = 18 Hz < $w_0/2\pi \approx$ 50 Hz) second signal has also confirmed the theoretical results. Figures 10 and 11 represent the function $S(\Omega)$ for the fast and the slow second signal, respectively. The results obtained are completely similar to those obtained for dichotomic noise.

4. CONCLUSIONS

Thus, we have theoretically and experimentally investigated a two-well nonlinear system with static asymmetry of the potential and with double stochasticity, i.e., under a certain constant bias and a simultaneous action of two types of noise, white and dichotomic. We have revealed that stochastic resonance is observed when control dichotomic noise is slow compared with the Kramers switching time in the nonlinear system and the noise amplitude is comparable to the value of the constant bias.

This result is associated with the fact that an external control action, for example, dichotomic noise or a square-wave signal with the amplitude equal to the constant bias (static asymmetry of the potential), restores the symmetry of the potential well within the time intervals when the sign of the external signal is opposite to that of the bias (dynamical restoration of symmetry).

ACKNOWLEDGMENTS

This work was supported by the Russian Foundation for Basic Research (project no. 02-02-16979) and by the State Programs *Quantum Macrophysics* and *Investigation of Cooperative and Quantum Phenomena in Condensed Media.*

REFERENCES

1. P. S. Landa and A. A. Zaikin, Zh. Éksp. Teor. Fiz. **111**, 358 (1997) [JETP **84**, 197 (1997)].
2. W. Horsthemke and R. Lefever, *Noise Induced Transitions. Theory and Applications in Physics, Chemistry, and Biology* (Springer, New York, 1984; Mir, Moscow, 1987).
3. B. McNamara, K. Wiesenfeld, and R. Roy, Phys. Rev. Lett. **60**, 2626 (1988).
4. M. I. Dykman, A. L. Velikovich, G. P. Golubev, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **53**, 182 (1991) [JETP Lett. **53**, 193 (1991)].
5. L. Gammaitoni, M. Martinelli, L. Pardi, and S. Santucci, Phys. Rev. Lett. **67**, 1799 (1991).
6. A. Simon and A. Libchaber, Phys. Rev. Lett. **68**, 3375 (1992).
7. R. N. Mantegna and B. Spagnolo, Phys. Rev. E **49**, R1792 (1994).
8. A. D. Hibbs, A. L. Singsaas, E. W. Jacobs, *et al.*, J. Appl. Phys. **77**, 2582 (1995).
9. A. Perez-Madrid and J. M. Rubi, Phys. Rev. E **51**, 4159 (1995).
10. Z. Neda, Phys. Lett. A **210**, 125 (1996).
11. A. E. Dubinov, K. E. Mikheev, Yu. B. Nizhegorodtsev, and V. D. Selemir, Izv. Ross. Akad. Nauk, Ser. Fiz. **60** (10), 76 (1996).
12. S. Fauve and F. Heslot, Phys. Lett. A **97A**, 5 (1983).
13. G. Debnath, T. Zhou, and F. Moss, Phys. Rev. A **39**, 4323 (1989).
14. L. Gammaitoni, F. Marchesoni, E. Menichella-Saeta, and S. Santucci, Phys. Rev. Lett. **62**, 349 (1989).
15. G. Vemuri and R. Roy, Phys. Rev. A **39**, 4668 (1989).
16. M. I. Dykman, R. Mannella, P. V. E. McClintock, and N. G. Stocks, Phys. Rev. Lett. **65**, 2606 (1990).
17. L. Gammaitoni, F. Marchesoni, E. Menichella-Saeta, and S. Santucci, Phys. Rev. Lett. **65**, 2607 (1990).
18. C. Presilla, F. Marchesoni, and L. Gammaitoni, Phys. Rev. A **40**, 2105 (1989).
19. L. Gammaitoni, E. Menichella-Saeta, S. Santucci, *et al.*, Phys. Rev. A **40**, 2114 (1989).
20. O. V. Gerashchenko, Pis'ma Zh. Tekh. Fiz. **29**, 82 (2003) [Tech. Phys. Lett. **29**, 256 (2003)].
21. D. S. Leonard and L. E. Reichl, Phys. Rev. E **49**, 1734 (1994).
22. M. I. Dykman, T. Horita, and J. Ross, J. Chem. Phys. **103**, 966 (1995).
23. W. Hohmann, J. Muller, and F. W. Scneider, J. Chem. Phys. **100**, 5388 (1996).
24. P. Babinec, Phys. Lett. A **225**, 179 (1997).
25. L. Gammaitoni, P. Hanggi, P. Jung, and F. Marchesoni, Rev. Mod. Phys. **70**, 223 (1998).
26. V. S. Anishchenko, A. V. Neiman, F. Moss, and L. Shimanski-Geier, Usp. Fiz. Nauk **169**, 7 (1999) [Phys. Usp. **42**, 7 (1999)].
27. A. Neiman, L. Schimansky-Geier, F. Moss, *et al.*, Phys. Rev. E **60**, 284 (1999).
28. R. Rosenfeld, A. Neiman, and L. Schimansky-Geier, Phys. Rev. E **62**, R3031 (2000).
29. P. Jung, Phys. Rep. **234**, 175 (1993).
30. M. I. Dykman, D. G. Luchinsky, R. Mannella, *et al.*, J. Stat. Phys. **70**, 463 (1993).
31. M. I. Dykman, D. G. Luchinsky, R. Mannella, *et al.*, J. Stat. Phys. **70**, 479 (1993).
32. M. I. Dykman, R. Mannella, P. V. E. McClintock, and N. G. Stocks, Phys. Rev. Lett. **68**, 2985 (1992).
33. M. I. Dykman, H. Haken, Gang Hu, *et al.*, Phys. Lett. A **180**, 332 (1993).
34. S. L. Ginzburg, M. A. Pustovoĭt, and O. V. Gerashchenko, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 672 (2001) [JETP Lett. **73**, 592 (2001)].
35. S. L. Ginzburg, M. A. Pustovoit, and O. V. Gerashchenko, Fluct. Noise Lett. **1**, L131 (2001).
36. O. V. Gerashchenko, Zh. Éksp. Teor. Fiz. **116**, 1477 (1999) [JETP **89**, 797 (1999)].
37. O. V. Gerashchenko, S. L. Ginzburg, and M. A. Pustovoit, Eur. Phys. J. B **19**, 101 (2001).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 836–845.
Translated from Zhurnal Éksperimental’noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 932–942.
Original Russian Text Copyright © 2003 by Kartsev.

A New Quantum Monte Carlo Algorithm in the Momentum Representation: The Sign Problem and the Hess–Fairbank Effect

P. F. Kartsev
Moscow Engineering Physics Institute (State University), Kashirskoe sh. 31, Moscow, 115409 Russia
e-mail: kash@picolab.mephi.ru

Received April 10, 2003

Abstract—An exact numerical algorithm based on the diagrammatic quantum Monte Carlo method in the momentum representation is proposed; in many cases, this algorithm is free of the sign problem and extends the class of models that can be analyzed by cluster methods. The weakening of the sign problem is demonstrated via the determination of the ground state of electrons on a chain in the Hubbard model. The algorithm is applied to the investigation of the behavior of a one-dimensional boson system with attraction in a rotating ring in the region of the Hess–Fairbank effect predicted by Ueda and Leggett. The existence of this effect for a comparatively small number of particles $N \sim 10$ is confirmed. An analytic boundary of this effect is determined in the limit as $N \longrightarrow \infty$. *© 2003 MAIK “Nauka/Interperiodica”.*

1. INTRODUCTION

Quantum Monte Carlo (QMC) methods, such as diagrammatic [1, 2] and variational, have been successfully applied to the study of many statistical models. For instance, the parameters of qubits based on Fe and Mn magnetic nanoclusters were determined by the diagrammatic QMC method [3] in situations when the decoherence time is large enough so that quantum computing devices can be implemented on the basis of these nanoclusters. Among important theoretical results, one should mention the investigation of a quasi-condensation of an interacting two-dimensional Bose gas [4] and the determination of a shift in the critical temperature of an interacting Bose gas [5].

One of the difficulties that restrict the application of the QMC methods—the so-called sign problem—arises due to the representation of a partition function by an alternating series; as a result, the error in the calculated mean values increases. The average sign <sign> of the statistical weight that determines the computation efficiency falls off exponentially as temperature decreases [6]; therefore, the study of the low-temperature properties of many prospective systems proves to be impossible.

The sign of a term of the series depends on the set of internal parameters of the method that correspond to the last term (the so-called Monte Carlo configuration), the Hamiltonian of a specific system, and the algorithm chosen. When simulating fermion systems, anticommutation of the wave function with respect to the permutation of particles also contributes to the sign.

Some models are free of the sign problem; however, this problem can arise under a comparatively small, unessential for the physical picture, perturbation of the Hamiltonian. Conversely, some systems can be reduced to a form that admits a simulation by the QMC method without the sign problem by a small variation or even by the identity transformation. For instance, a variation in the phase on a certain sublattice of a simple square (cubic) lattice [7] allows one to change the sign of the hopping matrix element t between nearest neighbors; this allows one to avoid the sign problem for $t > 0$.

Sometimes, the sign problem can be avoided by an appropriate choice of the numerical method. For instance, in some cases when electron systems with spin by a determinant method are investigated, all terms are positive [8]. However, the characteristic time of numerical calculations by determinant algorithms is proportional to the cube of the size L of a system, whereas the characteristic time of trajectory methods is linear in L; therefore, the latter methods prove to be more efficient when applied to the study of large systems ($L > 10$) even in the presence of the sign problem. On the other hand, for small systems ($L < 10$), the exact diagonalization of the Hamiltonian matrix [9, 10] is preferable in many cases.

There exists a correction method that allows one to determine the ground-state energy with high accuracy [6] even if the sign problem prevents one from reaching a sufficiently low temperature. However, no recipe is known for a similar correction of an arbitrary calculated mean value. Thus, the sign problem represents a serious obstacle in QMC methods.

In this paper, we present a new trajectory algorithm in the momentum representation that is developed on the basis of the diagrammatic QMC (DQMC) method [11].

This algorithm weakens the sign problem when simulating weakly interacting systems and allows one to analyze new models that do not admit the application of algorithms acting in a real space. One of such models is considered in Section 3.

2. DIAGRAMMATIC QUANTUM MONTE CARLO METHOD IN THE MOMENTUM REPRESENTATION

2.1. Formulation of the DQMC Method

According to [11], the Hamiltonian of a system is divided into two parts, $\hat{H} = \hat{H}_0 + \hat{V}$, where the ground-state energy $\hat{H}_0$ has a diagonal form in the representation of occupation numbers $\{n\}$:

$$\hat{H}_0|\{n\}^{(j)}\rangle = E_0^{(j)}|\{n\}^{(j)}\rangle,$$

and the perturbation is given by $\hat{V} = \sum_p \hat{Q}_p$.

Then, the partition function in the interaction representation is expressed as

$$\begin{aligned} Z = \sum_{m=0}^{\infty} \sum_{\substack{\{n^{(1)}\} \\ \{n^{(m)}\}}} \int_0^{\beta}(-d\tau_{m-1}) \int_0^{\tau_{m-1}} (-d\tau_{m-2}) \ldots \int_0^{\tau_1}(-d\tau_0) \\ \times \exp(-\beta E_0^{(0)}) \prod_{j=1}^{m} \exp(-\tau_j E_0^{(j-1)}) \\ \times \langle\{n\}^{(j-1)}|\hat{V}|\{n\}^{(j)}\rangle \exp(\tau_j E_0^{(j)}), \end{aligned} \tag{1}$$

where

$$\{n\}^{(m)} \equiv \{n\}^{(0)}, \quad \tau_m \equiv \tau_0 + \beta, \quad \beta = \frac{1}{k_B T}.$$

Here and in what follows, $\hbar = 1$. Calculations are carried out in continuous imaginary time $0 \le \tau \le \beta$, which is divided, to facilitate computer processing, into sufficiently small segments $\Delta\tau \sim 10^{-8}\beta$. Each term in (1) can be represented by a set of particle trajectories in the $(d + 1)$-dimensional space $(\mathbf{x}, \tau)$ and is expressed as

$$W \sim \prod_{j=1}^{m} (-\Delta\tau \langle\{n\}^{(j-1)}|\hat{Q}_{p_j}|\{n\}^{(j)}\rangle \exp(\ldots)), \tag{2}$$

where τ_j correspond to the times when the world lines are changed under perturbation $\hat{Q}_{p_j}$—the so-called kinks [11]—while the set of values p_j, $\{n\}^{(j)}$, and τ_j defines a Monte Carlo configuration.

The perturbation operator $\hat{V}$ is divided into terms $\hat{Q}_p$ for convenience. For instance, for a hopping term in the Hubbard-type model $\hat{Q}_{ij} = -t_{ij}\hat{a}_i^+\hat{a}_j$ ($\hat{a}_i^+$ and $\hat{a}_i^-$ are the creation and annihilation operators, respectively), a kink represents a jump of a trajectory from site j to site i at moment τ of imaginary time. On the other hand, the perturbation operator also admits diagonal terms that do not change the occupation numbers of the sites involved. The algorithm presented below employs both types of kinks.

As an elementary Monte Carlo step, we use addition and annihilation of kinks, as well as the variation of their position τ in imaginary time. Calculations are carried out by the Metropolis algorithm in which each configuration update has its inverse, whereas the probability R of adopting a transition between configurations A and B is chosen according to the expression

$$W_A R_{A \to B} f_{A \to B} = W_B R_{B \to A} f_{B \to A},$$

where W_A and W_B are the statistical weights of the old and new configurations, respectively, and $f_{A \to B}$ and $f_{B \to A}$ are the probabilities of choosing the transformations $A \longrightarrow B$ and $B \longrightarrow A$ in the algorithm, respectively. The simplest variant for choosing R is the so-called heath-bath algorithm [8]:

$$R_{A \to B} = \frac{W_B f_{B \to A}}{W_A f_{A \to B} + W_B f_{B \to A}}. \tag{3}$$

In addition, a natural requirement is imposed on the system of processes: there should exist a nonzero probability of transitions between any two admissible (i.e., with weight $W \neq 0$) Monte Carlo configurations in a finite number of steps.

2.2. DQMC Method in the Momentum Representation

The algorithm presented is developed for fermion and boson systems that have a Hubbard-type Hamiltonian in the momentum representation:

$$\hat{H} = \sum_p \epsilon_p \hat{a}_p^+ \hat{a}_p + \sum_{p, q, r, s} U_{pqrs} \hat{a}_p^+ \hat{a}_q^+ \hat{a}_r \hat{a}_s, \tag{4}$$

where ϵ_p is the energy of a particle possessing the momentum p. Further, we will consider a typical term of interaction

$$U_{pqrs} = U_{q-r}\delta_{p+q, r+s},$$

although momentum conservation is unessential for the algorithm described.

The interaction taken as perturbation $\hat{V}$ generates kinks that are shown in Fig. 1. A factor, corresponding to a kink, that appears in the weight of a configuration

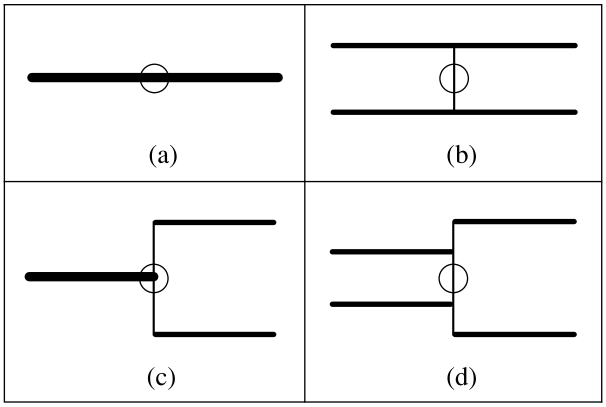

Fig. 1. Kinks generated by two-particle interaction; (a, b) diagonal and (c, d) nondiagonal kinks. Kinks (a) and (b) can be taken into account analytically in the case of contact interaction $U_{pqrs} = U_0\delta_{p+q,\,r+s}$; kinks (a) and (c) arise only during simulation of a Bose system. Here and below, the imaginary-time axis is directed horizontally, momenta are measured in the vertical direction, the occupation number is represented by the width of a line, and a kink is denoted by a circle.

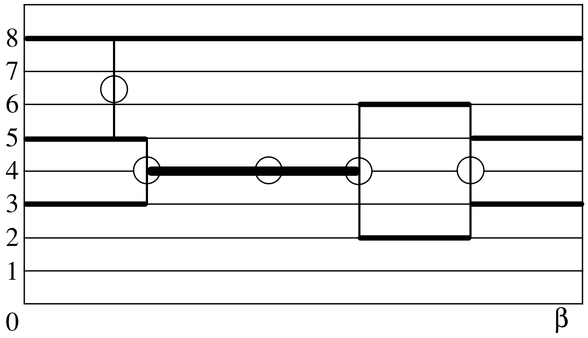

Fig. 2. An example of a trajectory configuration with all types of kinks (a–d) (see Fig. 1).

together with the exponential factor is given by a sum of appropriate terms $U_{pqrs}\sqrt{n_p n_q n_r n_s}$ of the series for all nonequivalent permutations of the momenta p, q, r, and s. For instance, when $U_{pqrs} = U_0\delta_{p+q,\,r+s}$, the factors $-\Delta\tau U_0 n_q(n_q - 1)$ and $-4\Delta\tau U_0 n_{q_1} n_{q_2}$ correspond to kinks (a) and (b).

An example of a configuration is shown in Fig. 2.

2.2.1. The sign problem in the new algorithm. The sign of a statistical weight is determined by the number of kinks and the relevant matrix elements of the perturbation $\hat{V}$. The average number of kinks in a configuration is estimated as $\beta|V|$ (see [11]); therefore, as applied to the simulation of weakly interacting systems, this algorithm proves preferable to conventional trajectory algorithms because of the weakening of the sign problem.

It is also important that, in the case of a point interaction of particles $U_{pqrs} = U_0\delta_{p+q,\,r+s}$, the diagonal part of the interaction gives only an energy shift; this allows us to simplify the algorithm by eliminating diagonal kinks (Figs. 1a and 1b) from it. Then, even in the case of repulsive interaction or Fermi statistics, a statistical weight may become negative only for configurations with at least three kinks; thus, at high temperatures, the sign problem (slowing down the calculation) is substantially weakened.

On the other hand, in the case of attractive interaction and the absence of antisymmetry, the algorithm is independent of sign because the contribution of any kink to the statistical weight is positive. This circumstance allows such systems to be studied with high precision even at low temperatures. (An example of such a problem is considered in the following section.)

2.2.2. Relation between the sign and the winding number. For a standard trajectory method, the fermion component of the sign is determined by the number of intersections of trajectories. Simple topological considerations yield [12]

$$\mathrm{sgn}_F W = (-1)^{\sum_i (Z_i - 1)}, \qquad (5)$$

where Z_i is the so-called winding number of the ith trajectory around the imaginary-time axis. This relation allows one to efficiently determine the sign of the statistical weight after the change of configuration.

However, in the algorithm considered, the trajectories of particles are linked in pairs via kinks, whereby the winding number of the trajectories is frequently undetermined. Therefore, to preserve the applicability of formula (5), kinks (b) and (d) (Fig. 1) are redefined as a simultaneous hopping of two particles, without intersection of trajectories, with an appropriate variation in the weight of a kink. For instance, the weight of kink (b) (Fig. 1) takes the form $-2\Delta\tau(U_0 - U_q)$.

2.2.3. Processes. Figure 3 shows the processes that are chosen to change a configuration of world lines, except for the “hole” version of process (b), and the motion of a kink in imaginary time. Here, it is also required to create and annihilate diagonal kinks in the case when the diagonal part of the perturbation has not been taken into consideration analytically as an energy shift. Note that, among all the processes listed, only the entanglement of a kink with a straight region of another worldline (Fig. 3c) may change the winding number of the trajectory and, hence (following (5)), the fermion component of the sign.

It is important that the processes considered locally preserve momentum and, hence, cannot change the total momentum K of the system. Simulation in the sector of the phase space with fixed K reduces the computation time of energy levels that are characterized by

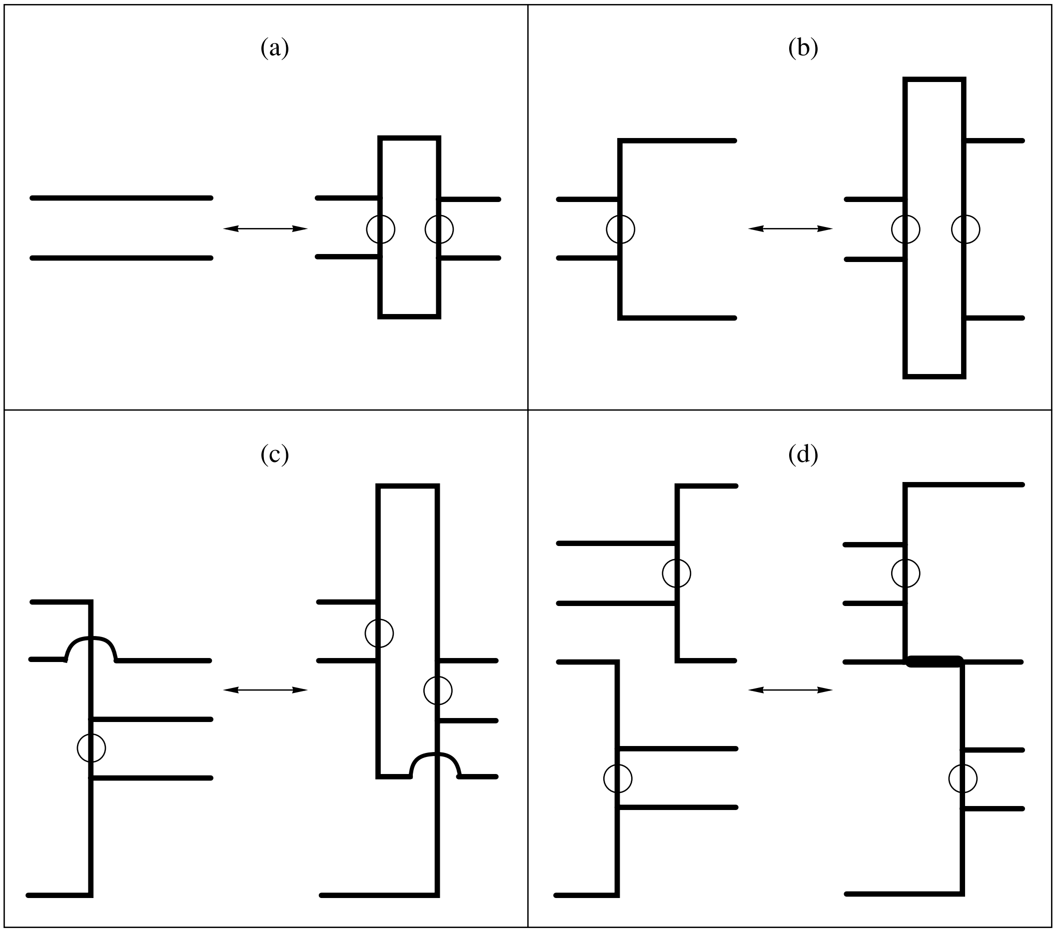

Fig. 3. Procedures for changing nondiagonal kinks: (a) creation/annihilation of a pair of kinks, (b) transformation of one kink into two, (c) entanglement of a kink and a straight region of a worldline, and (d) shift of a kink over an adjacent one.

certain momentum. However, to calculate other quantities, the total momentum of the system should take all possible values during the simulation. The easiest way of doing this is to introduce skewed kinks (Fig. 4) that correspond to a fictitious term,

$$\eta \sum_{\substack{pqrs \\ p+q \neq r+s}} \hat{a}_p^+ \hat{a}_q^+ \hat{a}_r \hat{a}_s,$$

in the Hamiltonian (see also the worm algorithm [13]). Just as in [13], the statistics is accumulated in the absence of skewed kinks, while the small quantity η is chosen from efficiency considerations.

2.3. The Fermi Hubbard Model on a Chain

To test the algorithm, we calculated the ground-state energy E_0 of a system of electrons on a chain of eight sites described by the Hamiltonian

$$\hat{H} = -t \sum_{\langle ij \rangle \sigma} (\hat{a}_{i\sigma}^+ \hat{a}_{j\sigma} + \text{H.c.}) + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \qquad (6)$$

for the particle number $N_\uparrow = N_\downarrow = 4$, $t = 1$, $U = -1$, and $K = 0$. Here, $\langle ij \rangle$ denotes a summation over neighboring sites. Since the system is small, we could test the algorithm with high accuracy. A value of $E_0 \approx -11.952(5)$ obtained for $\beta = 8.0$ coincides with the results obtained by a standard trajectory method in the conventional representation (the worm algorithm [13]), $E_0' \approx -11.(9)$ for $\beta = 2.0$, and by the method of exact diagonalization, $E_0 = -11.952326\ldots$.

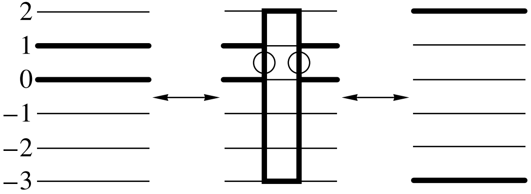

Fig. 4. Creation of a pair of skewed kinks and their subsequent collapse in a different order allow one to change the total momentum of the configuration from (a) $K = +1$ to (c) $K = -1$.

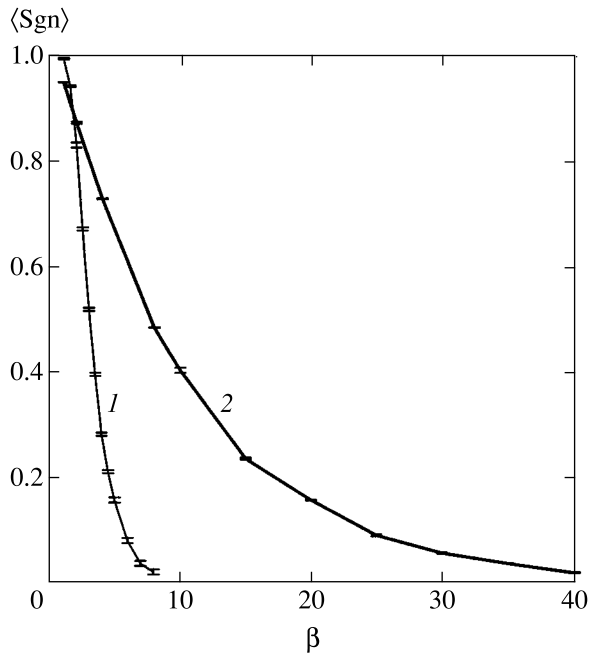

Fig. 5. The average sign as a function of temperature in calculations in (*1*) conventional and (*2*) momentum representations of a one-dimensional Hubbard model for electrons on a chain of eight sites with the number of particles $N_\uparrow = N_\downarrow = 4$; $U = -1$, and $t = 1$.

Such high accuracy of the new algorithm is attributed to the restriction of the phase space due to fixing the total momentum K. The ground ($E_0 = -11.952326\ldots$) and the subsequent ($E_1 = -11.901727\ldots$) energy levels are so close that they cannot be distinguished at practically achievable temperatures by standard Monte Carlo methods, whose accuracy is restricted to two or three decimal places due to the sign problem. However, the second of these levels corresponds to a nonzero momentum and therefore does not take part in the calculations performed by the algorithm proposed.

Calculations are possible when the average sign ⟨sign⟩ exceeds a few percent. The average sign as a function of inverse temperature β is shown in Fig. 5 for calculations by both QMC algorithms. The behavior of these functions confirms our assumption that a situation with the sign problem in the algorithm proposed is more optimistic. The highest possible value of β has increased from 7 to 40. Due to the attractive character of interaction, the sign has only a fermion component after passing to the momentum representation. (However, note that, in this system, repulsion ($U > 0$) can also be transformed to attraction by an electron–hole transformation for particles with one of the two values of spin σ.)

3. BOSE GAS WITH ATTRACTION IN A ROTATING RING

The algorithm described has been applied to studying the ground state of a quasi-one-dimensional Bose gas in a rotating ring of radius R with attractive interaction between particles [14, 15]. The Hamiltonian of the system has a form analogous to (4):

$$H = \hbar\omega_c \sum_k \left(k - \frac{\omega}{2\omega_c}\right)^2 n_k + \frac{g}{2}\sum_{k,l,q} \overset{+}{a}_k \overset{+}{a}_l a_{l-q} a_{k+q}, \qquad (7)$$

where ω is the angular velocity of rotation of a container, $\omega_c = \hbar/2mR^2$ is the critical angular velocity, $g = 2a\hbar^2/mRS < 0$ is the effective amplitude of interaction, a is the s-wave scattering length, m is the atomic mass, and $S = \pi r^2$ is the cross-section area of a torus.

An example of a Bose system with the negative scattering length is given by a ^{7}Li atomic gas, which was investigated in experiments on Bose condensation at ultralow temperatures [16]. For other alkaline metals, one can apply the Feshbach resonance [17], which allows one to vary the scattering length and, in particular, to change its sign. It is important that, while a condensate of attracting atoms is unstable in the bulk [18], in the one-dimensional case, a droplet represents the ground state and is stable [19]. The model considered becomes effectively one-dimensional at a sufficiently low temperature and for $r \ll R$. Variants of experimental realizations of such a system are demonstrated in [20, 21].

To simplify the analysis, we assume that $\hbar = 1$ and $\omega_c = 1$ and introduce a dimensionless interaction parameter

$$\gamma = \frac{|g|(N-1)}{\hbar\omega_c}.$$

3.1. Introduction to the Problem

Ueda and Leggett, who proposed this system as a model of an irrotational fluid, have shown in [14] that, in the macroscopic limit and for $\gamma < \gamma_c \sim 1$, there exists a cutoff angular velocity ω_* of a torus below which the gas is not dragged by the walls (the so-called Hess–Fairbank effect). In this case, all the particles occupy the same state with the angular momentum $k = 0$ and the graph of the momentum of the system versus the rotation speed has a horizontal plateau repeated periodically with period $2\omega_c$ (see Fig. 6). A similar result is obtained when the system is investigated in the limit as $\gamma \longrightarrow 0$ and $N \longrightarrow \infty$ by the numerical solution of the Gross–Pitaevskii equation.

On the other hand, the simulation carried out by Berman *et al.* [22] of the dynamics of a system in the basis of coherent states cast doubt on the stability of the condensate state in the region of the plateau even for small γ and, hence, on the applicability of standard methods of a weakly interacting Bose gas to this problem. Thus, until recently, there has been no unanimity of opinion about the true behavior of a quasi-one-dimensional system of bosons with attraction in a rotating vessel.

For the related problem of attracting bosons on an infinite line, there exist exact analytic solutions of the Gross–Pitaevskii [23] and Schrödinger [19] equations. In the limit as $N \longrightarrow \infty$, these equations yield identical results for the ground state of a system (see [24]); indirectly, this fact speaks in favor of standard methods and the Ueda–Leggett description for a system in the ring.

To conclude, we can calculate the ground state of a system by an exact numerical method. However, standard QMC methods acting on a lattice prove to be inefficient in this case because one has to discretize the space coordinate [25], which introduces a systematic error to the result [26]. When studying this problem, we applied the algorithm of the diagrammatic QMC method in the momentum representation, described in Section 2, which does not require the discretization of the coordinate. The second and the main advantage of the new algorithm in this problem is the complete absence of the sign problem.

Below in this section, we present the results of calculations by the QMC method for $N = 2, \ldots, 100$ for various values of the parameters ω and γ; these results confirm the existence of the Hess–Fairbank effect [15] in the system considered. An exact phase boundary of the effect is obtained analytically.

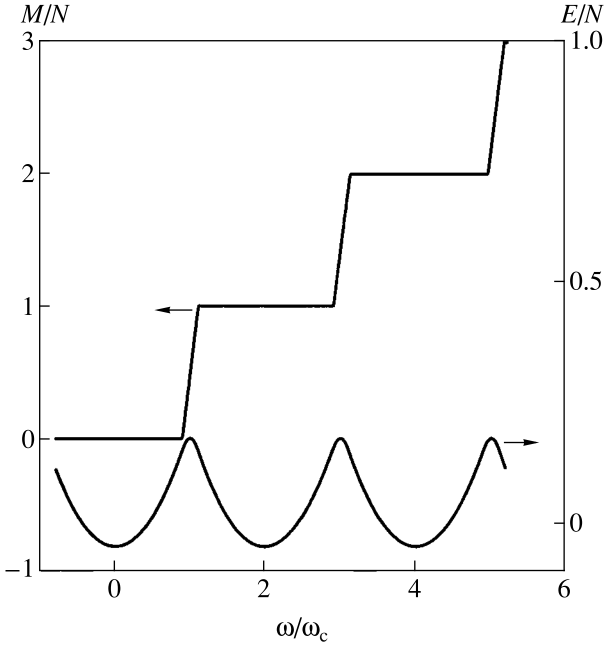

Fig. 6. Energy E and angular momentum M of the system in the ground state versus the rotation speed ω for $\gamma = 0.1$ and $N \longrightarrow \infty$, obtained according to macroscopic methods of an interacting Bose gas. Although the state corresponding to the horizontal plateau is unstable according to Berman *et al.* [22], the validity of the graphs is confirmed in the present paper by exact numerical calculations of a many-particle ground state.

3.2. A Macroscopic Case:
Exact Boundary of the Effect

The Hartree–Fock method in the basis of the states $\{\ldots, n_{-k}, \ldots, n_{-1}, n_0, n_1, \ldots, n_k, \ldots\}$ that was used in [14] is exact for $\gamma \longrightarrow 0$ because the Hamiltonian takes a diagonal form when only two rotational states with the minimal energies ϵ_k are taken into account. However, when $\gamma \sim 1$, corrections due to higher energy levels become essential.

Let us apply the Bogolyubov method [27] to take into account all momenta k. The plateau on the graph of $M(\omega)$ corresponds to the situation when all particles occupy the same rotational state with momentum $k = l$ (for definiteness, let $l = 0$). Then, up to corrections proportional to $1/N$, Hamiltonian (7) can be expressed as

$$\hat{H} = H_0 + \sum_k \epsilon_k \hat{n}_k - \frac{\gamma}{2}\sum_{k \neq 0}(2\hat{a}_k^+\hat{a}_k + \hat{a}_k^+\hat{a}_{-k}^+ + \hat{a}_k\hat{a}_{-k}),$$

where $\epsilon_k = k^2 - k\omega$. Changing $\hat{a}$, $\hat{a}^+ \longrightarrow \hat{b}$, $\hat{b}^+$ according to

$$\hat{a}_k = u_k\hat{b}_k + v_k\hat{b}_{-k}^+,$$

$$\hat{a}_k^+ = u_k^*\hat{b}_k^+ + v_k^*\hat{b}_{-k}$$

with the coefficients

$$u_k = \frac{1}{\sqrt{1-|L_k|^2}}, \quad v_k = \frac{L_k}{\sqrt{1-|L_k|^2}},$$

$$|u_k|^2 - |v_k|^2 = 1,$$

and requiring that the coefficients multiplying $\hat{b}_k\hat{b}_{-k}$ and $\hat{b}_k^+\hat{b}_{-k}^+$ should vanish, we obtain the following expressions:

$$L_k = \chi_k - \sqrt{\chi_k^2 - 1},$$

$$\chi_k = \frac{E_k}{\gamma} - 1,$$

$$E_k = \frac{\epsilon_k + \epsilon_{-k}}{2} = k^2.$$

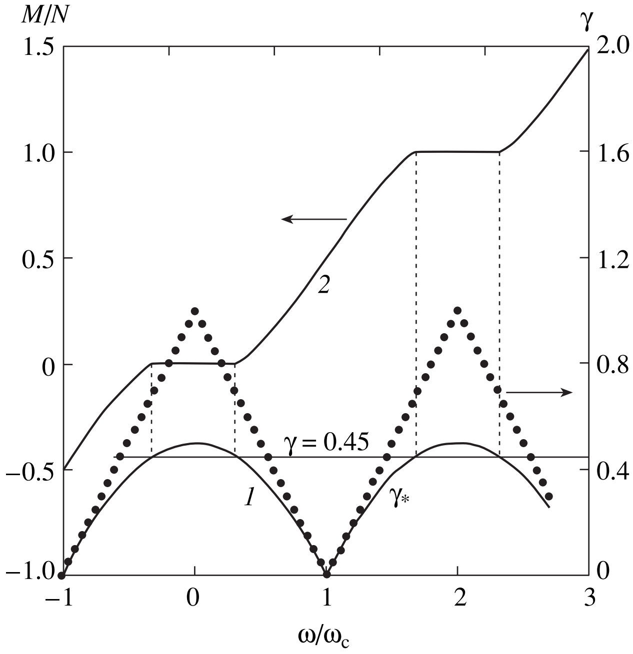

Fig. 7. (*1*) Exact phase diagram of the Hess–Fairbank effect obtained when analyzing the system in the limit as $N \longrightarrow \infty$ (formula (8); the effect exists for $\gamma < \gamma_*(\omega)$). The dots represent the approximate dependence [14]. When $\gamma \ll 1$, the graphs coincide; as γ increases, the width of the plateau rapidly decreases and the effect disappears at $\gamma = \gamma_c = 1/2$. (*2*) Angular momentum M of the system versus ω for $\gamma = 0.45$, obtained by numerically solving the Gross–Pitaevskii equation. The size of the plateau region on curve *2* corresponds to formula (8).

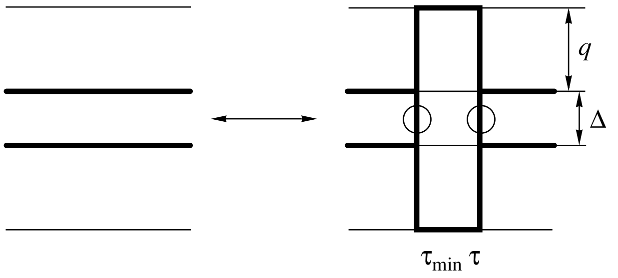

Fig. 8. Illustration to the notations of q and Δ in the text.

Note that such a solution is only possible when $\gamma < 1/2$. It is also clear that all particles have momentum $l = 0$ in the ground state only if the energy of quasiparticles, given by

$$\epsilon'(k) = \frac{1 - L_k}{1 + L_k} - \omega,$$

is greater than zero. Otherwise, the above arguments are not valid and the momentum $k = \pm 1$ proves to be occupied by a macroscopic number of particles; hence, the angular momentum M of the system becomes nonzero. The switching point that corresponds to the exact boundary of the Hess–Fairbank effect in this system (see Fig. 7) is given by

$$(\omega - 2l)^2 = 1 - 2\gamma. \tag{8}$$

This relation can also be obtained by analyzing the homogeneity of the solution to the Gross–Pitaevskii equation [28]; in terms of a macroscopic wave function, it is the violation of homogeneity that allows one to change the angular momentum. The solution to the Gross–Pitaevskii equation for this problem is given by the Jacobi elliptic function dn; although the parameters of this function are generally determined numerically [29], one can sufficiently easily derive an analytic solution for the moment when the nonhomogeneity of the ground state arises.

3.3. A Mesoscopic Case: Characteristic Features of the Algorithm

In all realizations of the diagrammatic Monte Carlo method, the time τ of a created or shifted kink is chosen, as in the worm algorithm [13], with the probability

$$\frac{\exp[-\delta E(\tau - \tau_{\min})]}{Z_1}.$$

This probability is proportional to the coefficient $\exp(-\delta E\tau)$ by which a statistical weight is multiplied after a change of configuration. The normalizing coefficient $Z_1(\delta E, \tau_{\max} - \tau_{\min})$ enters the expression $f_{A \to B}/f_{B \to A}$ (see (3)), which is used to determine the probability of adopting a process, and the update efficiency (the transition probability to a new configuration) decreases as δE increases. Due to the form of the spectrum in this problem ($\epsilon_k \sim k^2$), we had to take different values of the momentum k with different probabilities when determining the place where the procedure is applied.

For instance, in processes (a) and (b) (Fig. 3), the energy of trajectories in the interval considered is changed by an arbitrarily large value $\delta E = 2_q(q + \Delta)$ (see Fig. 8), and one has to choose q and Δ with the probability W corresponding to $Z_1(\delta E, \tau_{\max} - \tau_{\min})$. The approximation that was used to speed up computation is shown in Fig. 9. The details of each of the four processes are presented in [15].

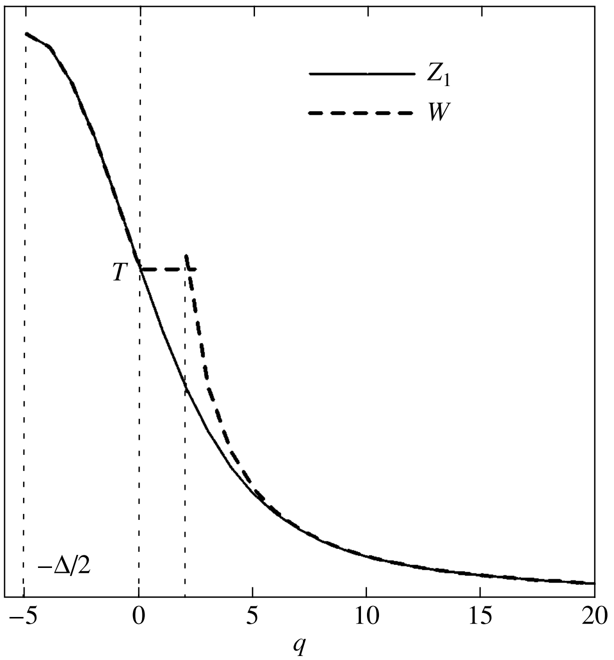

Fig. 9. Piecewise approximation W of the function $Z_1(\delta E(q), T)$ that was used to speed up the process of choosing momentum q in the course of processes (a) and (b) (Fig. 3). The value of momentum Δ is fixed when choosing the place of update.

It should also be noted that we used only nondiagonal kinks in the simulation (Figs. 1c and 1d) because the diagonal part of the interaction gives only an energy shift in this case:

$$\hat{H} = -\frac{|g|}{2}(2N^2 - N) + \sum_k \left[\left(k - \frac{\omega}{2}\right)^2 \hat{n}_k + \frac{|g|}{2}\hat{n}_k^2\right] - \frac{|g|}{2}\sum_{\substack{k,\,l,\,q\\ q \neq 0\\ q \neq l-k}} \hat{a}_k^+ \hat{a}_l^+ \hat{a}_{l-q}\hat{a}_{k+q}.$$

3.4. Mesoscopic Case: Results of Calculations

The energy of the ground state was calculated for a fixed angular momentum M of the system; then, we chose a branch with the minimal energy (see Figs. 10 and 11). Figure 12 shows the graphs of $M(\omega)/N$ for $N = 5$, 10, and 20 particles for $\gamma = 0.2$. For comparison, in the same figure we present a macroscopic function obtained by numerically solving the Gross–Pitaevskii

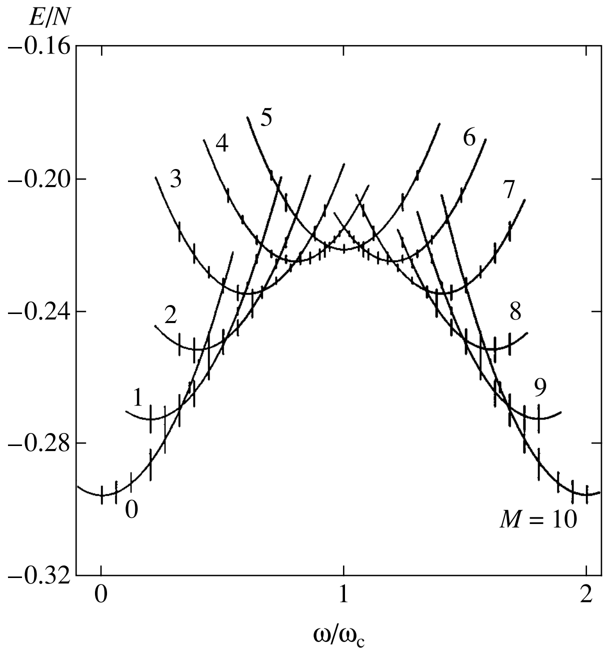

Fig. 10. Calculated ground-state energies of the system for $\gamma = \gamma_c = 1/2$ for 10 particles and various angular momenta $M = 0, \ldots, 10$. The value of angular momentum M is indicated by a number near a curve. In the absence of rotation, the ground state of the system corresponds to $M = 0$; as ω increases, the branches with $M = 1, 2$, etc., become ground states. Despite the error of the calculated points shown in the figure, the accuracy of the graphs is 10^{-2} because $E_0(M, \omega) = E_0^M + N(\omega/2 - M/N)^2$.

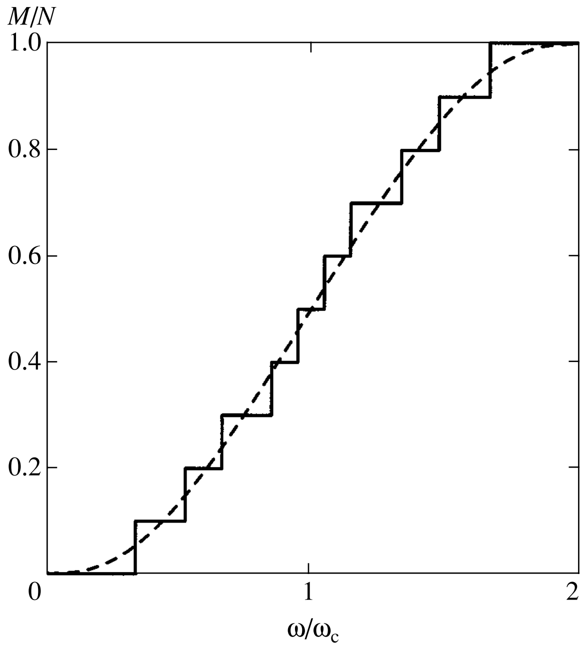

Fig. 11. Angular momentum of a system of ten particles in the ground state versus the speed of rotation for $\gamma = \gamma_c = 1/2$, calculated according to Fig. 10. The absolute error is 10^{-2}. The dashed curve represents a macroscopic dependence.

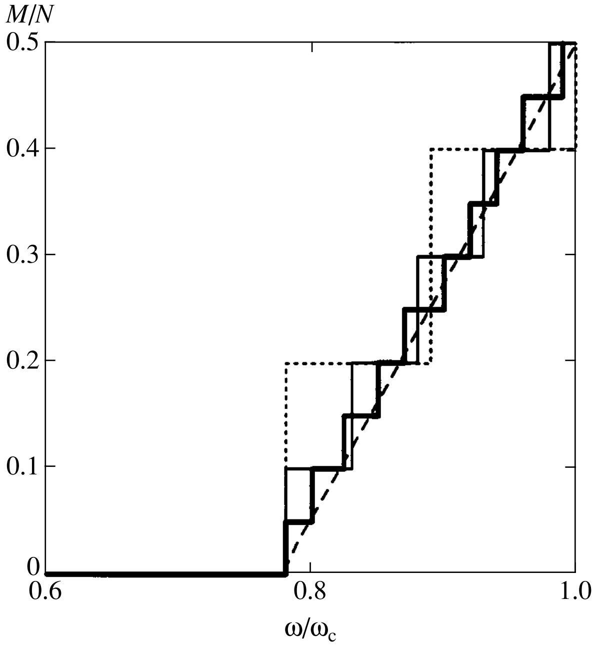

Fig. 12. The angular momentum of the system in the ground state versus the speed of rotation, calculated for 5, 10, and 20 particles for $\gamma = 0.2$. The absolute error is 10^{-2}. The dashed curve represents a macroscopic dependence. It is essential that the size of the plateau remains virtually unchanged as the number of particles is changed.

equation for the same value of γ. It is surprising that the width of the plateau is hardly changed as the number of particles decreases, and expression (8) proves to be applicable even for $N < 10$, although the width of intermediate steps becomes comparable to the size of the plateau for a small number of particles. The same situation occurs for other values of γ. (Note the difference between the definition $\gamma = |g|N/\hbar\omega_c$, which is given in [14], and $\gamma = |g|(N-1)/\hbar\omega_c$, which is used in the present paper; the second variant is more correct.)

Thus, the phase boundary of the Hess–Fairbank effect determined in the limit as $N \longrightarrow \infty$ remains unchanged when the number of particles is sufficiently small for computations by cluster methods.

As an illustration, we calculated the "density–density" correlation function

$$K(x) = \frac{1}{N(N-1)}\langle\Psi^{+}(x')\Psi^{+}(x'+x)\Psi(x'+x)\Psi(x')\rangle_{x'}$$

for homogeneous and nonhomogeneous (in the macroscopic limit) cases (Figs. 13 and 14, respectively). For a small number of particles, their positions strongly correlate in both cases and the functions are virtually identical; for large N, the graphs correspond to the solution of the Gross–Pitaevskii equation.

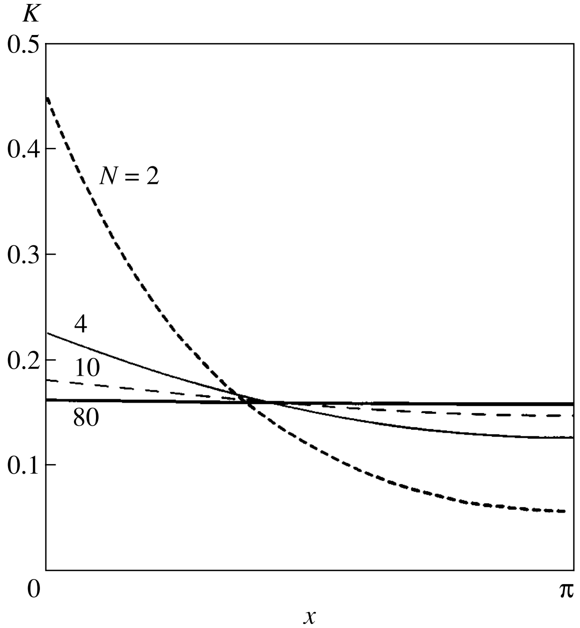

Fig. 13. The density–density correlation function of the system for $\omega = 0$ and $\gamma = 0.25 < \gamma_c = 1/2$ (within the plateau region) for 2, 4, 10, and 80 particles. Correlations increase as the number of particles increases.

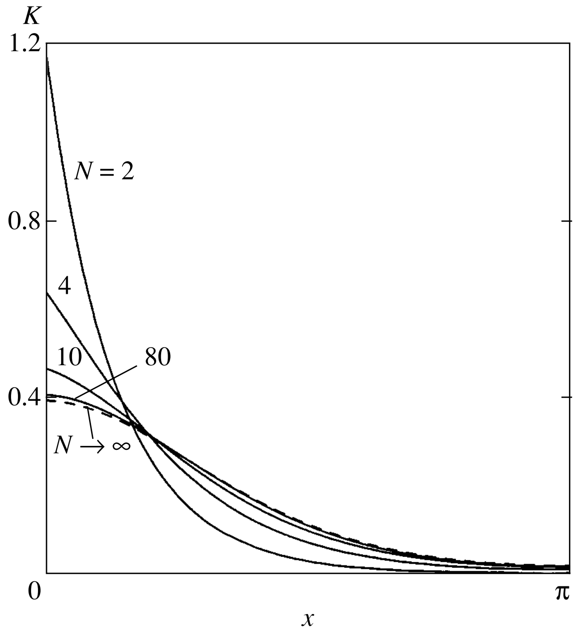

Fig. 14. The density–density correlation function of the system for $\omega = 0$ and $\gamma = 0.75 > \gamma_c = 1/2$ (outside the plateau region) for 2, 4, 10, and 80 particles. The dashed curve represents the solution to the Gross–Pitaevskii equation.

4. CONCLUSIONS

In this study, we have presented a variant of the DQMC method in the momentum representation, which is very efficient for investigating weakly interacting systems. This method can be applied to new quantum models that could not be studied by other cluster methods. We have presented arguments in favor of weakening the sign problem as compared with standard trajectory methods; some models are free of the sign problem. One of such models—a Bose gas with attraction in a thin rotating ring—has been investigated by the QMC method for the first time. The algorithm proposed has allowed us to follow up the variations of the ground state of this system, depending on the rotation speed of a container and the magnitude of interaction for a number of particles ranging from 2 to 100 and confirm the existence of the Hess–Fairbank effect [14] in this system. In the limit as $N \longrightarrow \infty$, we have obtained an exact equation for the phase boundary of the effect, also fitting well the case of a comparatively small number of particles ($N \sim 10$).

ACKNOWLEDGMENTS

I am grateful to V.A. Kashurnikov, B.V. Svistunov, and N.V. Prokof’ev for attention to the problem and valuable discussions.

This work was supported by the Russian Foundation for Basic Research, project nos. 03-02-16979 and 01-02-16508.

REFERENCES

1. N. V. Prokof’ev and B. V. Svistunov, Phys. Rev. Lett. **87**, 160601 (2001).
2. E. A. Burovski, A. S. Mishchenko, N. V. Prokof’ev, and B. V. Svistunov, Phys. Rev. Lett. **87**, 186402 (2001).
3. P. C. E. Stamp and I. S. Tupitsyn, cond-mat/0302015.
4. Yu. M. Kagan, V. A. Kashurnikov, A. V. Krasavin, *et al.*, Phys. Rev. A **61**, 043608 (2000).
5. V. A. Kashurnikov, N. V. Prokof’ev, and B. V. Svistunov, Phys. Rev. Lett. **87**, 120402 (2001).
6. N. Furukawa and M. Imada, J. Phys. Soc. Jpn. **60**, 810 (1991).
7. E. H. Lieb and F. Y. Wu, Phys. Rev. Lett. **20**, 1445 (1968).
8. J. E. Hirsch, D. J. Scalapino, R. L. Sugar, and R. Blankenbecler, Phys. Rev. Lett. **47**, 1628 (1981).
9. C. Lancsoz, J. Res. Natl. Bur. Stand. **45**, 244 (1950); E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).
10. J. E. Hirsch, Phys. Rev. B **67**, 035103 (2003).
11. N. V. Prokof’ev, B. V. Svistunov, and I. S. Tupitsyn, Zh. Éksp. Teor. Fiz. **114**, 570 (1998) [JETP **87**, 310 (1998)].
12. N. Kawashima and J. E. Gubernatis, Phys. Rev. B **50**, 136 (1994).
13. N. V. Prokof’ev, B. V. Svistunov, and I. S. Tupitsyn, Phys. Lett. A **238**, 253 (1998).
14. M. Ueda and A. J. Leggett, Phys. Rev. Lett. **83**, 1489 (1999).
15. P. F. Kartsev, cond-mat/0211356.
16. C. C. Bradley, C. A. Sackett, and R. G. Hulet, Phys. Rev. Lett. **78**, 985 (1997); C. A. Sackett and R. G. Hulet, Phys. Rev. Lett. **74**, 1315 (1995).
17. J. L. Roberts, N. R. Claussen, and S. L. Cornish, Phys. Rev. Lett. **86**, 4211 (2001).
18. R. A. Duine and H. T. C. Stoof, Phys. Rev. Lett. **86**, 2204 (2001).
19. J. B. McGuire, J. Math. Phys. **5**, 622 (1964).
20. E. M. Wright, J. Arlt, and K. Dholakia, Phys. Rev. A **63**, 013608 (2001).
21. A. S. Arnold and E. Riis, cond-mat/0110295.
22. G. P. Berman, A. Smerzi, and A. R. Bishop, Phys. Rev. Lett. **88**, 120402 (2002).
23. H. Hasimoto, J. Fluid Mech. **51**, 477 (1972).
24. Y. Castin and Ch. Herzog, cond-mat/0012040.
25. I. Carusotto and Y. Castin, Phys. Rev. Lett. **90**, 030401 (2003).
26. Z. Néda and Z. Dezcö, cond-mat/9912383.
27. E. M. Lifshitz and L. P. Pitaevskiĭ, *Statistical Physics* (Butterworth, Oxford, 1998; Fizmatlit, Moscow, 2001), Part 2.
28. E. P. Gross, Nuovo Cimento **20**, 454 (1961); L. P. Pitaevskiĭ, Zh. Éksp. Teor. Fiz. **40**, 646 (1961) [Sov. Phys. JETP **13**, 451 (1961)].
29. R. Kanamoto, H. Saito, and M. Ueda, cond-mat/0210229.

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, pp. 846–857.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 4, 2003, pp. 943–955.
Original Russian Text Copyright © 2003 by Krivitskiĭ, Kulik, Penin, Chekhova.

MISCELLANEOUS

Biphotons as Three-Level Systems: Transformation and Measurement

L. A. Krivitskiĭ, S. P. Kulik*, A. N. Penin, and M. V. Chekhova
Moscow State University, Vorob'evy gory, Moscow, 119992 Russia
**e-mail: skulik@qopt.phys.msu.su*

Received April 28, 2003

Abstract—Two algorithms for measuring the polarization state of the biphoton field prepared in the form of a three-level system (qutrit) are considered. On the basis of the general approach developed by Klyshko [21] for describing the polarization properties of single-mode electromagnetic fields in the fourth order in the field, a procedure for measuring the polarization density matrix of qutrits is proposed and implemented. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In a number of protocols of quantum information, use is made of states of light with definite (i.e., preset) properties. For instance, the Bell states [1] are used in quantum teleportation, high-density coding, or quantum cryptography. Such states are obtained as a result of spontaneous parametric down-conversion of light [2–4]. Polarization-squeezed states, in which the fluctuations of one or several Stokes parameters are suppressed, are widely discussed in the literature [5, 6]. A particular case of such states is "scalar light" for which fluctuations of all Stokes parameters are suppressed [7] (light of this type was recently obtained in experiment [8]). In quantum cryptography [9], the security in the key distribution increases when multilevel (in particular, three-level) states of light are used [10–12]. The manifestation of the geometrical phase for such (three-level) optical systems is of certain interest [13, 14].

Obviously, the application of states of light with preset properties in certain experiments presumes the solution of the following three interconnected problems:

(i) generation of such states;

(ii) their transformation during transmission over a communication channel;

(iii) application of a reliable procedure for controlling a state at a given instant.

In recent years, the procedure known as quantum tomography of light (*tomos* stands for layer and *grapho* means to write). This procedure aims at reconstructing the initial state of an electromagnetic field by measuring several projections of this state in different bases. Such a state can be recorded with the help of a wave function, density matrix, or quasiprobability function. It is probably more appropriate to apply the term "quantum tomography" to the quasiprobability function since this function permits a visual representation of the state in the form of a three-dimensional image. However, analysis of the literature shows that the term 'tomography' is being used now in a wider sense as reconstruction of the initial state. Among a large number of publications in this field, we will mention only those in which this procedure was used directly in experiment. For example, the reconstruction of the initial quasiprobability function was carried out in [15–17] during the measurement of states characterized by continuous variables, such as quadrature- or polarization-squeezed light. Tomography of the states of discrete variables, which are realized with the help of polarization-spatial qubits, was carried out in [18]. In this case, the polarization state of the biphoton field generated in the frequency-degenerate noncollinear mode can be treated as a pair of qubits. The first attempts at implementing qutrit tomography were made in [19].

2. BIPHOTON AS A THREE-LEVEL SYSTEM (QUTRIT)

This study is devoted to an analysis of the procedure for transformation and measurement of the states of light that can be described in terms of a three-level system. We consider the polarization state of frequency-degenerate biphoton radiation in a single spatial mode. In the general case, such states can be prepared using spontaneous parametric down-conversion of light from three nonlinear crystals oriented in a special way.

The states of light generated during spontaneous parametric down-conversion can be represented in the form

$$|\Psi\rangle = |\mathrm{vac}\rangle + \frac{1}{2}\sum_{k,k'} F_{k,k'}|1_k, 1_{k'}\rangle, \tag{1}$$

where $|\mathrm{vac}\rangle$ is the vacuum state, the quantity $F_{k,k'}$ is called the biphoton amplitude, and $|1_k, 1_{k'}\rangle$ is a state with one (signal) photon in mode k and one (idler) photon in mode k'. The modes are characterized by frequency, direction, and polarization. We will henceforth assume that the modes differ only in the polarization. The meaning of quantity $F_{k,k'}$ is that the square of its modulus gives the probability of recording two photons in the two polarization modes k and k'. Real states of the biphoton field possess fixed frequency and spatial (angular) spectra. When we speak of the single-mode approximation in frequency and angle, we assume that states (1) cannot be distinguished in respect to these parameters in a concrete experiment. In other words, for different frequencies and spatial components of such radiation, no delays exceeding the corresponding coherence scales of the field can emerge during its preparation and propagation.

Disregarding the vacuum component and taking into account all possible ways in which two photons can be distributed (that are indistinguishable in frequency and direction of propagation) between the two polarization modes, we obtain

$$\begin{aligned} |\Psi\rangle &= c_1|H, H\rangle + c_2|H, V\rangle + c_3|V, V\rangle \\ &= c_1|2, 0\rangle + c_2|1, 1\rangle + c_3|0, 2\rangle. \end{aligned} \qquad (2)$$

Here, we used two equivalent representations of state (1) in the Fock basis. For example, notation $|H, H\rangle = |2, 0\rangle$ indicates that both photons are in the horizontal polarization mode. The complex amplitudes of states ($c_i = d_i \exp(i\varphi_i)$, $i = 1, 2, 3$) satisfy the normalization condition

$$\sum_{i=1}^{3} |c_i|^2 = 1. \qquad (3)$$

State (2) describes a three-level system; consequently, the biphoton field can be treated as the object of investigation of the properties of this system. Since the state of a two-level system is known as a qubit in the quantum theory of information, state (2) will be referred to as a quantum trit, or qutrit. A detailed description of the characteristics of biphotons (qutrits) and their visual geometrical representation can be found in [20].

We assume that polarization properties of a single-mode electromagnetic field are described by the creation ($a_H^\dagger \equiv a^\dagger$, $a_V^\dagger \equiv b^\dagger$) and annihilation ($a_H \equiv a$, $a_V \equiv b$) operators for photons in the polarization modes H and V. These operators satisfy the conventional commutation relations

$$[a, a^\dagger] = [b, b^\dagger] = 1, \quad [a, a] = [b, b] = [b^\dagger, a] = 0.$$

It was shown in [21] that the parameters determining the polarization properties of a single-mode field in the fourth order can be combined into a Hermitian matrix known as the coherent matrix of the fourth order in the field,

$$K_4 = \begin{pmatrix} A & D & E \\ D^* & C & F \\ E^* & F^* & B \end{pmatrix}, \qquad (4)$$

with six moments as components. The diagonal components are formed by the real moments

$$A \equiv \langle a^{\dagger 2} a^2 \rangle, \quad B \equiv \langle b^{\dagger 2} b^2 \rangle, \quad C \equiv \langle a^\dagger a b^\dagger b \rangle. \qquad (5)$$

These moments characterize the correlation of intensities in polarization modes H and V. In the general case, the nondiagonal components are complex-valued:

$$D \equiv \langle a^{\dagger 2} a b \rangle, \quad F \equiv \langle a^\dagger b^\dagger b^2 \rangle, \quad E \equiv \langle a^{\dagger 2} b^2 \rangle. \qquad (6)$$

In particular, for a single-mode biphoton field, averaging is carried out over wave function (2), and moments (5) and (6) assume the form

$$\begin{gathered} A = 2|c_1|^2 = 2d_1^2, \quad B = 2|c_3|^2 = 2d_3^2, \\ C = |c_2|^2 = d_2^2, \end{gathered} \qquad (7)$$

$$D = \sqrt{2} c_1^* c_2, \quad F = \sqrt{2} c_2^* c_3, \quad E = 2 c_1^* c_3. \qquad (8)$$

It can be seen that moments (6) and (8) contain relative phases of three basis states. It should be noted that the moments defined in relations (7) and (8) coincide except for constants with the components of the density matrix (2),

$$\rho \equiv |\Psi\rangle\langle\Psi|, \quad \rho_{mk} = c_m c_k^*, \qquad (9)$$

which will be referred to as the polarization density matrix (i.e., the matrix written in the polarization basis). The condition $\rho^2 = \rho$ of the purity of the state imposes additional constraints on the moments. For example, we obtain the constraint

$$ABC = DFE^*, \qquad (10)$$

which makes it possible, for example, to eliminate moment E.

In a mixed state, averaging over the wave function should be supplemented with averaging over the classical probability distribution function P_j, where P_j is the probability of finding the system in a pure state Ψ_j, $\sum_j P_j = 1$.

Since the wave function is defined to within a constant phase factor, we have only two independent phases characterizing state (2), viz., $\varphi_{21} = \varphi_2 - \varphi_1$ and $\varphi_{31} = \varphi_3 - \varphi_1$.

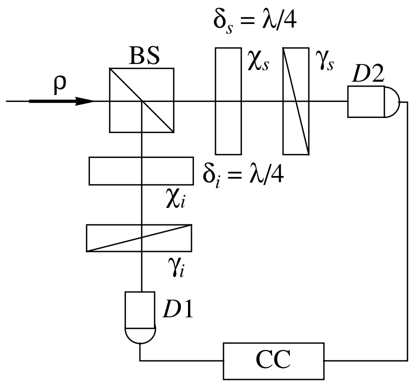

Fig. 1. Schematic diagram of the setup for measuring an arbitrary state ρ of the qutrit: BS is the beam splitter, and $D1$ and $D2$ are photodetectors with a quarter-wave plate and a rotating polarizer in front of each detector; photocurrent pulses from the detectors are fed to the pair coincidence circuit CC, detecting moment R.

Thus, in order to define a pure state of the three-level system (2), we must define $2S - 2 = 4$ real parameters. In order to define mixed states, $S^2 - 1 = 8$ real-valued parameters are required. Here, $S = 3$ is the dimension of the Hilbert space of a biphoton (qutrit).

3. QUTRIT TOMOGRAPHY

The idea of measuring the parameters of state (2) can be formulated as follows. The initial unknown state is subjected to preset polarization transformations so that the final state, which is fed to the input of the recording system, corresponds to the known combination of moments (5) and (6). Since we are speaking of the measurement of fourth-order moments in a field, the system of recording of biphotons (qutrits) must consist of a beam splitter and a pair of detectors with inputs connected to the photocount coincidence circuit (the Brown–Twiss circuit [22]). We can consider two methods for carrying out polarization transformations.

Method 1. First, the initial biphoton beam is split into two channels by a beam splitter (Fig. 1).[1] These channels will be referred to as the signal (s) and idler (i) channels. Further, in each spatial mode of the beam splitter, transformations on the signal photon and the idler photon forming the initial biphoton take place. Transformations are carried out with the help of quarter-wave phase plates and polarization prisms (polarizers). We can write these polarization transformations in the Heisenberg representation.

The action of the nonpolarizing beam splitter executing the spatial separation of signal and idler photons is described by the relation

$$\begin{pmatrix}a\\b\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}a\\b\end{pmatrix}^s + \frac{1}{\sqrt{2}}\begin{pmatrix}a\\b\end{pmatrix}^i. \qquad (11)$$

Then the complete transformation for each channel has the form

$$\begin{pmatrix}a'\\b'\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\sin^2\gamma & \cos\gamma\sin\gamma\\ \cos\gamma\sin\gamma & \cos^2\gamma\end{pmatrix}\begin{pmatrix}t & r\\ -r^* & t^*\end{pmatrix}\begin{pmatrix}a\\b\end{pmatrix}. \qquad (12)$$

Here, a and b are the photon annihilation operators in polarization modes H and V prior to transformation, γ is the angle of orientation of the polarizer relative to the vertical axis, and r and t are the amplitude coefficients of reflection and transmission of $\lambda/4$ phase plates, respectively,

$$t = \cos\delta + i\sin\delta\cos2\chi, \quad r = i\sin\delta\sin2\chi. \qquad (13)$$

Here we have introduced the parameter $\delta = \pi(n_o - n_e)h/\lambda$ of the plates (h is the plate thickness) and χ is the angle of their rotation relative to the vertical axis. This gives

$$t = (1 + i\cos2\chi)/\sqrt{2}, \qquad (14)$$

$$r = i\sin2\chi/\sqrt{2}. \qquad (15)$$

Thus, we have four parameters (two for each channel) defining the polarization transformations. These are the angles χ and γ of rotation of the plates and the polarizer.

In experiments with a biphoton field, we measure the number of photocount coincidences for the detectors mounted in the signal and idler channels or the fourth-order moment of the form

$$R_{s,i} \propto \langle (b'_s)^\dagger (b'_i)^\dagger b'_s b'_i \rangle = R(\chi_s, \gamma_s, \chi_i, \gamma_i), \qquad (16)$$

where b'_s (b'_i) is the photon annihilation operator in a mode tuned through angle γ_s (γ_i). In the general case, this moment is a combination of six moments forming matrix K_4.

The task of polarization tomography is to find six moments, (5) and (6), from relations of the form (16) upon a change in parameters χ_s, γ_s, χ_i, and γ_i of the phase plates and polarizers.

Several remarks can be made concerning the experimental procedure of measurement of a state.

1. We assume that the source generating biphotons prepares them in state (2), which is stationary. Each act of measurement is accompanied by the destruction of such a state. However, the experimenter has a sufficiently large set of copies of the initial state, so that the

---

[1] We will take into account only the events leading to the coincidence of photocounts, i.e., the events when the signal and idler photons fall into different input modes of the beam splitter. This takes place in half the total number of trials.

Table

ν	χ_s	γ_s	χ_i	γ_i	Field moment being measured
1	0°	–90°	0°	–90°	$A/4$
2	0°	–90°	0°	0°	$C/4$
3	0°	0°	0°	0°	$B/4$
4	45°	0°	0°	0°	$(B + C + 2\mathrm{Im}F)/8$
5	45°	–45°	0°	0°	$(B + C - 2\mathrm{Re}F)/8$
6	45°	–45°	0°	–90°	$(A + C - 2\mathrm{Re}D)/8$
7	45°	0°	0°	–90°	$(A + C + 2\mathrm{Im}D)/8$
8	–45°	–22.5°	45°	22.5°	$(A + B - 2\mathrm{Im}E)/16$
9	45°	–45°	45°	45°	$(A + B - 2\mathrm{Re}E)/16$

destruction of some of these states does not affect the remaining ones. Such an “ensemble” approach leads to the conclusion that in each next measurement we are dealing with the same unperturbed state and, hence, the measurement results provide information on the initial state.

2. The number of measurements required for complete reconstruction of arbitrarily defined state (2) is determined by the number of independent real numbers defining this state. However, in experiment, we are dealing with nonnormalized states and, hence, the number of measurements increases. Normalization is set for measuring moments A, B, and C defined by formulas (5) and (7). Using relations (3) and (7), we obtain

$$A + 2C + B = 2. \tag{17}$$

The remaining three moments, D, E, and F, carry information about phases $\varphi_{21} = \varphi_2 - \varphi_1$, $\varphi_{31} = \varphi_3 - \varphi_1$, and $\varphi_{32} = \varphi_{31} - \varphi_{21}$. However, since these moments are complex-valued, their real and imaginary parts (which are associated with cosine and sine of the phases, respectively) must be measured separately. Thus, the number of measurements required for complete reconstruction of state (2) is equal to seven for a pure state and nine for a mixed state.

3. In order to reduce the effect of errors on the parameters of the state being measured, we must minimize the number of moments appearing in relation (16).

Let us consider several examples.

I. Let us suppose that $\chi_s = 0$, $\gamma_s = -90°$, $\chi_i = 0$, and $\gamma_i = 0$. In this case, we have

$$t_s = \frac{1+i}{\sqrt{2}}, \quad r_s = 0, \quad t_i = \frac{1+i}{\sqrt{2}}, \quad r_i = 0.$$

Then it follows from relations (12) and (16) that $R_{s,\,i} \propto C/4$.

II. Let us suppose that $\chi_s = -45°$, $\gamma_s = -22.5°$, $\chi_i = 45°$, and $\gamma_i = 22.5°$. In this case, we have

$$t_s = \frac{1}{\sqrt{2}}, \quad r_s = -\frac{i}{\sqrt{2}}, \quad t_i = \frac{1}{\sqrt{2}}, \quad r_i = \frac{i}{\sqrt{2}}.$$

We find that $R_{s,\,i} \propto (A + B - 2\mathrm{Im}E)/16$.

The complete set of measurements required for determining the unknown state (2) can be represented in the form of a table. The experimental setup corresponding to this set contains a quarter-wave plate and a rotating polarizer each in the signal and idler channels. It should be noted that the number of polarization transformers appearing in each channel of the circuit (see Fig. 1) is equal to two and not to three, as proposed in [18], where an additional half-wave plate was used. This is possible since the action of a half-wave plate with parameter $\chi_{\lambda/2}$ and a polarizer with the fixed orientation $\gamma \equiv 90°$ is equivalent to the action of a single rotatable polarizer with parameter $\gamma = -2\chi_{\lambda/2}$.

Thus, the first seven rows of the table contain orientations of the plates and polarizers, which makes it possible to measure three real (A, B, C) and two complex (D, F) moments. These measurements are sufficient for complete reconstruction of a pure state. If the state is mixed, two more measurements must be made and the complex moment E must be determined. The verification of condition (10) can answer the question concerning the purity of the state.

Method 2. The polarization transformations are carried out on biphotons (qutrits) as integral objects. Then these objects are directed to the modified Brown–Twiss circuit, where their spatial separation and detection takes place (Fig. 2). The primary beam splitter (*PBS*) is of the polarization type: it completely reflects light with the vertical (*V*) polarization and transmits light with the horizontal (*H*) polarization. The nonpolarizing beam splitters *BS*1 and *BS*2 mounted behind the *PBS* spatially separate light of the same polarization, while detectors

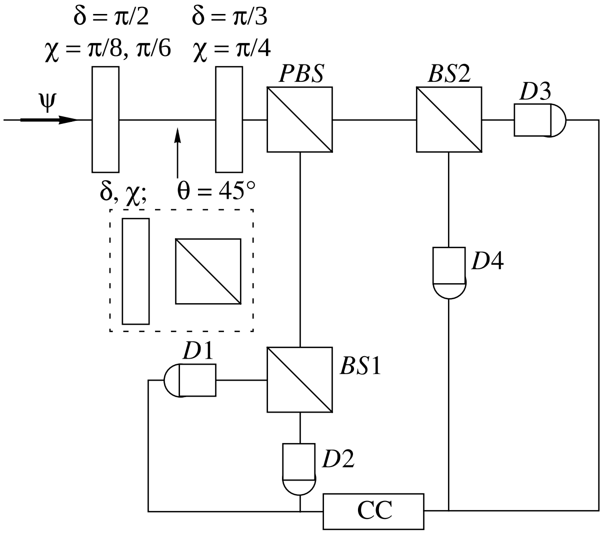

Fig. 2. Schematic diagram of a setup for partial tomography of qutrit state Ψ. The polarization transformations on biphotons as integral objects are carried out with the help of λ/2 and λ/3 phase plates. Beam splitters *PBS*, *BS*1, and *BS*2 are intended for the spatial separation of signal and idler photons with parallel or orthogonal polarizations. The coincidence circuit CC registers the number of coincidences between all possible pairs of detectors: *D*1–*D*2, *D*1–*D*3, *D*1–*D*4, *D*2–*D*3, *D*2–*D*4, and *D*3–*D*4. The dashed contour encircles a phase plate and a polarizer oriented at an angle of θ = 45°, which are used in an additional protocol.

are placed in their output modes. As a result, the coincidence counting rate between photocounts of the pair of detectors *D*1–*D*2 is proportional to the fourth-order moment of the field,

$$R_{D1-D2} \propto \langle b^{\dagger 2} b^2 \rangle \equiv B',$$

while that between detectors *D*3–*D*4 is proportional to the moment

$$R_{D3-D4} \propto \langle a^{\dagger 2} a^2 \rangle \equiv A'.$$

Coincidences between photocounts of any pairs of detectors (*D*1–*D*3, *D*1–*D*4, *D*2–*D*3, and *D*2–*D*4) give the value of the moment for fields with orthogonal polarizations:

$$R_{D1,\,2-D3,\,4} \propto \frac{\langle a^{\dagger} a b^{\dagger} b \rangle}{4} \equiv C'.$$

Let us consider the polarization transformations carried out in the circuit depicted in Fig. 2 in the Schrödinger representation. For this, we write the wave function (2) in the form of a column whose elements are normalized amplitudes c_i (i = 1, 2, 3). The transformations executed over these polarization states of the biphoton field by phase plates can be described by the unitary 3 × 3 matrix [23]

$$G \equiv \begin{pmatrix} t^2 & \sqrt{2}tr & r^2 \\ -\sqrt{2}tr^* & |t|^2 - |r|^2 & \sqrt{2}t^*r \\ r^{*2} & -\sqrt{2}t^*r^* & r^{*2} \end{pmatrix}, \qquad (18)$$

so that Ψ' = GΨ or

$$\begin{pmatrix} c'_1 \\ c'_2 \\ c'_3 \end{pmatrix} = G \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}. \qquad (19)$$

Coefficients t and r were introduced by relations (13).

In the Schrödinger representation, the moments A', B', and C' being measured contain averaging over the wave function Ψ' connected with the initial wave function Ψ via transformation (19). Our aim is to establish one-to-one correspondence between the results of measurements of moments A', B', and C' and moments (5) and (6) prior to the polarization transformations.

Let us consider the action of various transformers on biphotons.

1. The λ/2 plate: δ = π/2, χ = 22.5°. In this case, we have

$$t = r = \frac{i}{\sqrt{2}},$$

$$G \equiv \begin{pmatrix} -1/2 & -1/\sqrt{2} & -1/2 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/2 & 1/\sqrt{2} & -1/2 \end{pmatrix}. \qquad (20)$$

The three moments being measured can be expressed with the help of relations (19) and (20) in terms of the parameters of input state Ψ:

$$A' = 2|c'_1|^2 = \frac{1}{2}[d_1^2 + d_3^2 + 2d_1d_3\cos(\varphi_1 - \varphi_3)] + d_2^2 + \frac{2}{\sqrt{2}}d_1d_2\cos(\varphi_1 - \varphi_2) + \frac{2}{\sqrt{2}}d_2d_3\cos(\varphi_2 - \varphi_3), \qquad (21)$$

$$B' = 2|c'_3|^2 = \frac{1}{2}[d_1^2 + d_3^2 + 2d_1d_3\cos(\varphi_1 - \varphi_3)] + d_2^2 - \frac{2}{\sqrt{2}}d_1d_2\cos(\varphi_1 - \varphi_2) - \frac{2}{\sqrt{2}}d_2d_3\cos(\varphi_2 - \varphi_3), \qquad (22)$$

$$C' = \frac{1}{4}|c'_2|^2 = \frac{1}{8}[d_1^2 + d_3^2 - 2d_1d_3\cos(\varphi_1 - \varphi_3)]. \qquad (23)$$

It should be noted that the last expression demonstrates the polarization anticorrelation effect in explicit form.

The number of coincidences of photocounts between the detectors recording radiation in orthogonal polarization modes decreases to zero if the number of biphotons in the vertical polarization mode is equal to the number of biphotons in the horizontal polarization mode (7) (i.e., if $A = B$) and the phase shift between basis states $|2, 0\rangle$ and $|0, 2\rangle$ vanishes ($\varphi_1 = \varphi_3$).

2. The $\lambda/2$ plate: $\delta = \pi/2$, $\chi = 30°$. In this case, we have

$$t = \frac{i}{2}, \quad r = i\frac{\sqrt{3}}{2},$$

$$\begin{aligned} A' = 2|c'_1|^2 = \frac{1}{8}\{&d_1^2 + 6d_2^2 + 9d_3^2 \\ &- 2\sqrt{6}d_1d_2\cos(\varphi_1 - \varphi_2) + 6d_1d_3\cos(\varphi_1 - \varphi_3) \\ &- 6\sqrt{6}d_2d_3\cos(\varphi_2 - \varphi_3)\}, \end{aligned} \tag{24}$$

$$\begin{aligned} B' = 2|c'_3|^2 = \frac{1}{8}\{&9d_1^2 + 6d_2^2 + d_3^2 \\ &+ 6\sqrt{6}d_1d_2\cos(\varphi_1 - \varphi_2) + 6d_1d_3\cos(\varphi_1 - \varphi_3) \\ &+ 2\sqrt{6}d_2d_3\cos(\varphi_2 - \varphi_3)\}. \end{aligned} \tag{25}$$

3. The $\lambda/3$ plate: $\delta = \pi/3$, $\chi = 45°$. In this case, we have

$$t = \frac{1}{2}, \quad r = i\frac{\sqrt{3}}{2}.$$

$$\begin{aligned} A' = 2|c'_1|^2 = \frac{1}{8}\{&9d_1^2 + 6d_2^2 + d_3^2 \\ &- 6\sqrt{6}d_1d_2\sin(\varphi_1 - \varphi_2) - 6d_1d_3\cos(\varphi_1 - \varphi_3) \\ &- 2\sqrt{6}d_2d_3\sin(\varphi_2 - \varphi_3)\}, \end{aligned} \tag{26}$$

$$\begin{aligned} B' = 2|c'_3|^2 = \frac{1}{8}\{&d_1^2 + 6d_2^2 + 9d_3^2 \\ &+ 2\sqrt{6}d_1d_2\sin(\varphi_1 - \varphi_2) - 6d_1d_3\cos(\varphi_1 - \varphi_3) \\ &+ 6\sqrt{6}d_2d_3\sin(\varphi_2 - \varphi_3)\}. \end{aligned} \tag{27}$$

4. Finally, in the absence of a transformer, when $G \equiv I$ (I is the identity matrix), the moduli of the amplitudes of basis states (2),

$$A' \equiv A = 2d_1^2, \quad B' \equiv B = 2d_3^2, \quad 4C' \equiv C = d_2^2, \tag{28}$$

can be established in three dimensions.

Thus, for the circuit considered here, the protocol of polarization tomography of qutrits is as follows:

(i) transformer is absent; the real amplitudes of states d_1, d_2, and d_3 are determined from relations (28);

(ii) transformer $\delta = \pi/2$, $\chi = 22.5°$; from relation (23), we find $\cos(\varphi_1 - \varphi_3)$;

(iii) transformer $\delta = \pi/2$, $\chi = 30°$; from relations (24) and (25), we find $\cos(\varphi_1 - \varphi_2)$ and $\cos(\varphi_2 - \varphi_3)$;

(iv) transformer $\delta = \pi/3$, $\chi = 45°$; from relations (26) and (27), we find $\sin(\varphi_1 - \varphi_2)$ and $\sin(\varphi_2 - \varphi_3)$.

Comparing two of the above protocols, we note that the second protocol is less preferable from the viewpoint of the accuracy that can be attained. This is due to the fact that expressions (24)–(27) for the cosines and sines of the relative phases of the basis states contain a large number of quantities being measured, which increases the error in their determination. In addition, the second method is applicable only to pure states (2); in this case, we can speak only of partial tomography of qutrits. On the other hand, the circuit depicted in Fig. 2 is more convenient since it permits the simultaneous measurement of three moments (A', B', and C') for a given transformer and, hence, the entire protocol can be realized in a shorter time. This factor should probably be taken into account for using the results of tomography in a more complex protocol (e.g., in quantum cryptography on the basis of qutrits [10]).

It should also be noted that the choice of the polarization transformers in the second tomographic method is not limited to $\lambda/2$ and $\lambda/3$ plates. Any phase plate can ensure the relation between the observed (A', B', C') and input (A, B, C) moments. The only exception is the circuit with only one half-wave plate ($t = i\cos 2\chi$, $r = i\sin 2\chi$); in this case, only cosines of the corresponding phases can be determined.

The modified version of the circuit depicted in Fig. 2 gives the relations

$$A, B, C \xrightarrow{f(\delta, \chi;\, \theta = 45°)} A', B', C'$$

obtained using a single arbitrary phase plate with parameters δ and χ and a polarizer separating the linear polarization at an angle of 45°, which are mounted in front of the polarization beam splitter enclosed in the dashed contour in Fig. 2). However, the calculation of inverse transformers using such a circuit is quite cumbersome and is not included in this work; we just mention that such a method of qutrit tomography is possible in principle.

4. EXPERIMENT

Below, we describe the procedure of qutrit tomography, in which polarization transformations are carried out over the signal and idler photons separately after their spatial separation.

The schematic diagram of the experimental setup for single-mode biphoton tomography is shown in Fig. 3. Pumping is carried out by a cw argon laser with wavelength $\lambda = 351$ nm and power $P = 120$ mW. The pumping beam is directed to a nonlinear lithium iodate ($LiIO_3$) crystal $l = 1$ cm in length; in the bulk of this

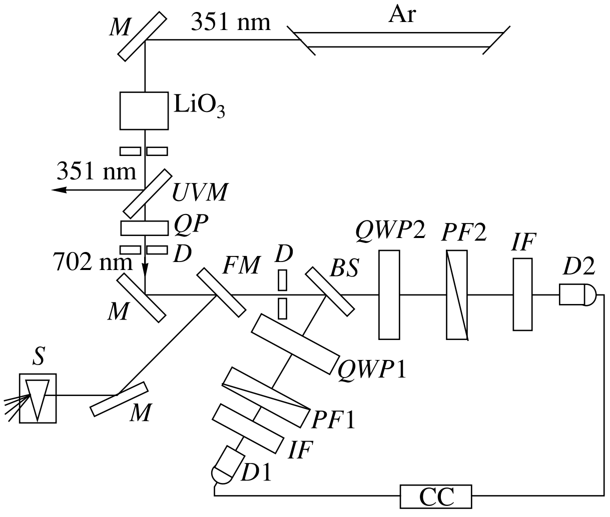

Fig. 3. Experimental setup: Ar is an argon laser; *M* are mirrors; LiIO$_3$ is a nonlinear crystal; *UVM* is a mirror reflecting the ultraviolet pumping radiation and transmitting the biphoton radiation; *QP* is the setting phase plate; *D* are diaphragms; *FM* is the folding mirror; *BS* is the beam splitter; *QWP*1 and *QWP*2 are quarter-wave plates; *PF*1 and *PF*2 are polarization prisms; *IF* are interference filters; *D*1 and *D*2 are photodetectors; and *S* is the spectrograph ISP-51 intended for controlling the biphoton field spectrum.

crystal, spontaneous parametric down-conversion takes place. The crystal is oriented so that biphotons are emitted in the frequency-degenerate collinear regime. The polarizations of both photons are orthogonal to the polarization of the pumping beam (type I phase-matching, or $e \longrightarrow oo$ interaction). In this regime, the optical axis of the crystal forms an angle of $\theta = 58°$ with the wave vector of pumping radiation. After the passage of pumping radiation through the crystal, it is extracted from the system by mirror *UVM* reflecting radiation at a wavelength of 351 nm and transmitting light at a double wavelength. A system of diaphragms *D* and an interference filter *IF* with a central wavelength of 700 nm and a full width at half maximum (FWHM) of 10 nm is used to separate one spatial and one frequency mode of the biphoton field. The spectral width of spontaneous parametric down-conversion for the given crystal is $\Delta\lambda \approx 20$ nm. The width of the spatial correlation function of spontaneous parametric down-conversion was determined in our experiments by the divergence of the pumping beam and was estimated as $\Delta\theta_p \approx 3 \times 10^{-4}$ rad. This quantity determined the diameter of diaphragms *D*. As a result of such a mode discrimination, the effect of the finite spatial frequency biphoton radiation spectrum on subsequent polarization transformations can be disregarded.

Various polarization states were prepared directly using a thin quartz plate *QP* (setting plate). Each value of the optical axis orientation of plate *QP* corresponds to a certain polarization state of biphotons fed to the input of the measuring system. The plate thickness is $h = 824 \pm 0.5$ μm; the radiation loss due to reflection from the plate faces amounts approximately to 8%, and the error in the orientation of its optical axis is approximately equal to 1°. The plate thickness was chosen so that the transformation executed by the plate depended on the wavelength insignificantly within the filter transmission band.

The measuring block is formed by an intensity interferometer with a 50% nonpolarization beam-splitting mirror *BS* and two FEU-79 detectors *D*1 and *D*2 (photomultipliers operating in the photon counting mode; the quantum efficiency is $\eta \sim 10^{-2}$). The beam-splitting mirror is mounted at a small angle (about 12°) to the beam so that the polarization state of light does not change as a result of reflection and transmission. A quarter-wave plate *QWP* and a rotating polarization filter *PF* (Glan-Thompson prisms) is installed in each arm of the interferometer. Zero-order quartz plates for a wavelength of $\lambda_s = 702$ nm with bleached faces are used. The loss introduced by the polarization filters amounts to 8–12%, and the error in the orientation of the optical elements is approximately equal to 2.0°. After amplification and amplitude discrimination, signals from detectors are supplied to the coincidence circuit CC with a resolution time of $T \approx 5$ ns. The frequency-angular spectrum of the biphoton field is monitored with the help of an ISP-51 spectrograph.

Since the SU2 transformation executed by the phase plate *QP* is used for preparing the initial states, it is impossible to obtain an arbitrary state of a biphoton (qutrit) in the experiment discussed here.[2] Nevertheless, the qutrit state behind the plate belongs to a subclass of states (2) with certain relations between parameters d_i and φ_i and is described by the expression

$$\begin{pmatrix} c_1 \\ c_3 \\ c_3 \end{pmatrix} = G \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} -0.9725\sin^2 2\alpha \\ \sqrt{2}\sin 2\alpha(0.9724\cos 2\alpha + i0.1636) \\ 0.0275 - 0.9724\cos^2 2\alpha + i0.3273\cos 2\alpha \end{pmatrix}, \qquad (29)$$

where matrix G for a *QP* plate with parameters $\Delta n = n_o - n_e = 0.0089$ and $\delta = 32.82$ was calculated using formula (18). Here, α is the angle of orientation of the plate relative to the vertical axis, which determines the state (29) of the qutrit and which is measured from the direction of polarization of the initial biphoton beam (not transformed by the quartz plate *QP*). In measurements 1–3 (see table), the diagonal components of matrix K_4 are determined, while a combination of diag-

---

[2] In order to obtain such a state, a transformation of the *SU*3 group is required; this can be done via nonlinear transformations which have an extremely low efficiency.

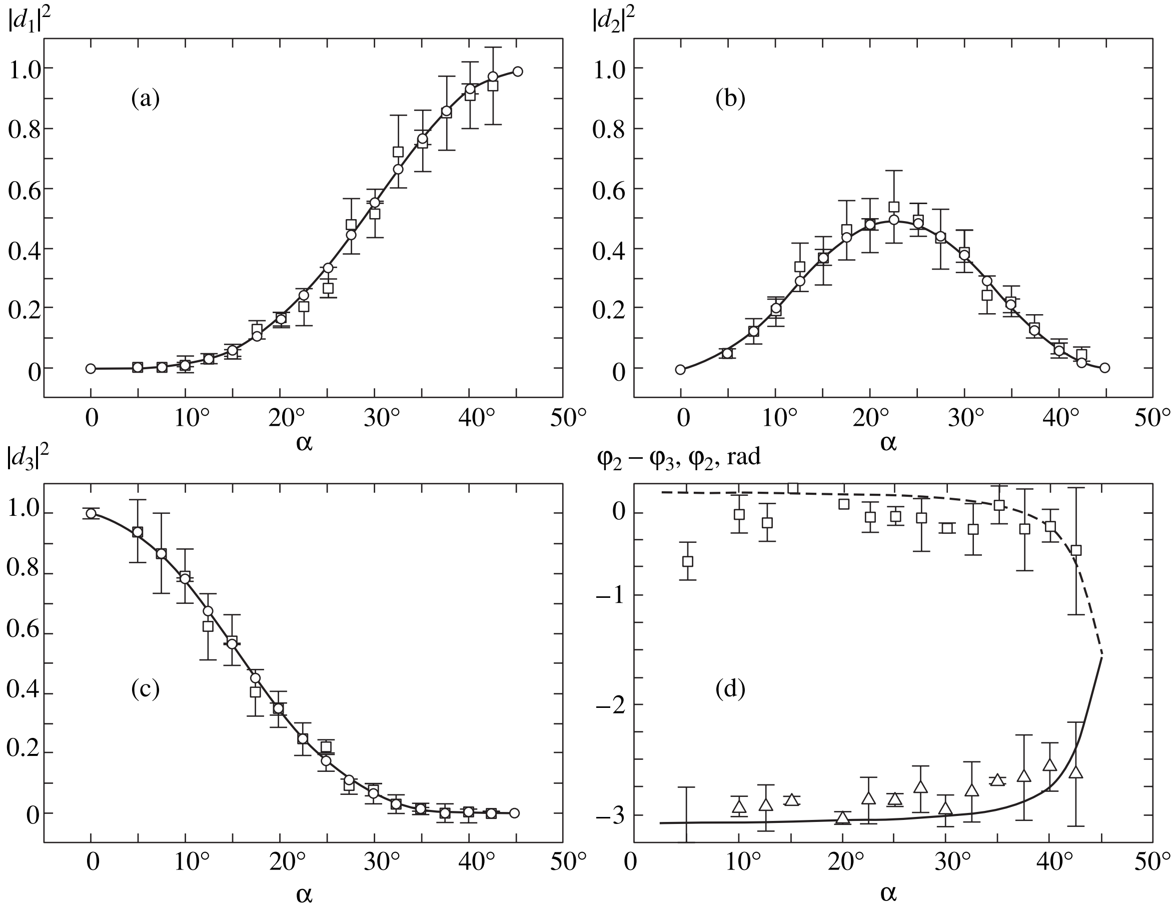

Fig. 4. The results of measurement of three amplitudes, d_1 (a), d_2 (b) and d_3 (c), and two phases, φ_2 (□), $\varphi_2 - \varphi_3$ (△) (d). Different initial states of qutrits correspond to different orientations α of setting plate *QP*. Solid and dashed curves correspond to calculated dependences.

onal and nondiagonal components is determined in measurements 4–9. It was shown in Section 3 that the first seven measurements are sufficient for a pure state. It should be noted that, proceeding to each next measurement, the orientation of only one polarization transformer has to be changed, which is convenient from the experimental point of view.

The results of experiments in the form of dependences of amplitudes d_i and phases φ_i of the states prepared by the quartz plate *QP* for various values of orientation α are represented in Fig. 4. Figure 5 shows the experimental dependences of the components of coherence matrix K_4 on α. The results of calculation of components d_i and φ_i (Fig. 4) and moments (Fig. 5) for different input states are shown for comparison. For a fixed value of the orientation $\alpha_0 = 25°$ of the setting plate, the polarization density matrix of the corresponding state was reconstructed experimentally:

$$\rho_{\text{exp}} = \begin{pmatrix} 0.271 & 0.345 + 0.074i & -0.24 - 0.114i \\ 0.345 - 0.074i & 0.508 & -00.316 - 0.075i \\ -0.24 + 0.114i & -0.316 + 0.075i & 0.221 \end{pmatrix}. \tag{30}$$

The set of eigenvalues for this matrix is as follows: $\lambda_1 = 0.99$, $\lambda_2 = -0.021$, and $\lambda_3 = 0.03$. The trace of the squared matrix (30) was found to be $\text{Tr}(\rho^2_{\text{exp}}) = 0.981$.

It can be seen from the curves shown in Figs. 4 and 5 that the maximal relative errors appear in the reconstructed values of phases of states in the nondiagonal elements of the density matrix also. This is due to the fact that the phases are calculated using simultaneously the results of several measurements. The errors are superimposed and make a considerable contribution to the result.

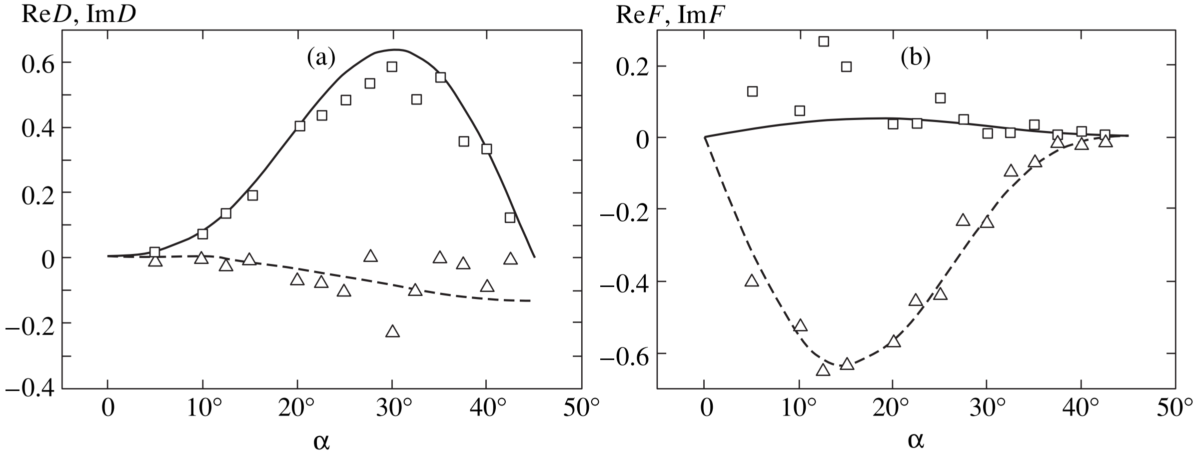

Fig. 5. The results of measurement of components of matrix K_4. Real (□) and imaginary (△) parts of moments *D* (a) and *F* (b). Different initial states of qutrits correspond to different orientations α of the setting plate *QP*. The solid and dashed curves correspond to calculated dependences.

Several remarks concerning the properties of the reconstructed density matrix ρ_{exp} are appropriate here. This matrix is Hermitian and normalized; i.e., $\rho_{\text{exp}}^{\dagger} = \rho_{\text{exp}}$ and $\text{Tr}(\rho_{\text{exp}}^2) = 1$. It follows from the general properties of the density matrix that it must have positive eigenvalues and satisfy the condition $0 \le \text{Tr}(\rho_{\text{exp}}^2) \le 1$. For the pure states studied here, there exists only one nondegenerate eigenvalue equal to unity and $\text{Tr}(\rho_{\text{exp}}^2) = 1$. Due to the effect of various experimental errors, reconstructed density matrix ρ_{exp} obviously does not meet the above requirements. In order to put into correspondence a "realistic" physical state satisfying the above properties to the experimentally measured matrix, we must carry out the procedure of the maximum likelihood estimation [18, 24]; this procedure and the results of its application are given in the Appendix.

Let us briefly consider the main sources of errors in our experiments.

First, there is a low quality of the phase plates (QWP) used for transforming the polarization state; the plate thickness does not satisfy the condition $\delta = \pi/4$. In addition, since experiments were performed with a biphoton field possessing a finite spectral width (10 nm), zero-order plates ensuring the independence of the transformation of the wavelength should be used. However, the setting plate *QP* was a higher-order plate and different spectral components of the biphoton field were transformed in different manners.

Second, we must mention the error in alignment of all polarization transformers. In our experiments, this error was about 2°, which gave rise to random errors in various realizations.

The above sources of errors are of technical rather than of fundamental nature. We believe that in the near future these errors will be eliminated from the measuring procedure discussed here.

5. CONCLUSIONS

We have considered two algorithms of measurement of the polarization state of a biphoton field prepared in the form of a three-level system (qutrit). On the basis of the general approach developed by Klyshko [21] for describing the polarization properties of single-mode electromagnetic fields in the fourth order in the field, we proposed and realized a procedure for measuring the polarization density matrix. As the initial states, we used a set of states that can be obtained with the help of spontaneous parametric down-conversion in a nonlinear crystal with a type I phase-matching in the frequency-degenerate collinear regime. We hope that the measurement procedure considered here will be useful in the realization of the quantum cryptography protocol based on qutrits [10] as well as in spectroscopic studies of the phase transition in ferroelectric crystals. Experimental data show that the anomalies observed in some crystals during the formation of the domain structure in the processes of elastic [25] and inelastic [26] small-angle scattering of light cannot be explained in the framework of the existing models. It would be interesting to apply the approach developed by us here in solving the inverse scattering problem for a phase transition.

ACKNOWLEDGMENTS

This study was supported financially by the Russian Foundation for Basic Research (project nos. 02-02-16664 and 03-02-16444) and INTAS (grant no. 2122-01).

APPENDIX

Procedure of Maximum Likelihood Estimation of Experimental Results

One of the results obtained in this study is the reconstruction of the density matrix of a given state from the experimental results. However, the reconstructed density matrix of the system does not correspond to a real pure biphoton state in view of various kinds of experimental errors. The procedure of maximum likelihood estimation (MLE) that will be described below makes it possible to indicate the pure biphoton state corresponding to the experimental result with the highest probability. We will apply the MLE method [18] realized for a system of two polarization-spatial qubits (the dimension of space $S = 4$). The algorithm considered below was realized for three-level systems, or qutrits (the dimension of space $S = 3$).

The MLE algorithm consists of the following three main parts.

1. Obtaining the expression for the density matrix of the three-level system in the general form. This matrix is a function of nine real variables (we denote this matrix as $\rho_{ph}(q_1, \ldots, q_9)$ and satisfies the following three properties: it is nonnegative definite, normalized, and Hermitian.

2. Introducing the likelihood function that shows the extent to which the obtained experimental data are close to the real "physical" state described by the density matrix $\rho_{ph}(q_1, \ldots, q_9)$. The likelihood function $\Im(q_1, \ldots, q_9; R_1, \ldots, R_9)$ depends on nine arguments $\{q_1, \ldots, q_9\}$ of the ideal density matrix and on nine experimental values $\{R_\nu\}$, where R_ν is the coincidence count rate of photocounts (16) and $\nu = 1, 2, \ldots, 9$ is the number of the corresponding measurement in the table.

3. Using the standard numerical optimization methods for the available experimental data $\{R_1, \ldots, R_9\}$, we obtain the set of variables $\{q_1^{\text{opt}}, \ldots, q_9^{\text{opt}}\}$ for which function $\Im(q_1, \ldots, q_9; R_1, \ldots, R_9)$ assumes the minimal value. In this case, the best estimate for the experimental density matrix has the same form as the ideal density matrix of the optimized set of variables; i.e., $\rho_{ph}(q_1^{\text{opt}}, \ldots, q_9^{\text{opt}})$. Let us consider the MLE procedure in greater detail.

"Physical" Density Matrix

The property of nonnegative definiteness for an arbitrary matrix can be represented in the form

$$\langle\Psi|U|\Psi\rangle \geq 0. \qquad \text{(A.1)}$$

The matrix written in the form $U = Q^\dagger Q$ is nonnegative definite and Hermitian. Indeed, substituting this matrix into inequality (A.1), we obtain

$$\langle\Psi|Q^\dagger Q|\Psi\rangle = \langle\Psi'|\Psi'\rangle \geq 0, \qquad \text{(A.2)}$$

where $|\Psi'\rangle = Q|\Psi\rangle$. The Hermiticity of the matrix is also obvious:

$$(Q^\dagger Q)^\dagger = Q^\dagger (Q^\dagger)^\dagger = Q^\dagger Q.$$

The matrix can be normalized as follows:

$$u = Q^\dagger Q/\mathrm{Tr}(Q^\dagger Q). \qquad \text{(A.3)}$$

A system with three degrees of freedom has a density matrix of dimension 3×3 with eight independent real-valued variables. For the convenience of subsequent computations, matrix Q is defined in the triangular form

$$Q(q_1, \ldots, q_9) = \begin{pmatrix} q_1 & 0 & 0 \\ q_4 + iq_5 & q_2 & 0 \\ q_8 + iq_9 & q_6 + iq_7 & q_3 \end{pmatrix}. \qquad \text{(A.4)}$$

Then, in accordance with Eq. (A.3), the physical density matrix can be written in the form

$$\rho_{ph}(q_1, \ldots, q_9) = Q^\dagger Q/\mathrm{Tr}(Q^\dagger Q). \qquad \text{(A.5)}$$

It is also necessary for subsequent computations to express the elements of matrix Q in terms of the elements of matrix ρ_{ph}, i.e., to obtain a relation inverse to (A.5):

$$Q \approx \begin{pmatrix} \sqrt{\rho_{11} - \dfrac{M_{12}M_{21}}{\rho_{33}M_{11}} - \dfrac{\rho_{31}\rho_{13}}{\rho_{33}}} & 0 & 0 \\ \dfrac{M_{12}}{\sqrt{M_{11}\rho_{33}}} & \sqrt{\dfrac{M_{11}}{\rho_{33}}} & 0 \\ \dfrac{\rho_{31}}{\sqrt{\rho_{33}}} & \dfrac{\rho_{32}}{\sqrt{\rho_{33}}} & \sqrt{\rho_{33}} \end{pmatrix}. \qquad \text{(A.6)}$$

Here, M_{ij} is the first-order minor of matrix ρ_{ph}, i.e., the determinant of the matrix obtained as a result of crossing out the ith row and the jth column of matrix ρ_{ph}.

Likelihood Function

In a system with physical density matrix (A.5), the expected value of the number of coincidences being recorded is equal to

$$\bar{R}_\nu^{ph} = N\langle[(b'_s)^\dagger (b'_i)^\dagger b'_s b'_i]_\nu\rangle,$$

where N is the normalization factor. We assume that the spread in the number of coincidences has a Gaussian statistics. Then the probability that the given set of values $R_1, \ldots, R_9$ will be obtained is given by

$$P(R_1, \ldots, R_9) = \frac{1}{N_{\text{norm}}}\prod_{\nu=1}^{9}\exp\left[-\frac{(R_\nu - \bar{R}_\nu^{ph})^2}{2\sigma_\nu^2}\right], \quad (\text{A.7})$$

where $\sigma_\nu \sim \sqrt{\bar{R}_\nu^{ph}}$ is the standard deviation and N_{norm} is the normalization constant. From the viewpoint of mathematical procedure, the problem of determining the maximum of functional (A.7) is equivalent to determining the minimum of the exponent. This function is called the likelihood function of the MLE method. This function can be written explicitly in the form

$$\Im(q_1, \ldots, q_9; R_1, \ldots, R_9) = \sum_{\nu=1}^{9}\frac{(\bar{R}_\nu^{ph} - R_\nu)^2}{2\bar{R}_\nu^{ph}}. \quad (\text{A.8})$$

Numerical Optimization

In order to find the set of parameters $\{q_1^{\text{opt}}, \ldots, q_9^{\text{opt}}\}$ for which the probabilistic function (A.8) with the preset experimental values $\{R_1, \ldots, R_9\}$ assumes its minimal value, we used the standard procedure with the Mathcad 2000-MINIMIZE package. It was found that the state with density matrix $\rho_{ph}\{q_1^{\text{opt}}, \ldots, q_9^{\text{opt}}\}$ describes the experimental results with the highest probability. The required initial estimate of parameters $\{q_1, \ldots, q_9\}$ was obtained using relation (A.6) with a density matrix reconstructed from experiment. It should be noted that, generally, matrix $\rho_{ph}\{q_1^{\text{opt}}, \ldots, q_9^{\text{opt}}\}$ does not necessarily correspond to a class of pure states; it must only satisfy the physical requirements imposed on the density matrix. In order to find out which pure state corresponds to the experimentally reconstructed matrix with the highest probability, we must require that $\rho_{ph}\{q_1, \ldots, q_9\}$ satisfies the condition $\text{Tr}(\rho_{ph}^2(q_1, \ldots, q_9)) = 1$. This reduces the number of independent variables to four and simplifies the numerical procedure significantly.

The MLE procedure considered above was carried out on the experimentally reconstructed density matrix ρ_{exp} in the form of (30). As a result, the optimized density matrix

$$\rho_{\text{opt}} = \begin{pmatrix} 0.33 & 0.329 - 0.207i & -0.263 - 0.01i \\ 0.329 + 0.207i & 0.459 & -0.257 - 0.176i \\ -0.263 + 0.01i & -0.257 + 0.176i & 0.211 \end{pmatrix}$$

was obtained; the trace of this matrix is $\text{Tr}(\rho_{\text{opt}}^2) = 0.9994$ and its eigenvalues are $\lambda_1 = 0.99969$, $\lambda_2 = 00031$, and $\lambda_3 = 0$.

REFERENCES

1. *The Physics of Quantum Information: Quantum Cryptography, Quantum Teleportation, Quantum Computation*, Ed. by D. Bouwmeester, A. K. Ekert, and A. Zeilinger (Springer, Berlin, 2000; Postmarket, Moscow, 2002).
2. P. G. Kwiat, K. Mattle, H. Weinfurter, *et al.*, Phys. Rev. Lett. **75**, 4337 (1995).
3. P. G. Kwiat, E. Waks, A. White, *et al.*, Phys. Rev. A **60**, R773 (1999).
4. Y. H. Kim, M. V. Chekhova, S. P. Kulik, *et al.*, Phys. Rev. A **63**, 062301 (2001).
5. A. S. Chirkin, A. A. Orlov, and D. Yu. Parashchuk, Kvantovaya Élektron. (Moscow) **20**, 999 (1993).
6. N. Korolkova, G. Leuchs, R. Loudon, *et al.*, Phys. Rev. A **65**, 052306 (2002).
7. V. P. Karasev and A. V. Masalov, Opt. Spektrosk. **74**, 928 (1993) [Opt. Spectrosc. **74**, 551 (1993)].
8. A. V. Burlakov, G. O. Rytikov, S. P. Kulik, and M. V. Chekhova, Zh. Éksp. Teor. Fiz. **122**, 738 (2002) [JETP **95**, 639 (2002)].
9. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, Rev. Mod. Phys. **74**, 145 (2002).
10. H. Bechmann-Pascuinucci and A. Peres, Phys. Rev. Lett. **85**, 3313 (2000).
11. D. Kazlikowski, D. K. L. Oi, M. Christandl, *et al.*, Phys. Rev. A **67**, 012310 (2003).
12. T. Durt, N. J. Cerf, N. Gisin, and M. Zukovski, Phys. Rev. A **67**, 012311 (2003).
13. J. Ben-Aryeh, quant-ph/0301047.
14. B. C. Sanders, H. De Guise, S. D. Barlett, and W. Zhang, Phys. Rev. Lett. **86**, 369 (2001).
15. D. T. Smithey, M. Beck, and M. G. Raymer, Phys. Rev. Lett. **70**, 1244 (1993).
16. G. M. D'Ariano and M. G. A. Paris, J. Opt. B: Quantum Semiclassical Opt. **2**, 113 (2000).

17. P. A. Bushuev, V. P. Karasev, A. V. Masalov, and A. A. Putilin, Opt. Spektrosk. **91**, 558 (2001) [Opt. Spectrosc. **91**, 526 (2001)].
18. D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, Phys. Rev. A **64**, 052312 (2001).
19. A. V. Burlakov, L. A. Krivitskiĭ, S. P. Kulik, *et al.*, Opt. Spektrosk. **94**, 744 (2003) [Opt. Spectrosc. **94**, 684 (2003)].
20. A. V. Burlakov and M. V. Chekhova, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 505 (2002) [JETP Lett. **75**, 432 (2002)].
21. D. N. Klyshko, Zh. Éksp. Teor. Fiz. **111**, 1955 (1997) [JETP **84**, 1065 (1997)].
22. R. H. Brown and R. Q. Twiss, Nature **177**, 27 (1956).
23. A. V. Burlakov and D. N. Klyshko, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 795 (1999) [JETP Lett. **69**, 839 (1999)].
24. K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, Phys. Rev. A **61**, 010304 (2000).
25. A. V. Belinsky, G. Kh. Kitaeva, S. P. Kulik, and A. N. Penin, Phys. Rev. B **51**, 3362 (1995).
26. S. P. Kulik, G. Kh. Kitaeva, and A. N. Penin, Ferroelectrics **172**, 469 (1995).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 4, 2003, p. 858.

ERRATA

Erratum: “Violation of the Factorization Theorem in Large-Angle Radiative Bhabha Scattering” [*Zh. Éksp. Teor. Fiz.* 115, 392 (1999); *JETP* 88, 213 (1999)]

A. B. Arbuzov, E. A. Kuraev, and B. G. Shaikhatdenov

Factor –1/2 was lost in the contribution of box diagrams. Quantity Δ_B in Eq. (22) should read

$$\Delta_B = \frac{1}{F}\left[(\Phi + \Phi_P)\ln\frac{uu_1}{ss_1} + (\Phi_Q + \Phi_R)\ln\frac{uu_1}{tt_1}\right].$$

Operations $\hat{P}$, $\hat{Q}$, and $\hat{R}$ were applied to vertex contributions incorrectly. Quantity $\Delta_G + \Delta_{\Gamma\Pi}$, which appears in Eq. (23), should be defined as

$$\Delta_G + \Delta_{\Gamma\Pi} = \frac{1}{F}\left[(\Phi + \Phi_P)\ln\frac{s^2}{tt_1} + (\Phi_Q + \Phi_R)\ln\frac{s}{s_1}\right].$$

To analyze the corrected result of our calculations, we rewrite the soft photon contribution from Eq. (29) in the form

$$\delta^{\text{soft}} = \frac{\alpha}{\pi}\left\{(L_s - 1)\left[4\ln\frac{m}{\lambda} + 2\ln\frac{\Delta\varepsilon}{\varepsilon} + \ln\frac{\Delta\varepsilon}{\varepsilon_1'} + \ln\frac{\Delta\varepsilon}{\varepsilon_2'}\right]\right.$$
$$\left. + L_s^2 + L_s\ln\frac{tt_1s_1}{uu_1s} + \mathcal{O}(1)\right\}.$$

Expression (31) should be therefore rewritten as

$$d\delta^{\text{soft+virt}}$$
$$= d\sigma_0\frac{\alpha}{\pi}\left[L_s\left(\ln\frac{(\Delta\varepsilon)^4}{\varepsilon_2\varepsilon_1'\varepsilon_2'} + \Delta_L\right) + \Delta(y_1, y_2, c_1, c_2)\right],$$

$$\Delta_L = 3.$$

This formula is in agreement with the predictions of the renormalization group approach presented in Eq. (35). Thus, the factorization of the leading logarithmic contributions in radiative Bhabha scattering does hold. It is worth mentioning that the contribution of vacuum polarization to radiative Bhabha scattering, which was omitted in Eqs. (31), can be taken from paper [8].

The above changes in the formulas lead to a change in Table 1: $\Delta_L = 3$ remains constant, and the values of $\Delta(y_1, y_2, c_1, c_2)$ are –8.89, 2.00, –1.47, and 7.80 for points 1, 2, 3, and 4, respectively.

1063-7761/03/9704-0858$24.00 © 2003 MAIK “Nauka/Interperiodica”