Continuity of magnetohydrodynamic flows near singular points

V. S. Beskin and I. V. Kuznetsova

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia;
Moscow Physicotechnical Institute, 141700 Dolgoprudnyi, Moscow Region, Russia
(Submitted 14 May 1997)
Zh. Éksp. Teor. Fiz. **113**, 771–785 (March 1998)

The continuity of hydrodynamic and magnetohydrodynamic axisymmetric stationary flows near singular points of acoustic surfaces is investigated when the flows are far from spherical symmetry. It is shown that the flow is still continuous at the time the characteristics undergo a bifurcation, but may become discontinuous when the flow is distorted further. © *1998 American Institute of Physics.* [S1063-7761(98)00103-6]

1. INTRODUCTION

The problem of flow continuity near singular surfaces is one of the classical problems in theoretical physics.[1–5] However, even in the simplest case of ideal hydrodynamics, the problem reduces to a nonlinear second-order partial differential equation which changes from elliptical to hyperbolic at an acoustic surface. Recall that equations of this form, which go back to the classical equation of Tricomi, have been discussed since the beginning of this century in connection with transonic hydrodynamic flows (nozzles, wings). In particular, for plane flows the hodograph transformation method has been extremely fruitful (yielding a linear Chaplygin equation) and has led to important advances in understanding these processes.[1–3] Despite this, the question of constructing a rigorous analytic theory for these flows is, in fact, still open.

The difficulty is, first of all, that the very statement of the direct problem in terms of an equilibrium equation is nontrivial. In other words, the problem of constructing a solution for given boundary conditions is nontrivial. Thus, even for one of the simplest variants, planar hydrodynamic flow, there is no rigorous way of constructing a solution to the direct problem (finding a solution from the shape of a nozzle or wing). As a result, in most cases it has been necessary to be satisfied with a solution of the inverse problem (finding the shape of a nozzle from a known flow).[1–3]

Yet another difficulty with this approach lies in the fact that, for example, in the hydrodynamic limit the equilibrium equation contains three integrals (constants) of motion, which must be determined from the boundary conditions. This means that to determine the structure of a transonic flow on some surface, one must specify four boundary conditions, e.g., two thermodynamic functions, along with two components of the velocity. However, in order to determine the Bernoulli integral, without which the equilibrium equation cannot, of course, be solved, all three components of the velocity must be known, which is impossible since the third component of the velocity must be obtained from the solution. This inconsistency is one of the fundamental complications in the method being described here. As a result, the construction of a solution to the direct problem encounters substantial difficulties. On the other hand, it has been shown that a flow which passes smoothly through the acoustic and separatrix surfaces is analytic near singular points—the points of tangency of the acoustic and separatrix surfaces.[3] This important result makes it possible to analyze the properties of a flow when it differs only slightly from a known exact solution to the problem by expanding the solution in terms of small deviations from the singular point immediately in the physical plane.

As for astrophysical applications, exact solutions have also been constructed only for a number of model problems. In most cases, these studies have been carried out with the aid of various self-similar formulations[6–10] or numerically.[11–17] Thus, Bogovalov[18] has recently shown with a numerical simulation of a cold magnetized wind from a rotating star with a quasimonopole magnetic field that, when the angular velocity of rotation of the star is increased, a discontinuity shows up in the region of a singular point behind the fast magnetoacoustic surface and does not vanish as the angular velocity is increased further. In that paper it was also suggested that this kind of discontinuity arises as soon as the characteristics undergo a bifurcation, when the structure of the characteristic surfaces near the singular point changes suddenly as the flow parameters vary discontinuously. The purpose of this paper is actually to verify this assumption.

In the second section of our paper we formulate the basic equations for axisymmetric stationary flows. The third section is devoted to a discussion of the conditions under which the characteristics undergo a bifurcation. In the fourth section, a flow near a singular point is examined in the hydrodynamic limit. It is shown that the flow remains continuous at the time the characteristics undergo a bifurcation, but in principle can become discontinuous when the flow is distorted further. For this to happen, however, the shape of the acoustic surface must already differ substantially from spherical. In the fifth section, a program of this sort is carried out for the magnetohydrodynamic problem examined by Bogovalov. Here also there is no singularity when the bifurcation occurs. On the other hand, it is shown that the instability found above could be the cause of the discontinuity found by Bogovalov.

2. BASIC EQUATIONS

First of all, we state the basic equations that we shall use to study axisymmetric stationary flows. For hydrodynamic flows, one can introduce the stream function $\Phi(r,\theta)$, which is defined by the equation

$$n\mathbf{v}_p=\frac{1}{2\pi\varpi}\,\nabla\times\Phi\mathbf{e}_\varphi. \tag{1}$$

Here $\varpi=r\,\sin\,\theta$ is the distance from the axis of rotation, while n and $\mathbf{v}_p$ are the density and poloidal velocity of the material. The curves $\Phi(r,\theta)=\text{const}$ just define the stream lines and, thanks to the condition $d\Phi=n\mathbf{v}\cdot d\mathbf{S}$, the quantity $\Phi(r,\theta)$ coincides with the flux of material through the surface bounded by the circle $r=\text{const}$, $\theta=\text{const}$, $0<\varphi<2\pi$. In particular, the total flux through a sphere of radius r is $\Phi(r,\pi)$. Here the continuity equation $\nabla\cdot(n\mathbf{v})=0$ is satisfied automatically because of the definition (1).

In the hydrodynamic limit there are only three integrals of motion, which are constant on the surfaces $\Phi=\text{const}$. These are the energy flux (Bernoulli integral)

$$E(\Phi)=\frac{v^2}{2}+w+\varphi_G(r,\theta) \tag{2}$$

(w is the specific enthalpy and φ_G is the gravitational potential), the component $L(\Phi)$ of the angular momentum given by

$$L(\Phi)=v_\varphi r\,\sin\,\theta \tag{3}$$

and the entropy $s(\Phi)$, which is also constant on the streamlines in ideal hydrodynamics. For simplicity, in the following we shall consider only hydrodyamical flows with zero angular momentum, $L=0$, and constant entropy s and energy E throughout all space. This situation arises during the accretion of a gas to a moving gravitational center.[19] In this case the equilibrium equation of the streamlines can be rewritten in the form[20]

$$-\varpi^2\nabla_k\left(\frac{1}{\varpi^2}\,\nabla^k\Phi\right)-\frac{(\nabla^i\Phi)(\nabla^k\Phi)(\nabla_i\nabla_k\Phi)}{D(\nabla\Phi)^2}+\frac{(\nabla\varpi^2)\cdot(\nabla\Phi)}{2D\varpi^2}-4\pi^2\varpi^2n^2\,\frac{(\nabla\varphi_G)\cdot(\nabla\Phi)}{D(\nabla\Phi)^2}=0, \tag{4}$$

where

$$D=-1+\frac{c_s^2}{v_p^2}, \tag{5}$$

c_s is the speed of sound, and all covariant differentiation operators ∇_k act in a flat space. Equation (4) contains only one free surface, the acoustic surface, which is defined by the condition $D=0$.

Equation (4) is a generalization to the axisymmetric case of the well known equation[3,5]

$$\varphi_{xx}+\varphi_{yy}+\frac{\varphi_y^2\varphi_{xx}+2\varphi_x\varphi_y\varphi_{xy}+\varphi_x^2\varphi_{yy}}{D(\nabla\varphi)^2}=0, \tag{6}$$

which describes planar isentropic flow for a potential $\varphi(x,y)$ that specifies the velocity of the material in accordance with the definition $\mathbf{v}=\nabla\varphi$. Like Eq. (6), Eq. (4) is actually the Euler equation, now rewritten for the stream function $\Phi(r,\theta)$. We emphasize that the equilibrium equation (4), like Eq. (6), only contains the stream function Φ and the two integrals of motion E and s. In fact, with the aid of Eq. (1) (and for $L=0$), Eq. (2) can be rewritten in the form

$$2E-2\varphi_G(r,\theta)-2w=\frac{(\nabla\Phi)^2}{4\pi^2r^2n^2\,\sin^2\,\theta}. \tag{7}$$

Since the enthalpy w can be expressed as a function of n and s, Eq. (7) also determines, although implicitly, the density n and, together with it, all the remaining thermodynamic quantities, as functions of the potential Φ and the two integrals of motion. Equation (7) can also be rewritten in the form

$$\nabla_k n=n\,\frac{N_k}{D}, \tag{8}$$

where

$$N_k=-\frac{(\nabla^i\Phi)(\nabla_i\nabla_k\Phi)}{(\nabla\Phi)^2}+\frac{1}{2}\,\frac{\nabla_k\varpi^2}{\varpi^2}-4\pi^2\varpi^2n^2\,\frac{\nabla_k\varphi_G}{(\nabla\Phi)^2}, \tag{9}$$

while D is given by Eq. (5).

Similarly, in the framework of magnetohydrodynamics we can introduce the magnetic stream function $\Psi(r,\theta)$ through

$$\mathbf{B}_p=\frac{1}{2\pi\varpi}\,\nabla\Psi\times\mathbf{e}_\varphi, \tag{10}$$

where $\mathbf{B}_p$ is the longitudinal magnetic field. Accordingly, the velocity of the particles can be written

$$\mathbf{v}=\frac{\eta}{n}\,\mathbf{B}+\Omega_F\varpi\mathbf{e}_\varphi. \tag{11}$$

In this case, for a cold plasma we have $s=0$ (and neglecting gravitation) the equilibrium equation for the magnetic surfaces can be written in the form[20] (see Refs. 9, 10, and 13, as well)

$$A\left[\nabla_k\left(\frac{1}{\varpi^2}\,\nabla^k\Psi\right)+\frac{1}{\varpi^2(\nabla\Psi)^2}\,\frac{(\nabla^a\Psi)(\nabla^b\Psi)(\nabla_a\nabla_b\Psi)}{D}\right]-\frac{A}{\varpi^2(\nabla\Psi)^2}\,\frac{1}{2D}\,(\nabla_k'F)(\nabla^k\Psi)+\frac{1}{\varpi^2}\,(\nabla_k'A)\times(\nabla^k\Psi)+\frac{\Omega_F}{c^2}\,\frac{d\Omega_F}{d\Psi}\,(\nabla\Psi)^2+\frac{64\pi^4}{\tilde{\omega}^2}\,\frac{1}{2M^2}\,\frac{\partial}{\partial\Psi}\left(\frac{G}{A}\right)-16\pi^3\mu n\,\frac{1}{\eta}\,\frac{d\eta}{d\Psi}=0. \tag{12}$$

Here μ is the relativistic enthalpy, which coincides with the mass of the particles in a cold plasma; $M^2=4\pi\eta^2\mu/n$; the gradient ∇_k' acts on all quantities except M^2; and the derivative $\partial/\partial\Psi$ acts only on the integrals of motion. In addition,

$$D=\frac{A}{M^2}+\frac{1}{M^2}\,\frac{B_\varphi^2}{B_p^2}. \tag{13}$$

while the Alfvén factor A is

$$A=1-\Omega_F^2\varpi^2/c^2-M^2. \tag{14}$$

Finally, we have

$$G=\varpi^2(E-\Omega_F L/c)^2+M^2L^2-M^2\varpi^2E^2 \tag{15}$$

and

$$F=\frac{64\pi^4}{M^4}\frac{K}{A^2}-\frac{64\pi^4}{M^4}\varpi^2\eta^2\mu^2 \tag{16}$$

and all the operators ∇ again act in a flat space. Here the toroidal magnetic field B_φ and electric field $\mathbf{E}$ are defined as

$$B_\varphi=-\frac{2I}{c\varpi}; \tag{17}$$

and

$$\mathbf{E}=-\frac{\Omega_F}{2\pi c}\nabla\Psi. \tag{18}$$

Here $I(r,\theta)$ is the total electric current inside the region $\Psi<\Psi(r,\theta)$.

As in the hydrodynamic limit, the equilibrium equation (12) contains only the function $\Psi(r,\theta)$ and four integrals of motion, specifically, the energy flux $E(\Psi)$ and angular momentum $L(\Psi)$, given by

$$E=E(\Psi)=\frac{\Omega_F I}{2\pi c^4}+\mu\eta+\mu\eta\frac{v^2}{2c^2} \tag{19}$$

and

$$L=L(\Psi)=\frac{I}{2\pi c^3}+\mu\eta\varpi\frac{u_\varphi}{c}, \tag{20}$$

as well as the rotational angular velocity $\Omega_F(\Psi)$ of the magnetic surfaces and the ratio of the particle flux to the magnetic flux, $\eta(\Psi)$. Here the electric current is given in terms of the integrals of motion by[20]

$$I=2\pi c^3\frac{L-\Omega_F\varpi^2E/c}{1-\Omega_F^2\varpi^2/c^2-M^2}. \tag{21}$$

On the other hand, the equilibrium equation (12) contains two singular surfaces, the Alfvén surface on which $A=0$, and the fast magnetoacoustic surface on which $D=0$. Thus, Eq. (12) requires that four boundary conditions be specified on the surface of a rotating body.

Finally, the coupling condition equivalent to the Bernoulli equation of hydrodynamics, which just makes it possible to express M^2 in terms of the integrals of motion, is rewritten in the form

$$\frac{\varpi^2(E-\Omega_F L/c)^2(1-\Omega_F^2\varpi^2/c^2-2M^2)+M^4(\varpi^2E^2-L^2)}{\varpi^2A^2}=\frac{1}{64\pi^4}\frac{M^4(\nabla\Psi)^2}{\varpi^2}+\eta^2\mu^2, \tag{22}$$

and the derivative $\nabla_k M^2$ is written as

$$\nabla_k M^2=\frac{N_k}{D}, \tag{23}$$

where now the numerator N_k has the form

$$N_k=-\frac{A}{(\nabla\Psi)^2}\nabla^b\Psi\nabla_k\nabla_b\Psi+\frac{A}{2}\frac{\nabla_k' F}{(\nabla\Psi)^2}. \tag{24}$$

As we can see, in both cases the equilibrium equation is a second-order equation which is linear in the higher derivatives. In other words, it can be written in the canonical form

$$\mathscr{A}\frac{\partial^2\Psi}{\partial r^2}+2\mathscr{B}\frac{\partial^2\Psi}{\partial r\partial\theta}+\mathscr{C}\frac{\partial^2\Psi}{\partial\theta^2}+\mathscr{F}=0, \tag{25}$$

where the coefficients $\mathscr{A}-\mathscr{F}$ are independent of the second derivatives. Calculating the standard combination $\mathscr{D}=\mathscr{A}\mathscr{C}-\mathscr{B}^2$, for the MHD flows we obtain

$$\mathscr{D}=A^2D(D+1) \tag{26}$$

and, accordingly, $\mathscr{D}=D(D+1)$ in hydrodynamics. The equilibrium equation, therefore, changes from elliptic to hyperbolic at the singular surfaces on which the poloidal velocity of the material equals either the fast or the slow magnetoacoustic speed, $D=0$, or with the cusp velocity, $D=-1$.[20] In the cold-plasma approximation, the condition $D=0$ is satisfied only on the fast magnetoacoustic surface, while the slow and cusp surfaces do not exist. In the hydrodynamic limit, on the other hand, the only singular surface will be the acoustic surface, $D=0$.

3. BIFURCATION OF THE CHARACTERISTICS

We now consider the bifurcation of the characteristics, which can occur when the flow differs greatly from spherically symmetric. To demonstrate this, we write down the equation

$$\frac{dr}{d\theta}=\frac{\mathscr{B}\pm\sqrt{\mathscr{B}^2-\mathscr{A}\mathscr{C}}}{\mathscr{A}}, \tag{27}$$

which describes the trajectories of the characteristic surfaces in the hyperbolic region of the equilibrium equation (25). Now, using the explicit form of the coefficients $\mathscr{A}-\mathscr{C}$, we have, for example, in the hydrodynamic limit

$$\frac{dr}{d\theta}=\frac{r^2(\partial\Phi/\partial r)(\partial\Phi/\partial\theta)\pm[r^2(\partial\Phi/\partial r)^2+(\partial\Phi/\partial\theta)^2]\sqrt{-D(D+1)r^2}}{(\partial\Phi/\partial\theta)^2+D^2[r^2(\partial\Phi/\partial r)^2+(\partial\Phi/\partial\theta)^2]}. \tag{28}$$

Equation (28) shows that, in general, at the acoustic surface $D=0$, the derivative $d(r-r_s)/d\theta$ ($r_s=r_s(\theta)$ is the radius of the acoustic surface) is nonzero. Thus, two characteristics will pass through every point.[5] However, at singular

points where the derivative satisfies $\partial\Phi/\partial r=0$, determining the behavior of the characteristic surfaces requires a more detailed treatment which can be carried out by expanding the solution near the singular point $r=r_*$. In the end, in terms of the dimensionless units

$$R=\frac{r_s-r}{r_* D_1}, \tag{29}$$

where $D_1=r_*(\partial D/\partial r)$ for $r=r_*$, Eq. (28) can be rewritten in the form

$$\frac{dR}{d\theta}=a\,\theta\pm\sqrt{R}, \tag{30}$$

where

$$a=-\frac{2(\partial D/\partial\theta^2)_{r=r_*,\theta=0}-r_*(D_1/\Phi_*)(\partial^2\Phi/\partial r\,\partial\theta^2)_{r=r_*,\theta=0}}{D_1^2}. \tag{31}$$

The exact solution of Eq. (30) can be found using the substitution

$$R(\theta)=k^2(\theta)\,\theta^2. \tag{32}$$

Substituting Eq. (32) in Eq. (30), we obtain the implicit expression

$$k(\theta)=k_1+C^{(k_1-k_2)/k_1}[k_2-k(\theta)]^{k_2/k_1}\theta^{(k_2-k_1)/k_1}. \tag{33}$$

Here C is an integration constant, while k_1 and k_2 are two fixed values of $k(\theta)$ which are independent of the angle θ. They can be found directly from Eq. (30). As a result, we have

$$2k^2\pm k-a=0, \tag{34}$$

and, therefore,

$$k_{1,2}=\frac{1\pm\sqrt{1+8a}}{4}. \tag{35}$$

The behavior of the characteristic surfaces is exhibited in Figs. 1 and 2. Here the condition $a>0$ corresponds to the classical case[1–5] in which four characteristics, which correspond to the two branches of the roots of Eq. (35), pass through a singular point. In fact, for $k\approx k_1$, Eq. (33) yields

$$k(\theta)\approx k_1+C_1\theta^m, \tag{36}$$

where the exponent is

$$m=-\frac{1+8a+\sqrt{1+8a}}{4a}. \tag{37}$$

Thus, for $a>0$, when $m<0$ holds the second term in Eq. (36) diverges as $\theta\to 0$. Therefore, only the characteristic corresponding to $C_1=0$ will pass through the origin when $k\approx k_1$. On the other hand, in the case $a<0$ (which cannot be realized for a plane hydrodynamic flow in the absence of a gravitational field), the situation is substantially different. Thus, when the condition $-1/8<a<0$ corresponding to the real roots (35) is satisfied, infinitely many characteristics pass through the singular point, since now we have $m>0$ and $k(\theta)\to k_1$ as $\theta\to 0$ for all values of the constant C_1. On the other hand, for $a<-1/8$, when the roots (35) become complex, the structure of the characteristic surfaces is sharply different, so that not even one characteristic passes through the singular point. The spherically symmetric case, as such, corresponds to the condition $a=0$.

Therefore, when the flow differs enough from spherically symmetric, the entire structure of the characteristic surfaces, including the separatrix characteristic (which, as is known, just separates two regions that are not causally coupled[1–3]), changes in a discontinous fashion. In particular, if for $a>-1/8$ the separatrix characteristic joins two singular points, then for $a<-1/8$ it no longer does. As a result, the region lying immediately above a singular point with $a<0$ also begins to affect the elliptic region of the equilibrium equation in a discontinous fashion, since, as can be seen from Fig. 2, a perturbation from this region now reaches the

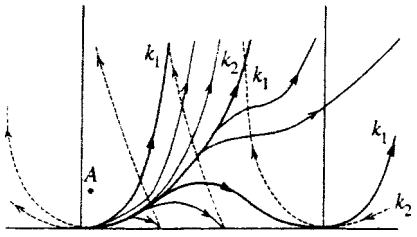

FIG. 1. The behavior of the characteristic surfaces in the case where the parameter a for a nonstationary singular point (left) satisfies the condition $-1/8<a<0$. The heavy curve denotes the separatrix characteristic, as well as the solution corresponding to $k=k_1$ and $k=k_2$. A classical singular point $a>0$ is indicated on the right. The point A has no effect on the subsonic region.

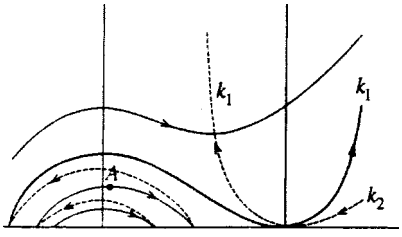

FIG. 2. The same, for $a<-1/8$. A perturbation from point A reaches the acoustic surface along the characteristic.

acoustic surface along the characteristic. Now it can be understood why Bogovalov proposed that a discontinuity as such in an outflow from a rotating magnetized body originates when the characteristics undergo a bifucation.[18]

4. HYDRODYNAMIC FLOW

We now proceed to a rigorous examination of the behavior of the solution near the singular points of the equilibrium equation (25). First of all, we consider accretion in the hydrodynamic limit with a nonrelativistic velocity in the neighborhood of the point $\theta=0$, $r=r_*$, where r_* is the radius of the acoustic surface at $\theta=0$. This situation can be realized during accretion at a moving black hole.[20] Following Frankl',[3] we write down the first terms of the expansion of the stream function Φ, density n, and velocity v in powers of $r-r_*$ and θ in the form

$$\Phi=\Phi_*\left[\frac{1}{2}\,\theta^2+\frac{1}{2}\,bh\theta^2-\frac{q}{24}\,\theta^4+\frac{p_1}{720}\,\theta^6+\frac{p_2}{24}\,\theta^4 h+\frac{p_3}{2}\,h^2\theta^2\right], \tag{38}$$

$$n=n_*(1+\eta_1 h+\eta_2\theta^2), \tag{39}$$

and

$$v=v_*(1+w_1 h+w_2\theta^2), \tag{40}$$

where $h=(r-r_*)/r_*$. It is clear that with $b=0$, $q=1$, $p_1=1$, $p_2=0$, and $p_3=0$, for $\eta_2=0$ and $w_2=0$, the function Φ (38) corresponds to the spherically symmetric flow

$$\Phi=\Phi_0(1-\cos\,\theta). \tag{41}$$

On the other hand, it is clear that the expansion (38) does not require any limitations on the values of the coefficients b, q, and p_1-p_3. Recall that at a classical singular point, the entire structure of a transonic flow (for given constant values of the integrals E and s) was determined with the aid of just one more function. In the direct formulation, this function might be the shape of the flow boundary, e.g., a nozzle or wing. However, as mentioned above, serious difficulties arise in constructing a solution. On the other hand, because the transonic flow is analytic,[3] the solution near the acoustic surface can be constructed by expanding in a series when one further function characterizing the flow is specified. Thus, for a plane symmetric flow, this function might be the longitudinal velocity on the symmetry axis, $f(x)=v_x(x,0)$.[3] In our case, it is convenient to take the flux $nv_r(r,0)$ on the axis, $\theta=0$, as the defining function. According to Eqs. (1) and (38), this means that the role of ''boundary conditions'' will be played by b and p_3, which just determine the deviation of the flow from spherical symmetry. Thus, for accretion to a black hole moving at velocity v_∞,[20] we obtain

$$b=2k_1(\Gamma)\varepsilon_1, \tag{42}$$

and

$$p_3=\frac{2b}{\sqrt{10-6\Gamma}}\left(\frac{4-\sqrt{10-6\Gamma}}{\Gamma+1}-1\right), \tag{43}$$

where the small parameter is $\varepsilon_1=v_\infty/c_s(\infty)$ and the numerical coefficient $k_1(\Gamma)\sim 1$ has been determined elsewhere.[20] All the remaining coefficients of the expansions in Eqs. (38)–(41) should ultimately be expressed in terms of these quantities. In fact, our problem will just consist in verifying the continuity of this procedure for $a=-1/8$.

As a result, using Eqs. (4) and (9) for N_r, we obtain the following relations for the coefficients in the expansion of Φ and n:

$$\eta_1=\frac{2b-4+\sqrt{(4-2b)^2-(\Gamma+1)(6-6b+2p_3+b^2)}}{\Gamma+1}, \tag{44}$$

$$q=1+\frac{3}{4}\,b^2, \tag{45}$$

$$w_1=b-2-\eta_1, \tag{46}$$

and

$$w_2=-\eta_2, \tag{47}$$

where the + sign in front of the square root in Eq. (44) corresponds to accretion. Here

$$b=2-\frac{GM}{r_* v_*^2}, \tag{48}$$

where $v_*=\Phi_*/2\pi n_* r_*^2$ is the flow velocity at the point $\theta=0$, $h=0$, so that for a spherically symmetric flow,[22] when $r_*=GM/2v_*^2$, we have $b=0$. For simplicity, here we have used the polytropic equation of state $P=kn^\Gamma$, $1<\Gamma<5/3$, for which the specific enthalpy is $w=c_s^2/(\Gamma-1)$, where $c_s^2=(1/m_p)k\Gamma n^{\Gamma-1}$ is the square of the speed of sound. We emphasize that for $b=0$ and $p_3=0$, Eq. (44) for η_1 transforms to the well known expression[22,23]

$$\eta_1=\frac{-4+\sqrt{10-6\Gamma}}{\Gamma+1} \tag{49}$$

for a spherically symmetric flow.

Furthermore, the first terms in the expansion in powers of h and θ for the θ-component of Eq. (8) give

$$\eta_2=-\frac{\left(\frac{1}{60}\,p_1+\frac{1}{36}\,q^2+\frac{1}{15}-\frac{1}{9}\,q+\frac{1}{12}\,b^2+\frac{1}{24}\,bp_2\right)\theta^2+\frac{1}{2}\left(\frac{1}{2}\,b^2+\frac{2}{3}\,b+\frac{1}{3}\,p_2+bp_3-\frac{1}{3}\,qb\right)h}{(\Gamma+1)\eta_2\theta^2+[(\Gamma+1)\eta_1+(4-2b)]h}. \tag{50}$$

It is clear from Eq. (50) that the following expressions hold for p_1 and p_2:

$$p_1=-60(\Gamma+1)\eta_2^2-\frac{5}{3}q^2-4+\frac{20}{3}q-5b^2-\frac{5}{2}bp_2 \tag{51}$$

and

$$p_2=-6[(\Gamma+1)\eta_1+4-2b]\eta_2-2(q-1+b)+qb-3bp_3. \tag{52}$$

In addition, in this case of a potential flow, we have

$$\nabla\times v=0. \tag{53}$$

As a result of Eq. (53), we obtain an expression for the coefficient η_2,

$$\eta_2=\frac{1-q}{6}-\frac{1}{4}\eta_1 b+\frac{p_3}{2}. \tag{54}$$

Equations (44)–(54) also make it possible to determine all the rest of the coefficients in the expansion for the stream function Φ and density n in terms of b and p_3.

Finally, the coefficient a in Eq. (30) is expressed in terms of b and p_3 as follows:

$$a=-\frac{2b+p_3(\Gamma+1)}{D_1^2}, \tag{55}$$

where

$$D_1^2=(4-2b)^2-(\Gamma+1)(6-6b+2p_3+b^2) \tag{56}$$

is the expression under the square root sign in Eq. (44). It is clear from condition (55) that at the point at which the characteristics undergo a bifurcation, $a=-1/8$, there are no changes in the procedure for determining the coefficients in the expansion of the stream function Φ and density n in terms of b and p_3. In other words, the flow remains continuous. However, as the flow becomes more distorted, when $D_1^2\to 0$ and, therefore, $a\to-\infty$, the possibility of representing the solution of the equation in the form of the analytic series (38) is lost, because η_1 and, therefore, all the other expansion coefficients, become complex quantities. This also means that a continuous flow regime becomes impossible.[5]

In a similar fashion, we can obtain relationships for the coefficients in the expansions of the stream function and density near the equator. In this case, which is realized, for example, during accretion at a rotating black hole,[20,21] we have

$$\Phi=\Phi_*\left[1+\beta-\frac{q}{6}+bh\beta+\frac{p_1}{120}\beta^5+\frac{p_2}{6}\beta^3h+p_3\beta h^2\right], \tag{57}$$

and

$$n=n_*(1+\eta_1 h+\eta_2\beta^2), \tag{58}$$

where $\beta=\theta-\pi/2$.

The equations for the expansion coefficients now have the form

$$q=1+b^2, \tag{59}$$

$$\eta_1=\frac{2b-4+\sqrt{(4-2b)^2-(\Gamma+1)(6-6b+2p_3+b^2)}}{\Gamma+1}, \tag{60}$$

$$p_1=-12(\Gamma+1)\eta_2^2-8-12b^2+12q-3q^2-4bp_2, \tag{61}$$

$$p_2=-b-2b^2-4p_3b+b^3 -2(\Gamma+1)\eta_1\eta_2-4(2-b)\eta_2, \tag{62}$$

and

$$\eta_2=\frac{1}{2}(2p_3-b^2-\eta_1 b). \tag{63}$$

In this case the situation is clearly analogous to that examined above.

We conclude, therefore, that when the characteristics undergo a bifurcation, the procedure for finding the solution near a singular point is still valid. In other words, to determine the solution we have, as before, to determine one other free function besides the integrals of motion E and s. On the other hand, the continuity may break down when the flow is distorted further, i.e., for sufficiently large values of b and p_3, when the expression D_1^2 under the square root in Eq. (44) goes negative. It is true, as our analysis has shown, that for the values given by Eqs. (42) and (43) this becomes possible only for $\varepsilon_1>1$, when a shock inevitably develops in the entering flow. Recall that for a spherically symmetric flow $D_1^2=10-6\Gamma$, so that $D_1^2>0$ for $\Gamma<5/3$.

5. MAGNETOHYDRODYNAMIC FLOW

We now consider the nonrelativistic MHD flow of a cold plasma (and neglecting gravitation) in the neighborhood of the singular point $\theta=\pi/2$, $h=0$. As noted above, in order to determine the transonic flow we have to specify four functions on the surface of the body, for example, the density and velocity of the material, as well as the rotational angular velocity and the normal component of the magnetic field. For simplicity, as in the case of the flow considered by Bogovalov,[18] we shall assume that all these quantities are constant on the surface:

$$n(R,\theta)=n_{\rm in}=\text{const}, \tag{64}$$

$$v(R,\theta)=v_{\rm in}=\text{const}, \tag{65}$$

$$\Omega_F(R,\theta)=\Omega=\text{const}, \tag{66}$$

and

$$B_n(R,\theta)=B_n=\text{const}. \tag{67}$$

We now write the first terms in the expansions of the flux Ψ and density n in powers of h and β in the form

$$\Psi=\Psi_*\left[1+\beta+\alpha_1 h\beta-\frac{\alpha_2}{6}\beta^3 +\frac{\alpha_3}{120}\beta^5+\frac{\alpha_4}{6}\beta^3h+\alpha_5h^2\beta\right] \tag{68}$$

and

$$n=n_*(1+\eta_1 h+\eta_2\beta^2). \tag{69}$$

For $\alpha_1=0$, $\alpha_2=1$, $\alpha_3=1$, $\alpha_4=0$, and $\alpha_5=0$, the flux Ψ goes over to the spherically symmetric solution of Eq. (12) which occurs for an angular velocity $\Omega_F=0$ and angular momentum $L=0$:

$$\Psi=\Psi_0(1-\cos\,\theta). \tag{70}$$

In this case the Alfvén surface coincides with the fast magnetoacoustic surface. For nonrelativistic flow velocities, the coupling condition (22) has the form

$$E-\mu\eta=\frac{M^2(\nabla\Psi)^2}{128\pi^4\varpi^2\mu\eta}+\frac{K}{A^2}, \tag{71}$$

where now

$$A=1-M^2, \tag{72}$$

and

$$K=(2\Omega_F L\mu\eta\varpi^2-\Omega_F^2\varpi^4\mu^2\eta^2)(1-2M^2)+M^4L^2. \tag{73}$$

As in Bogovalov's paper, we shall consider an outflow of cold plasma with the constant integrals of motion Ω_F, $E=\mu\eta(1+v_{in}^2/2c^2)$,

$$\eta=\frac{2\pi n_{\rm in}v_{\rm in}R^2}{\Psi_0} \tag{74}$$

and angular momentum

$$L(\Psi)=L_0\left[1-\frac{(\Psi-\Psi_*)^2}{\Psi_*^2}\right]. \tag{75}$$

As a result, from Eqs. (12) and (71), it is possible to obtain the following relationships:

$$\alpha_1=1-2(\omega^2+1)\xi_*+\xi_*^2\,\frac{X_A^2}{X_*^2}\,\frac{M_*^2-1}{\alpha^2}, \tag{76}$$

$$\alpha_2=(2\xi_*^2\omega^2+2\xi_*)\,\frac{M_*^2}{X_*^2}+\alpha_1^2+\alpha_1-1, \tag{77}$$

$$\eta_1=-\frac{4}{3}\,\xi_*-\frac{2}{3}\,\omega^2(1-\alpha_1)-D_1, \tag{78}$$

and

$$D_1^2=\left[\frac{4}{3}\,\xi_*+\frac{2}{3}\,\omega^2(1-\alpha_1)\right]^2-\frac{4(2M_*^2-1)}{3M_*^2}\xi_*^2$$
$$-\omega^2(1-\alpha_1)+\frac{\omega^2}{3}\,(1-\alpha_1)^2-\frac{2}{3}\,\omega^2\alpha_5, \tag{79}$$

where we have introduced the dimensionless quantities

$$\alpha=\frac{\Omega_F r_A}{v_A},\quad \omega^2=\frac{M_*^2-1}{M_*^2},\quad \xi_*=\frac{\Omega_F r_*^2\mu\eta}{I_*M_*^2}, \tag{80}$$

and

$$X_*=\frac{r_*}{r_A},\quad X_A=\frac{r_a}{r_A}. \tag{81}$$

Here $I_*=I(r_*,\pi/2)/2\pi$, $I(r,\theta)$ is the total current flow in the region $\Psi<\Psi(r,\theta)$, with

$$\frac{I}{2\pi}=\frac{L-\Omega_F\varpi^2\mu\eta}{1-M^2}, \tag{82}$$

while $v_A=v_{in}$ is the Alfvén speed on the acoustic surface for $\Omega=0$, given by

$$v_A^2=\frac{\Psi_0^2}{64\pi^4 r_A^4\mu^2\eta^2}. \tag{83}$$

Finally, r_A and r_a are the Alfvén radii for a spherically symmetric flow and for a flow with $\alpha\neq0$. For a flow that is close to spherically symmetric, naturally we have $\alpha\to0$, $\xi\to1$, $X_*\to1$, and $\Psi_*\to\Psi_0$.

Then, the condition $D=0$ is written in terms of these quantities as

$$\omega^2=\frac{\Psi_0^2}{\Psi_*^2}\,\frac{X_*^6\alpha^2}{\xi_*^2M_*^6}, \tag{84}$$

and the coupling equation (71) is written as

$$1=\frac{M_*^4}{X_*^4}\,\frac{\Psi_*^2}{\Psi_0^2}+\alpha^2(2X_A^2-X_*^2)+\frac{\alpha^2(X_A^2-X_*^2)^2}{X_*^2\omega^4}. \tag{85}$$

Finally, the coefficient η_2 can be determined from the equilibrium equation (22):

$$\eta_2=-\omega^2\left[\frac{\alpha_2}{2}-\alpha_5-\frac{1}{2}-\frac{\alpha^2X_A^2X_*^2}{M_*^2(1-M_*^2)}\,\frac{\Psi_0^2}{\Psi_*^2}\right.$$
$$\left.\times(X_*^2-X_A^2M_*^2)\right]-\frac{1}{2}\,\eta_1\alpha_1. \tag{86}$$

As in the case of a hydrodynamic flow, therefore, the expressions for all the expansion coefficients in Eqs. (76)–(78) have a square-root singularity. Unfortunately, Eq. (78) for η_1 requires that four quantities characterizing the flow be specified, but these cannot be determined directly from the boundary conditions (64)–(67). Thus, the value of the angular velocity α at which the expression under the square-root sign becomes negative (if it happens at all) cannot be determined with certainty. For small $\alpha\ll1$, however, Eq. (78) can be written in the following form:

$$\eta_1=-\frac{4}{3}-\frac{2}{3}\,\alpha^2-\sqrt{\frac{4}{9}-\frac{2}{9}\,\alpha^2}. \tag{87}$$

In particular, for $\alpha=0$ Eq. (87) gives $\eta_1=-2$, which just corresponds to free outflow of the material. As we see, Eq. (87) shows that the expression under the square-root sign decreases as the angular rotation speed increases. Thus, it is possible that as a spherically symmetric flow is distorted further, it may become discontinuous. Furthermore, according to Eq. (87), the expression under the square-root sign becomes negative right at $\alpha\sim1$, which corresponds exactly to Bogovalov's result.

We note, finally, that for $\alpha\ll1$, the coefficient a in Eq. (30) has the form

$$a=-\frac{6\alpha^2}{D_1^2}. \tag{88}$$

Clearly, as in the case of a hydrodynamic flow, at the moment the characteristics undergo a bifurcation, the flow ceases to be discontinuous. The transition to a regime in which the flow becomes discontinous corresponds to $a=-\infty$.

6. CONCLUSION

In this paper we have shown that the solution of the equilibrium equation near the singular points remains continuous while the characteristics undergo a bifurcation, but may become discontinuous as the flow is distorted further. The possibility of shock wave formation, obtained numerically by Bogovalov, is thereby demonstrated. On the other hand, the continuity of the procedure for determining the flow parameters at $a=-1/8$ shows that a sudden change in the position of the separatrix characteristic (which separates regions that are not coupled causally) does not affect the number of boundary conditions required to determine the flow structure.

We thank A. V. Gurevich for his interest in this work, useful discussions, and support, as well as V. B. Baranov and S. V. Bogovalov for fruitful discussions. This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-18217).

[1] K. G. Guderley, *Theory of Transonic Flow*, Pergamon, Oxford (1962).
[2] R. Mises, *Mathematical Theory of Compressible Fluid Flow*, Izd. Inostr. Lit., Moscow (1961); Academic, New York (1958).
[3] F. I. Frankl', *Selected Papers on Gas Dynamics* [in Russian], Nauka, Moscow (1970).
[4] M. M. Smirnov, *Equations of Mixed Type*, Nauka, Moscow (1986); Amer. Math. Soc., Providence (1978).
[5] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics*, Nauka, Moscow (1986); Pergamon, Oxford (1987).
[6] R. D. Blandford and D. G. Payne, Mon. Not. R. Astron. Soc. **199**, 883 (1982).
[7] B. C. Low and K. Tsynganos, Astrophys. J. **302**, 163 (1986).
[8] G. S. Bisnovatyĭ-Kogan, Ya. M. Kazhdan, A. A. Klypin *et al.*, Astron. Zh. **56**, 359 (1979) [Sov. Astron. **23**, 201 (1979)].
[9] K. Tsinganos and C. Sauty, Astron. Astrophys. **255**, 405 (1992).
[10] R. V. E. Lovelace, J. C. L. Wang, and M. E. Sulkanen, Astrophys. J. **315**, 504 (1987).
[11] J. F. Hawley, L. L. Smarr, and J. R. Wilson, Astrophys. J. **277**, 296 (1984).
[12] G. W. Pneuman and R. A. Kopp, Sol. Phys. **18**, 258 (1971).
[13] Zh.-Yu Li, T. Chuen, and M. C. Begelman, Astrophys. J. **394**, 459 (1992).
[14] R. Hunt, Mon. Not. R. Astron. Soc. **198**, 83 (1979).
[15] L. I. Petrich, S. Shapiro, R. F. Stark, and S. Teukolsky, Astrophys. J. **336**, 313 (1989).
[16] M. Ruffert and D. Arnett, Astrophys. J. **427**, 351 (1994).
[17] X. Cao and H. C. Spruit, Astron. Astrophys. **287**, 80 (1994).
[18] S. V. Bogovalov, Mon. Not. R. Astron. Soc. **280**, 39 (1996).
[19] V. S. Beskin and Yu. N. Pidoprygora, Zh. Éksp. Teor. Fiz. **107**, 1025 (1995) [JETP **80**, 575 (1995)].
[20] V. S. Beskin and V. I. Par'ev, Usp. Fiz. Nauk **163**, 95 (1993) [Phys. Usp. **36**, 529 (1993)].
[21] V. I. Pariev, Mon. Not. R. Astron. Soc. **283**, 1264 (1996).
[22] Ya. B. Zeldovich and I. D. Novikov, *Relativistic Astrophysics*, Univ. of Chicago Press, Chicago (1971), vol. 1.
[23] S. Shapiro and S. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars: the Physics of Compact Objects*, Wiley, N. Y. (1983).

Translated by D. H. McNeill

Annihilation of a relativistic positron and *K*-electron to yield a photon and second *K*-electron

A. I. Mikhaĭlov and I. A. Mikhaĭlov*)

B. P. Konstantinov St. Petersburg Institute of Nuclear Physics, Russian Academy of Sciences, St. Petersburg, 188350 Gatchina, Leningrad Region, Russia

(Submitted 15 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 786–804 (March 1998)

We study the annihilation of a fast positron and a *K*-electron resulting in the emission of a photon and a second *K*-electron. It is assumed that all electrons and positrons move in the Coulomb field of the nucleus and that the Coulomb parameter αZ is much less than unity ($\alpha = 1/137$ is the fine-structure constant and Z is the atomic number). The electron–electron interaction, which is responsible for the ejection of the electron by the atom, is taken into account in the first order of perturbation theory. We calculate the differential and total cross sections of the process and construct the ratio of the cross sections of double and single ionization as a function of the energy of the incident positron. Finally, we establish the high-energy limit of this ratio, equal to $0.34/Z^2$. © *1998 American Institute of Physics.* [S1063-7761(98)00203-0]

1. INTRODUCTION

The annihilation of a positron and a free electron is only possible if two or more photons are emitted, since otherwise it is impossible to conserve energy and momentum. But since annihilation with an atomic electron involves a third object (the nucleus), it can be two-photon, one-photon, or even radiationless if the atom contains more than one electron:

$$\mathrm{A}+e^+\to\mathrm{A}^+ + 2\gamma, \tag{1}$$

$$\mathrm{A}+e^+\to\mathrm{A}^+ + \gamma, \tag{2}$$

$$\mathrm{A}+e^+\to\mathrm{A}^{++} + e^-, \tag{3}$$

where A is a neutral atom or ion, and A^+ and A^{++} are ions whose charges exceed that of A by one or two units.

The reaction (1) is the crossing-symmetry channel for the Compton scattering of photons by the atom and can proceed with low momentum transfer to the nucleus, as in the case of Compton scattering. This means that among the above reactions, this one has the largest cross section, equal in order of magnitude to the photon scattering cross section of a free electron, $\sigma\sim r_0^2$, where $r_0 = 2.82\times 10^{-13}$ cm is the classical electron radius. Two-photon annihilation on atoms has yet to be studied more thoroughly. We know of only one experimental paper[1] and one theoretical paper,[2] and the results of these two studies are quite different.

One-photon annihilation (2) is the crossing-symmetry channel for the atomic photoeffect. The reaction proceeds with high momentum transfer $q\sim m$ (m is the electron mass)[1)] and has cross section $\sigma\sim r_0^2\alpha^4 Z^5$. One-photon annihilation on atoms is thoroughly studied in Refs. 3–5.

Radiationless annihilation (3) is studied in Refs. 6–9. In this reaction, the energy liberated in the electron–positron annihilation process is transferred to another electron that leaves the atom. As a result, a doubly charged ion and free electron emerge with definite energy. Reaction (3) proceeds with large momentum transfer $q>2m$ to the nucleus and large energy transfer $\Delta E>2m$ to the liberated electron. The reaction therefore has a small cross section $\sigma\sim r_0^2(\alpha Z)^8$ and occurs at small distances $r\sim 1/q$ from the nucleus.[8,9]

More complicated annihilation processes on an atom are also possible. Among these is one-photon annihilation accompanied by ejection of an electron (the name of the process corresponds to the number of photons actually emitted in the reaction). As a result the atom becomes doubly ionized, and the process can be called double ionization with one-photon annihilation:

$$\mathrm{A}+e^+\to\mathrm{A}^{++} + e^- + \gamma. \tag{4}$$

The energy of this reaction is divided between the photon and the electron. Hence, in contrast to (2) and (3), here we have photon and electron energy spectra.

We show below that the cross section of this reaction for relativistic positrons is $\sigma\sim r_0^2\alpha(\alpha Z)^3$. Notwithstanding the fact that there emerges, in comparison to reaction (3), a new small factor $\sim\alpha$ related to photon emission, the cross section exceeds that of radiationless annihilation. The released energy is divided between the photon and the electron very unequally. Two regions of the energy spectrum contribute most to the cross section: the edge region, where the emission of a hard photon and low-energy electron with high momentum transfer to the nucleus is observed, and the central region, where both the photon and the electron have high energies, while the momentum transferred to the nucleus is low. The edge region is extremely narrow (its width is of order of the binding energy of the *K*-electron), but its contribution to the total cross section is comparable to the contribution of the rest of the spectrum. The central region comprises about half the spectrum, but the section of this region

that contributes most to the cross section corresponds to ejection of slower electrons.

The electron–electron interaction plays the most important role in reaction (4), with the result that this reaction together with double photoionization

$$A+\gamma\rightarrow A^{++}+2e^{-} \tag{5}$$

are of interest to studies of electron correlations in the relativistic collision-energy range.[10] Granted, experimental studies of relativistic processes on low-Z elements or on the outer shells of high-Z atoms, where correlation effects are most in evidence, are hindered by the smallness of the corresponding cross sections. However, the behavior of cross sections in the relativistic range has its peculiarities and requires separate study. For instance, while the ratio of the cross section of double photoionization to that of single photoionization in the limit of high but still nonrelativistic energies is constant,[11–13] in the relativistic range it increases with photon energy and approaches a new limit.[10]

In this paper we examine the double ionization of the K shells of atoms and ions in the one-photon annihilation of high-energy positrons with atomic electrons. We assume that the electrons and the positron move in the Coulomb field of a nucleus with charge Z. The electron–electron interaction is taken into account in the first order of perturbation theory, which is justified for $Z\gg 1$. At the same time, we assume that $\alpha Z\ll 1$, and all analytic expressions are derived by series expansion in this parameter. However, the problem contains one more Coulomb parameter, $\xi=\alpha ZE/p$ (E and p are the electron energy and momentum), which is not always small. In such cases the dependence on ξ is taken into account exactly. We find the energy spectrum, the angular distribution of electrons and photons, and the total cross section of the process. We calculate the energy dependence of the ratio of the cross sections of double and single ionization. We also give simple formulas for the high-energy limit in the relativistic and nonrelativistic region. Finally, we show that at nonrelativistic positron energies, the double-ionization cross section can exceed the single ionization cross section in the case of annihilation on light atoms.

2. AMPLITUDE AND CROSS SECTION OF DOUBLE IONIZATION WITH ONE-PHOTON ANNIHILATION IN THE EDGE REGION OF THE ENERGY SPECTRUM

To first order in the electron–electron interaction, the process of double ionization with one-photon annihilation is represented by eight Feynman diagrams, four of which are depicted in Fig. 1. The other four diagrams differ from these in the sign and interchange of the states ψ_α and ψ_β of the initial electrons.[2)]

When a photon with energy $\omega\ll m$ is emitted, all diagrams are of the same order, and their contribution to the total cross section is α^{-1} times smaller than the cross section of the radiationless annihilation reaction (3). Due to the smallness of this contribution, we do not consider the low-energy region of the phonon spectrum.

When a hard photon and a low-energy electron are ejected, diagrams a and b are dominant, since their photon

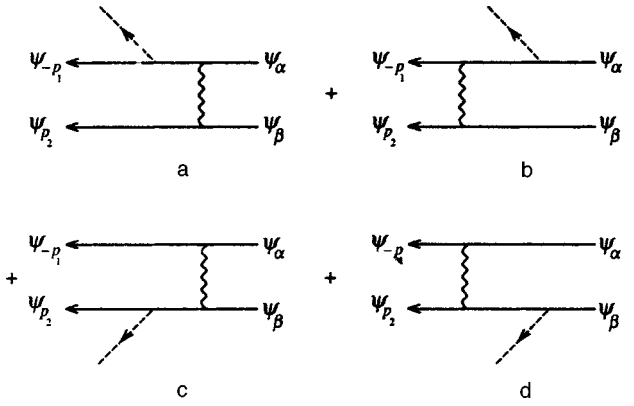

FIG. 1. Feynman diagrams for the process (4). Solid lines denote electrons in the Coulomb field of the nucleus, dashed lines denote photons, and wavy lines denote the electron–electron interaction (photon propagator).

propagator is much larger than the photon propagator of diagrams c and d. To elaborate, in diagrams a and b, the virtual photon carries energy $E_2-E_\beta=\varepsilon_2+I$ (here $\varepsilon_2=E_2-m$ is the kinetic energy of the ejected electron, and $I=m\alpha^2Z^2/2$ is the binding energy of the K electron), while in the diagrams c and d the transferred energy is $E_1+E_\alpha\simeq E_1+m$ (E_1 is the positron energy). Accordingly, diagrams a and b represent terms that are much larger than those represented by c and d.

When a hard photon and high-energy electron are ejected, all four diagrams in Fig. 1 together with all four diagrams obtained from the former by interchanging the initial (or final) states must be taken into account

Thus, in the high-energy region of the photon spectrum, there are two regions that make the largest contributions to the total cross section: the edge region, adjacent to the limiting value of the photon energy $\omega_{\max}=E_1+m$ or the region of low-energy electrons (with kinetic energies $\varepsilon_2\sim I$), and the central region, where the reaction energy is apportioned to the photon and electron is such a way that their total momentum differs from the positron momentum only by a quantity of order $\eta=m\alpha Z$ (the momentum q transferred to the nucleus is low: $q\sim\eta$). These two regions must be examined separately, since in the first the nonrelativistic approximation remains valid to a great extent, while in the second only a fully relativistic approach can be used.

We start with the edge region of the photon spectrum. Here we discard diagrams c and d in view of their extreme smallness. Let us compare diagrams a and b. The significant disparity between these diagrams is due to the different values of the electron propagators G_a and G_b in the region of the spectrum where the electron carries away only a small amount of energy, $\varepsilon_2\sim I$. The propagator G_a is the Coulomb Green's function of the low-energy electron, while G_b is the Green's function of the electron with a large negative energy ($-E_1$). Since the energy enters into the denominator of the propagator, diagram b is much smaller than a, and can be discarded. The amplitude of double ionization with one-photon annihilation in the edge region is represented by diagram a and the diagram obtained from a by interchanging the initial states ψ_α and ψ_β:

$$M^{++}_{\text{edge}}=M_{\alpha\beta}-M_{\beta\alpha}. \tag{6}$$

In the coordinate representation we can write the amplitude $M_{\alpha\beta}$ as follows:[3)]

$$M_{\alpha\beta}=\int \overline{\phi}(\mathbf{r}')\gamma^{\mu}\psi_{\alpha}(\mathbf{r}')d\mathbf{r}' \times\int \frac{e^{iR\Delta}}{4\pi R}\,\overline{\psi}_{p_2}(\mathbf{r})\gamma_{\mu}\psi_{\beta}(\mathbf{r})d\mathbf{r}, \tag{7}$$

where

$$\overline{\phi}(\mathbf{r}')=\int \overline{\psi}_{-p_1}(\mathbf{r}'')\hat{e}^{*}\,\exp(-i\mathbf{k}\cdot\mathbf{r}'')G_C^E(\mathbf{r}'',\mathbf{r}')d\mathbf{r}'', \tag{8}$$

$$R=|\mathbf{r}-\mathbf{r}'|,\quad \Delta=E_2-E_{1s}=\varepsilon_2+I,$$

$$E=E_{1s}-\Delta=m-\varepsilon_2-2I, \tag{9}$$

$2I$ is the ionization energy of the K shell, $\hat{a}=\gamma_0 a_0-\boldsymbol{\gamma}\cdot\mathbf{a}$, γ_0 and $\boldsymbol{\gamma}$ are Dirac matrices, and $G_C^E(\mathbf{r}'',\mathbf{r}')$ is the relativistic Coulomb Green's function for an electron with energy E. We assume that (7) is summed over μ.

We are not concerned with polarization phenomena, so that we limit ourselves to the case of linear polarization of photons with momentum $\mathbf{k}$ and polarization vector $\mathbf{e}$. Then

$$\hat{e}^{*}=\hat{e}=-\boldsymbol{\gamma}\cdot\mathbf{e},\quad \mathbf{e}\cdot\mathbf{k}=0. \tag{10}$$

The Dirac-conjugate wave function $\overline{\psi}$ is equal to $\psi^{\dagger}\gamma_0$, the functions $\overline{\psi}_{-p_1}$ and $\overline{\psi}_{p_2}$ are the wave functions of the continuous-spectrum positron and electron, E_1 and $\mathbf{p}_1$ (E_2 and $\mathbf{p}_2$) are the energy and momentum of the positron (emitted electron), and ψ_{α} and ψ_{β} are the wave functions of K electrons with different orientation of spin. The wave functions of all of the electrons and the positron are of the Coulomb type, which is a fairly good approximation for ions and atoms, since the integrals with respect to $\mathbf{r}$ and $\mathbf{r}'$ become saturated at distances $r\sim r'\sim 1/\eta$ ($\eta=m\alpha Z$ is the mean momentum of a K electron), where the field of the nucleus is only slightly screened by the electron. The integral with respect to r'' converges at even smaller distances, for $r''\sim 1/|\mathbf{k}-\mathbf{p}|\sim 1/m$, where the field is of the Coulomb type.

Since in the edge region the emitted electron has low momentum ($p_2\sim\eta$), the vector part of the electron current is small compared to the scalar part:

$$\overline{\psi}_{p_2}\boldsymbol{\gamma}\psi_{\beta}\sim\frac{\mathbf{p}_2}{m}\varphi_{p_2}^{*}\varphi_{\beta},\quad \overline{\psi}_{p_2}\gamma_0\psi_{\beta}\sim\varphi_{p_2}^{*}\varphi_{\beta}, \tag{11}$$

where φ_{p_2} and φ_{β} are the nonrelativistic Coulomb wave functions of a free and bound electron, respectively.

Subsequent calculations will be carried out in the momentum representation. If we allow for (11), the expression for the amplitude $M_{\alpha\beta}$ simplifies considerably:

$$M_{\alpha\beta}=\int\frac{d\mathbf{f}}{(2\pi)^3}F_1(-\mathbf{p}_1,\mathbf{f},\alpha)D(\mathbf{f})F_2(\mathbf{p}_2,\mathbf{f},\beta), \tag{12}$$

where

$$F_1(-\mathbf{p}_1,\mathbf{f},\alpha)=\int\frac{d\mathbf{f}'d\mathbf{f}_1}{(2\pi)^6}\langle\psi_{-p_1}|\mathbf{f}'-\mathbf{k}\rangle\hat{e}\langle\mathbf{f}'|G_C|\mathbf{f}_1\rangle \times\gamma_0\langle\mathbf{f}_1+\mathbf{f}|\psi_{\alpha}\rangle, \tag{13}$$

$$F_2(\mathbf{p}_2,\mathbf{f},\beta)=\int\frac{d\mathbf{f}_2}{(2\pi)^3}\langle\psi_{p_2}|\mathbf{f}_2\rangle\gamma_0\langle\mathbf{f}_2-\mathbf{f}|\psi_{\beta}\rangle, \tag{14}$$

$$D(\mathbf{f})=\frac{1}{\mathbf{f}^2-\Delta^2-i0},\quad \Delta=\varepsilon_2+I, \tag{15}$$

with $\langle\mathbf{f}'|G_C|\mathbf{f}\rangle\equiv G_C^E(\mathbf{f}',\mathbf{f})$ the relativistic Coulomb Green's function in the momentum representation.

Since the integrals (12)–(14) saturate at $f\sim f_1\sim f_2\sim\eta$, in the lowest-order approximation in αZ we can put

$$D(\mathbf{f})=\frac{1}{f^2} \tag{16}$$

and use nonrelativistic Coulomb wave functions for the K electrons and the ejected electron:

$$|\psi_{\alpha}\rangle=|1s\rangle u_0(\alpha),\quad |\psi_{\beta}\rangle=|1s\rangle u_0(\beta),\quad \langle\psi_{p_2}|=\overline{u}_{p_2}\langle\varphi_{p_2}|,$$

$$u_0(\alpha)=\begin{pmatrix}w(\alpha)\\0\end{pmatrix},\quad u_0(\beta)=\begin{pmatrix}w(\beta)\\0\end{pmatrix},$$

$$w(\alpha)=\begin{pmatrix}1\\0\end{pmatrix},\quad w(\beta)=\begin{pmatrix}0\\1\end{pmatrix},$$

$$\overline{u}_{p_2}=(w_2^{+},0),\quad w_2\equiv w(\lambda_2),\quad w^{+}(\lambda)w(\lambda)=1,$$

$$|1s\rangle=N_1\left(-\frac{\partial}{\partial\eta}\right)V_{i\eta}|0\rangle,\quad N_1^2=\frac{\eta^3}{\pi},\quad \eta=m\alpha Z,$$

$$\langle\mathbf{f}'|V_{i\eta}|\mathbf{f}\rangle=\frac{4\pi}{(\mathbf{f}'-\mathbf{f})^2+\eta^2}. \tag{17}$$

Here u is a Dirac bispinor, w is a Pauli spinor, and $|0\rangle$ is a plane-wave state with zero momentum.

Using (17), we can derive the following expression for F_2 (see Eq. (14)):

$$F_2(\mathbf{p}_2,\mathbf{f},\beta)=w_2^{+}w(\beta)N_1\left(-\frac{\partial}{\partial\eta}\right)\langle\varphi_{p_2}|V_{i\eta}|\mathbf{f}\rangle. \tag{18}$$

The calculation of F_1 of Eq. (13) is much more complicated, since the function contains the relativistic Coulomb Green's function with low kinetic energy $E-m\sim I$. For this Green's function, the Coulomb parameter $\xi=\alpha ZE/p$ is roughly equal to unity and cannot be used as a small expansion parameter. On the other hand, for a relativistic positron, there is an expansion in powers of αZ and $\xi_1=\alpha ZE_1/\,p_1\approx\alpha Z$ (see Ref. 14):

$$\langle\psi_{-p_1}|=\overline{u}_{-p_1}\{\langle-\mathbf{p}_1|-\alpha Z\langle-\mathbf{p}_1|\hat{V}_0G_1+\cdots\}, \tag{19}$$

where G_1 is the relativistic propagator of a free electron with energy $-E_1$, and $\overline{u}_{-p_1}=u_{-p_1}^{\dagger}\gamma_0$; u_{-p_1} is a bispinor with four-momentum $-p_1$. The first term in this expansion (the zeroth approximation in αZ) is a plane wave. We keep two terms in (19), since plugging only the first term into (13) yields zero in the limit $\alpha Z\to 0$. Denoting the contribution of the plane wave to the integral (13) by F_{10}, a calculation that uses the function ψ_{α} from (17) yields

$$F_{10}=N_1\bar{u}_{-p_1}\hat{e}U(\mathbf{f})u_0, \quad U(\mathbf{f})=-\frac{\partial}{\partial\eta}\langle\boldsymbol{\kappa}|G_CV_{i\eta}|-\mathbf{f}\rangle, \tag{20}$$

where

$$\boldsymbol{\kappa}=\mathbf{k}-\mathbf{p}_1.$$

Equation (20) contains the matrix element of the relativistic Coulomb Green's function for an electron with low kinetic energy. However, this Green's function cannot be replaced by the nonrelativistic one, since it is calculated for a large value of one of the momenta (based on energy conservation, we can show that $\kappa>2m$). We transform $U(\mathbf{f})$ in such a way that the Green's operator G_C is sandwiched between the nonrelativistic momenta $f\sim f'\sim\eta$. To this end we use the Lippmann–Schwinger equation[15]

$$G_C=G-\alpha ZG\hat{V}_0G_C. \tag{21}$$

Here $\hat{V}_0=\gamma_0V_0$, $-\alpha ZV_0$ is the operator of the interaction of an electron and the Coulomb field of the nucleus, and G is the relativistic Green's operator in zero field. The matrix element of this operator is defined to be

$$\langle\mathbf{f}|G|\mathbf{f}'\rangle=G(\mathbf{f})(2\pi)^3\delta(\mathbf{f}-\mathbf{f}'),$$

$$G(\mathbf{f})=\frac{E\gamma_0-\boldsymbol{\gamma}\cdot\mathbf{f}+m}{p^2-f^2+i0}, \quad p^2=E^2-m^2. \tag{22}$$

Note that for $E-m\ll m$ and $f\sim\eta$, the relativistic operator $G(\mathbf{f})$ is simply related to the nonrelativistic operator by

$$G(\mathbf{f})=G^{nr}(\mathbf{f})\frac{\gamma_0+1}{2}, \quad G^{nr}(\mathbf{f})=\frac{2m}{p^2-f^2+i0}. \tag{23}$$

Plugging (21) into (20), we obtain

$$U(\mathbf{f})u_0=-\frac{\partial}{\partial\eta}G(\boldsymbol{\kappa})\left\{\langle\boldsymbol{\kappa}|V_{i\eta}|-\mathbf{f}\rangle-\alpha Z\int\frac{d\mathbf{f}'}{(2\pi)^3}\times\langle\boldsymbol{\kappa}|\hat{V}_0|\mathbf{f}'\rangle\langle\mathbf{f}'|G_CV_{i\eta}|-\mathbf{f}\rangle\right\}u_0. \tag{24}$$

The dominant contribution to the integral in (24) is provided by the region where $f'\sim\eta$. At such values of $\mathbf{f}'$ the factor $\langle\boldsymbol{\kappa}|V_0|\mathbf{f}'\rangle\approx4\pi/\kappa^2$ can be taken outside the integral sign, and the relativistic function G_C reduces to the nonrelativistic Coulomb Green's function G_C^{nr}. Indeed, iterating the Lippmann–Schwinger equation, we obtain

$$\begin{aligned}G_Cu_0&=(G-\alpha ZG\hat{V}_0G+\cdots)u_0\\&=(G^{nr}-\alpha ZG^{nr}V_0G^{nr}+\cdots)u_0\\&=G_C^{nr}u_0,\end{aligned} \tag{25}$$

where we have used (23) and the equalities

$$\gamma_0^2=1, \quad \left(\frac{\gamma_0+1}{2}\right)^n=\frac{\gamma_0+1}{2}, \quad \frac{\gamma_0+1}{2}u_0=u_0. \tag{26}$$

In (24), the two terms in braces are quantities of the same order,[4)] but after we have taken the derivative with respect to η, the first term is multiplied by $2\eta/\kappa^2$, while the second is multiplied by a quantity $\sim1/\eta$. The term in (24) should therefore be discarded. Then

$$U(\mathbf{f})u_0\simeq\frac{4\pi\alpha Z}{\kappa^2}G(\boldsymbol{\kappa})\frac{\partial}{\partial\eta}\int\frac{d\mathbf{f}'}{(2\pi)^3}\times\langle\mathbf{f}'|G_C^{nr}V_{i\eta}|-\mathbf{f}\rangle u_0. \tag{27}$$

Using the results obtained in Refs. 15 and 16, we obtain a simple expression for the integral in (27):

$$\begin{aligned}J(\eta,\mathbf{f})&=\int\frac{d\mathbf{f}'}{(2\pi)^3}\langle\mathbf{f}'|G_C^{nr}V_{i\eta}|-\mathbf{f}\rangle\\&=\frac{2ipm}{4\pi}I_y\langle0|V_{py+i\eta}|\mathbf{f}\rangle,\\I_y&=\int_1^\infty dy\left(\frac{y+1}{y-1}\right)^{i\xi}, \quad \xi=\frac{\alpha Zm}{p}.\end{aligned} \tag{28}$$

The final expression for F_{10} assumes the form

$$F_{10}=N_1\bar{u}_{-p_1}\hat{e}\left(1+\frac{\tilde{\kappa}}{2m}\right)u_0\frac{4ip\eta m}{\kappa^4}\times\left(-\frac{\partial}{\partial\eta}\right)I_y\langle\mathbf{f}|V_{py+i\eta}|0\rangle, \tag{29}$$

where $\tilde{a}=\boldsymbol{\alpha}\cdot\mathbf{a}$, and $\boldsymbol{\alpha}=\gamma_0\boldsymbol{\gamma}$ is the Dirac matrix.

We now evaluate F_{11}, the contribution of the second term in (19) to the integral (13):

$$\begin{aligned}F_{11}&=\alpha ZN_1\bar{u}_{-p_1}\langle\boldsymbol{\kappa}|V_0|0\rangle G_1(-\mathbf{k})\hat{e}\frac{\partial}{\partial\eta}J(\eta,\mathbf{f})u_0\\&=\frac{4\pi\alpha ZN_1}{\kappa^2(\omega^2-p_1^2)}\frac{\partial J(\eta,\mathbf{f})}{\partial\eta}\bar{u}_{-p_1}(\omega-\tilde{k})\hat{e}u_0,\end{aligned} \tag{30}$$

where ω is the photon energy.

Summing the contributions (29) and (30), we find the amplitude F_1 (Eq. (13)):

$$F_1(-\mathbf{p}_1,\mathbf{f},\alpha)=F_{10}+F_{11}=-T(\lambda_1,\alpha)\frac{\partial J(\eta,\mathbf{f})}{\partial\eta}, \tag{31}$$

$$T(\lambda_1,\alpha)=N_1\frac{8\pi\eta}{\kappa^4}\bar{u}_{-p_1}(\lambda_1)\hat{e}\times\left(1+\frac{\tilde{\kappa}}{2m}+\frac{\tilde{k}-\omega}{2m}\frac{\kappa^2}{\omega^2-p_1^2}\right)u_0(\alpha). \tag{32}$$

Here $T(\lambda_1,\alpha)$ is the one-photon annihilation amplitude of a positron with polarization λ_1 and a K electron with polarization α, calculated to lowest order in αZ.

Inserting (16), (18), and (31) into (12), we obtain

$$M_{\alpha\beta}=A_{\alpha\beta}I_y\frac{\partial^2}{\partial\eta\partial\eta'}\int\frac{d\mathbf{f}}{(2\pi)^3}\times\langle\varphi_{p_2}|V_{i\eta'}|\mathbf{f}\rangle\frac{1}{f^2}\langle\mathbf{f}|V_{py+i\eta}|0\rangle, \tag{33}$$

$$A_{\alpha\beta}=\frac{ipm}{2\pi}N_1T(\lambda_1,\alpha)w_2^+w(\beta). \tag{34}$$

After taking the derivatives in (33), we can put $\eta'=\eta$. Using the equality

$$\frac{1}{f^2}\langle \mathbf{f}|V_{ia}|0\rangle = \frac{1}{a^2}\langle \mathbf{f}|V_0 - V_{ia}|0\rangle$$

and the operator identity[14]

$$\frac{\partial}{\partial a}V_{ia}V_{ib} = V_{i(a+b)},$$

we can transform (33) into

$$M_{\alpha\beta} = A_{\alpha\beta}I_y \frac{\partial}{\partial\mu}\frac{1}{(\mu-\eta')^2}\langle \varphi_{p_2}|V_{i\mu} - V_{i\eta}|0\rangle,$$

$$\mu = 2\eta + ry, \quad r = -ip = \sqrt{p_2^2 + 2\eta^2}. \tag{35}$$

For real values of a, the matrix element $\langle \varphi_{p_2}|V_{ia}|0\rangle$ with the Coulomb function of the continuous spectrum, φ_{p_2}, has the form[17]

$$\langle \varphi_{p_2}|V_{ia}|0\rangle = \frac{4\pi N_{p_2}}{p_2^2 + a^2}\exp\left(-2\xi_2 \arctan\frac{p_2}{a}\right),$$

$$N_{p_2}^2 = \frac{2\pi\xi_2}{1-\exp(-2\pi\xi_2)}, \quad \xi_2 = \frac{\eta}{p_2}. \tag{36}$$

We introduce a dimensionless parameter ν and express the other parameters of the problem in terms of ν:

$$\nu = \frac{\varepsilon_2}{I} = \frac{P_2^2}{\eta^2}, \quad \xi_2 = \frac{\eta}{p_2} = \frac{1}{\sqrt{\nu}}, \quad i\xi = \frac{\eta}{r} = \frac{1}{\sqrt{\nu+2}}. \tag{37}$$

In (35) we transform to a new integration variable,

$$t = \frac{1-y}{1+y}.$$

Substituting (36) into (35) yields

$$M_{\alpha\beta} = -K(\nu)T(\lambda_1,\alpha)w^+(\lambda_2)w(\beta), \tag{38}$$

where

$$K(\nu) = N_1 N_{p_2}\frac{m}{\eta^4}J(\nu), \tag{39}$$

$$J(\nu) = \frac{8\zeta^2}{(1+\zeta)^3}\left(\frac{I_1}{\nu+1} - \frac{I_2}{\nu+2}\right), \quad \zeta = \frac{1}{\sqrt{\nu+2}}, \tag{40}$$

$$I_1 = \exp\left(-\frac{2}{\sqrt{\nu}}\arctan\sqrt{\nu}\right)\int_0^1 dt\,\frac{t^{-\zeta}(1-t)}{(1+st)^3}, \tag{41}$$

$$I_2 = \int_0^1 dt\,\frac{t^{-\zeta}(1-t)^3}{(1+st)^3}\,\phi_1(t)\phi_2(t), \quad s = \frac{1-\zeta}{1+\zeta}, \tag{42}$$

$$\phi_1(t) = \exp\left[-\frac{2}{\sqrt{\nu}}\arctan\frac{\sqrt{\nu}(1-t)}{a+bt}\right],$$

$$a = \sqrt{\nu+2}+2, \quad b = \sqrt{\nu+2}-2,$$

$$\phi_2(t) = \frac{(3\zeta^2+1)(1-t)^2 + 6\zeta(1-t^2) + 2(1+t)^2}{[(2\zeta^2+1)(1-t)^2 + 4\zeta(1-t^2) + (1+t)^2]^2}. \tag{43}$$

The differential cross section of double ionization with one-photon annihilation summed over the polarizations of the photon (λ_k) and the final electron (λ_2) and averaged over the positron polarization (λ_1) assumes the form[5)]

$$d\sigma^{++} = \frac{(4\pi\alpha)^3}{2\omega j}\overline{|M^{++}|^2}\frac{d\mathbf{p}_2 d\mathbf{k}}{(2\pi)^5}\delta(\omega + E_2 - E_1 - 2m), \tag{44}$$

where $j = p_1/E_1$ is the positron flux and M^{++} is the amplitude of double ionization with one-photon annihilation.

Equation (44) can be used to obtain the cross section for the edge region by replacing $\overline{|M^{++}|^2}$ with

$$\begin{aligned}\overline{|M_{\text{edge}}^{++}|^2} &= \frac{1}{2}\sum_{\lambda_k\lambda_1\lambda_2}|M_{\alpha\beta} - M_{\beta\alpha}|^2 \\ &= K^2(\nu)\frac{1}{2}\sum_{\lambda_k\lambda_1}[|T(\lambda_1,\alpha)|^2 + |T(\lambda_1,\beta)|^2] \\ &= 2K^2(\nu)\frac{1}{4}\sum_{\lambda_k\lambda_1\lambda_0}|T(\lambda_1,\lambda_0)|^2 \\ &= 2K^2(\nu)\overline{|M^+|^2}.\end{aligned} \tag{45}$$

Here $\overline{|M^+|^2}$ is the square of the one-photon annihilation amplitude summed over the polarization of the photon (λ_k) and averaged over the polarizations of the positron (λ_1) and the bound electron (λ_0). The differential cross section of one-photon annihilation on two K electrons can be expressed in terms of this quantity:

$$d\sigma^+ = \frac{4\pi\alpha}{\omega j}\overline{|M^+|^2}\frac{d\mathbf{k}}{(2\pi)^2}\delta(\omega - E_1 - m). \tag{46}$$

Equation (44) is valid in any region of the spectrum. But a specific feature of the edge region is the simple relationship between the cross section of double and single ionization, which follows from (44)–(46):

$$d\sigma_e^{++} = \frac{2}{\pi}\alpha^2 K^2(\nu)d\mathbf{p}_2 d\sigma^+. \tag{47}$$

We direct the z axis along the positron momentum $\mathbf{p}_1$, denote the polar and azimuthal angles that specify the direction of the outgoing photon with momentum $\mathbf{k}$ (the electron with momentum $\mathbf{p}_2$) by θ_k and φ_k (by θ_2 and φ_2), respectively, and the solid angle into which the photon (electron) is ejected by $d\Omega_k$ ($d\Omega_2$). Then[6)]

$$d\mathbf{p}_2 = mp_2 d\varepsilon_2 d\Omega_2 = \frac{\eta^3}{2}\sqrt{\nu}\,d\nu d\Omega_2, \tag{48}$$

$$\pi d\sigma^+ = \sigma_0 Z^5 S(\theta_k)d\Omega_k, \quad \sigma_0 = \pi r_0^2\alpha^4, \tag{49}$$

$$S(\theta_k) = 2^5(E_1+m)\frac{p_1 m^4}{\kappa^6}\left(\frac{E_1+2m}{4m} - \frac{m^2}{\kappa^2}\right)\sin^2\theta_k,$$

$$\kappa = |\mathbf{p}_1 - \mathbf{k}|. \tag{50}$$

Plugging (48) and (49) into (47) yields an expression for the triple differential cross section

$$\frac{d\sigma_e^{++}}{d\nu\,d\Omega_2 d\Omega_k} = \frac{Q(\nu)}{4\pi Z^2}\frac{d\sigma^+}{d\Omega_k},$$

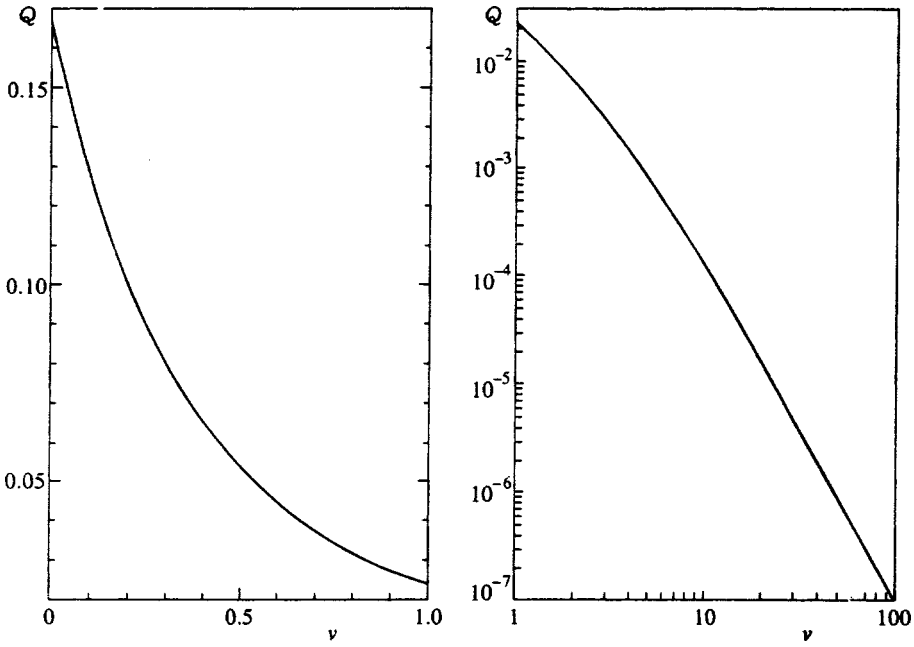

FIG. 2. The energy spectrum of low-energy electrons, $Q(\nu)=Z^2 d\sigma_e^{++}/\sigma^+ d\nu$, with $\nu=\varepsilon_2/I$.

$$Q(\nu)=\frac{8J^2(\nu)}{1-\exp(-2\pi/\sqrt{\nu})}. \tag{51}$$

A characteristic feature of (51) is the fact that the cross section is independent of the direction in which an electron is ejected, i.e., low-energy electrons produced by double ionization with one-photon annihilation are distributed isotropically in solid angle. Integrating (51) over the photon and electron ejection angles, we obtain an expression for the energy distribution of the low-energy electrons ($\varepsilon_2 \ll m$, or $\nu \ll (\alpha Z)^{-2}$):

$$\frac{d\sigma_e^{++}}{d\nu}=\frac{Q(\nu)}{Z^2}\sigma^+, \quad \sigma^+=\sigma_0 Z^5 \varphi(E_1), \tag{52}$$

$$\varphi(E_1)=\frac{4m^3}{p_1(E_1+m)^2}\left(\frac{E_1^2}{m^2}+\frac{2}{3}\frac{E_1}{m}+\frac{4}{3} - \frac{E_1+2m}{p_1}\ln\frac{E_1+p_1}{m}\right). \tag{53}$$

The function $Q(\nu)$ is plotted in Fig. 2. Employing Eqs. (40)–(42), we can obtain $Q(\nu)$ for small and large values of ν:

$$Q(0)=0.168, \quad Q(\nu)=\frac{4}{\pi}\nu^{-7/2} \quad \text{for} \quad \nu \gg 1. \tag{54}$$

Since $Q(\nu)$ rapidly decreases with increasing ν, in calculating the contribution of the entire edge region, we can extend the region of integration with respect to ν to infinity:

$$\sigma_e^{++}=\frac{B}{Z^2}\sigma^+=\sigma_0 Z^3 B\varphi(E_1),$$

$$B=\int_0^\infty Q(\nu)d\nu=0.090. \tag{55}$$

The ratio $Z^2\sigma_e^{++}/\sigma^+=B$ is a constant that depends neither on the energy of the incident positron nor on Z.

The angular distributions of the electrons and photons belonging to the edge region are very simple:

$$\frac{d\sigma_e^{++}}{d\Omega_2}=\frac{\sigma_e^{++}}{4\pi}, \quad \frac{d\sigma_e^{++}}{d\Omega_k}=\frac{B}{Z^2}\frac{d\sigma^+}{d\Omega_k}=B\sigma_0 Z^3 S(\theta_k). \tag{56}$$

Figure 3 depicts the angular function $S(\theta_k)$ for different positron energies. The function has a peak when $\theta<\pi/2$. The peak narrows and shifts toward smaller angles as the positron energy rises. There is no forward emission of photons, i.e., in the direction of the incident positron beam.

3. AMPLITUDE AND CROSS SECTION OF DOUBLE IONIZATION WITH ONE-PHOTON ANNIHILATION IN THE CENTRAL REGION OF THE ENERGY SPECTRUM

We call the region of the energy spectrum in which the momentum transferred to the nucleus is low ($q \sim \eta$) the central region. A rather broad middle region exists in the spectrum where $q=0$. Its boundaries are dictated by the compatibility of the two equations

$$\mathbf{p}_2+\mathbf{k}=\mathbf{p}_1, \quad E_2+\omega=E_1+2m. \tag{57}$$

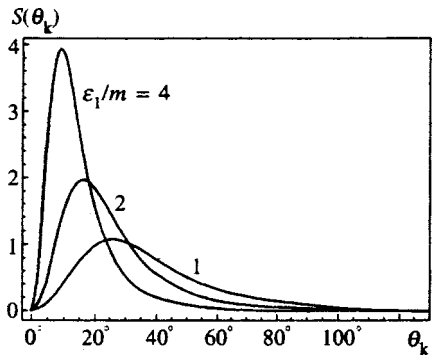

FIG. 3. Angular distribution $S(\theta_k)=\pi(\sigma_0 Z^5)^{-1}d\sigma^+/d\Omega_k$ of photons emitted in process (2).

Solving these equations, we find that

$$\frac{2m\omega_0}{E_0+p_1}\leqslant\omega\leqslant\frac{2m\omega_0}{E_0-p_1},$$
$$\omega_0=E_1+m,\quad E_0=E_1+2m. \tag{58}$$

Here ω_0 and E_0 are the maximum energies that the photon and electron emitted in the process of double ionization with one-photon annihilation can have. In the central region all four diagrams in Fig. 1 must be taken into account, as well as the diagrams with ψ_α and ψ_β interchanged.

We first calculate diagrams a and b in Fig. 1 (diagrams c and d are obtained from a and b by the interchanges $\psi_{-p_1}\leftrightarrow\psi_{p_2}$ and $\psi_\alpha\leftrightarrow\psi_\beta$). Since the ejected electron is relativistic, the virtual photon transfers a large amount of energy. The energy of the electron in the intermediate state (in the Green's function) is also high. For this reason, we can assume plane waves for the positron and electron wave functions, and a free-particle relativistic Green's function G. The wave functions of the bound electrons are still of the Coulomb type. The amplitudes M_a and M_b for these diagrams are

$$M_a=\overline{u}_{-p_1}\hat{e}G(\boldsymbol{\kappa})\gamma^\mu\int\frac{d\mathbf{f}}{(2\pi)^3}\langle\mathbf{q}-\mathbf{f}|\psi_\alpha\rangle \times D(\mathbf{p}_2-\mathbf{f})\overline{u}_{p_2}\gamma_\mu\langle\mathbf{f}|\psi_\beta\rangle, \tag{59}$$

$$M_b=\overline{u}_{-p_1}\gamma^\mu\int\frac{d\mathbf{f}}{(2\pi)^3}G(\mathbf{q}-\mathbf{k}-\mathbf{f})\hat{e}\langle\mathbf{q}-\mathbf{f}|\psi_\alpha\rangle \times D(\mathbf{p}_2-\mathbf{f})\overline{u}_{p_2}\gamma_\mu\langle\mathbf{f}|\psi_\beta\rangle. \tag{60}$$

The photon propagator $D(\mathbf{f})$ is defined in (15). In the energy range where $q\sim\eta$, the main contribution to the integrals (59) and (60) is provided by values of f close to η. Keeping only the leading terms, we obtain

$$D(\mathbf{p}_2-\mathbf{f})\simeq\frac{1}{2m\varepsilon_2},\quad G(\mathbf{q}-\mathbf{k}-\mathbf{f})\simeq G(-\mathbf{k}),$$
$$M_a=\frac{\phi(\mathbf{q})}{2m\varepsilon_2}\,\overline{u}_{-p_1}\hat{e}G(\boldsymbol{\kappa})\gamma^\mu u_0(\alpha)\overline{u}_{p_2}\gamma_\mu u_0(\beta), \tag{61}$$
$$M_b=\frac{\phi(\mathbf{q})}{2m\varepsilon_2}\,\overline{u}_{-p_1}\gamma^\mu G(-\mathbf{k})\hat{e}u_0(\alpha)\overline{u}_{p_2}\gamma_\mu u_0(\beta), \tag{62}$$
$$\phi(\mathbf{q})=\int\frac{d\mathbf{f}}{(2\pi)^3}\langle\mathbf{q}-\mathbf{f}|1s\rangle\langle\mathbf{f}|1s\rangle=\left(\frac{4\eta^2}{q^2+4\eta^2}\right)^2, \tag{63}$$
$$G(\boldsymbol{\kappa})=\frac{\gamma_0E_a-\boldsymbol{\gamma}\cdot\boldsymbol{\kappa}+m}{E_a^2-\boldsymbol{\kappa}^2-m^2}\approx\frac{\gamma_0E_a+\boldsymbol{\gamma}\cdot\mathbf{p}_2+m}{-4m\varepsilon_2}, \tag{64}$$
$$E_a=m-\varepsilon_2,$$
$$G(-\mathbf{k})=\frac{\gamma_0E_b+\boldsymbol{\gamma}\cdot\mathbf{k}+m}{E_b^2-\mathbf{k}^2-m^2}\approx\frac{\gamma_0E_b+\boldsymbol{\gamma}\cdot\mathbf{k}+m}{-2m\omega},$$
$$E_b=m-\omega. \tag{65}$$

In deriving (61) and (62), we put $q=0$ everywhere except in the factor $\phi(\mathbf{q})$. The cross section of the process is proportional to $\phi^2(\mathbf{q})$. This factor is responsible for the special status of the central region, since $\phi^2(\mathbf{q})\sim\eta$ there, while outside this region $\phi^2(\mathbf{q})\sim(\alpha Z)^8$. Note that for $q\gg\eta$ Eqs. (61) and (62) become invalid, but these regions still make but a small contribution to the cross section (except in the very narrow edge region $p_2\sim\eta$ considered in Sec. 2).

Primed letters denote diagrams in Fig. 1 with initial states interchanged. If in diagrams a and b (a′ and b′) we interchange the states ψ_{-p_1} and ψ_{p_2}, we obtain diagrams c′ and d′ (c and d), respectively. The total amplitude of double ionization with one-photon annihilation that allows for the contributions of all eight Feynman diagrams can be written

$$M^{++}=M(-p_1,p_2)-M(p_2,-p_1).$$
$$M(-p_1,p_2)=M_a+M_b-M_{a'}-M_{b'},$$
$$M(p_2,-p_1)=M_{c'}+M_{d'}-M_c-M_d. \tag{66}$$

We denote the amplitude of the process in the central region by M^{++}_{central}. It can be obtained from (66) if for M_a,M_b, etc. we take expressions (61), (62), and the like. After an involved procedure of calculating traces, the square of the amplitude (66) summed and averaged over the particle polarizations acquires the surprisingly compact form

$$\overline{|M^{++}_{\text{central}}|^2}=\frac{\phi^2(\mathbf{q})W(\omega)}{(2m)^4E_1E_2}, \tag{67}$$

where

$$W(\omega)=\left(\frac{\varepsilon_2+\omega_0}{\varepsilon_2\omega_0}\right)^2\left\{E_1E_2-m^2-p_{1n}p_{2n} +\left(\frac{m\omega}{\varepsilon_2\omega_0}\right)^2(E_1E_2+m^2+p_{1n}p_{2n}-2\varepsilon_2\omega_0)\right\}, \tag{68}$$

$$p_{1n}=\mathbf{p}_1\cdot\mathbf{n}=\frac{E_1^2-E_2^2+\omega^2}{2\omega},$$
$$p_{2n}=\mathbf{p}_2\cdot\mathbf{n}=\frac{E_1^2-E_2^2-\omega^2}{2\omega},$$
$$\mathbf{n}=\frac{\mathbf{k}}{k},\quad\varepsilon_2+\omega=\omega_0. \tag{69}$$

Equation (68) can be further simplified[7)] by introducing the dimensionless quantities $x=\omega/\omega_0$ and $\gamma=m/\omega_0$:

$$W(\omega)\equiv W(x)=\gamma\left(\frac{2-x}{1-x}\right)^2\left\{2(x-\gamma)+\left[1-\left(\frac{\gamma x}{1-x}\right)^2\right] \times\left[4\frac{1-x}{x}-\gamma\left(\frac{2-x}{x}\right)^2\right]\right\}. \tag{70}$$

Substituting the expression (67) for $\overline{|M^{++}|^2}$ into (44) we obtain a formula for the differential cross section of double ionization with one-photon annihilation in the central region of the spectrum:

$$d\sigma_c^{++}=\frac{\alpha r_0^2}{16\pi^2}\phi^2(\mathbf{q})W(x)\frac{d\Gamma}{m^2p_1},$$

$$d\Gamma=\frac{1}{E_2\omega}d\mathbf{p}_2 d\mathbf{k}\ \delta(E_2+\omega-E_0),\quad E_0=E_1+2m. \quad (71)$$

Before we perform an integration involving the energy delta function, it is convenient to transform from the variable $\mathbf{p}_2$ to the variable $\mathbf{q}=\mathbf{p}_2+\mathbf{k}-\mathbf{p}_1$. For fixed Ω_k (direction of ejection of the photon) and $\mathbf{q}$, the electron energy E_2 depends on ω. After we have integrated with respect to ω, we must therefore make the substitution

$$d\omega\ \delta(E_2(\omega)+\omega-E_0)\to\left|\frac{\partial E_2}{\partial\omega}+1\right|^{-1}, \quad (72)$$

where the derivative $\partial E_2/\partial\omega$ is evaluated at a value of ω that is the root of the equation

$$E_2(\omega)+\omega-E_0=0. \quad (73)$$

Bearing in mind that only small $q\sim\eta$ (the central region) are important in integrals over $\mathbf{q}$, we find that

$$\omega=\omega(t_k)=\frac{2m\omega_0}{E_0-p_1t_k},\quad t_k=\cos\ \theta_k, \quad (74)$$

$$d\Gamma=\frac{\omega^2}{2m\omega_0}d\mathbf{q}d\Omega_k. \quad (75)$$

Substituting (75) into (71) and integrating over $\mathbf{q}$, we obtain for the angular distribution of photons

$$\frac{d\sigma_c^{++}}{d\Omega_k}=\frac{\sigma_0Z^3}{32\pi}\frac{\omega_0}{p_1}x^2W(x),\quad x=x(\theta_k). \quad (76)$$

Equation (74) makes it possible to relate dt_k to dx,

$$dt_k=\frac{2m}{p_1}\frac{dx}{x^2}, \quad (77)$$

and to go from the angular distribution (76) to the energy distribution

$$\frac{d\sigma_c^{++}}{dx}=\frac{\sigma_0Z^3}{8}\frac{m}{\varepsilon_1}W(x),\quad \varepsilon_1=E_1-m. \quad (78)$$

On the basis of (74), we conclude that $x_1\leqslant x\leqslant x_2$, where

$$x_1=\frac{2m}{E_0+p_1},\quad x_2=\frac{2m}{E_0-p_1}. \quad (79)$$

Integrating (78) with respect to x, we find the contribution of the entire central region to the cross section of double ionization with one-photon annihilation:

$$\sigma_c^{++}=\sigma_0Z^3f(E_1),\quad f(E_1)=\frac{1}{8}\frac{m}{\varepsilon_1}\int_{x_1}^{x_2}W(x)dx. \quad (80)$$

Since in the region where $\mathbf{q}=0$ the angles θ_k and θ_2 and the photon energy ω are related by

$$t_k=\frac{p_1-p_2t_2}{\omega}\quad\left(t_2=\cos\ \theta_2=\frac{\mathbf{p}_2\cdot\mathbf{p}_1}{p_2p_1}\right), \quad (81)$$

which follows from (57), we can make the independent variable the angle θ_2 of electron ejection and obtain the electron angular distribution. To this end we express ω and $d\omega$ in terms of t_2 and dt_2:

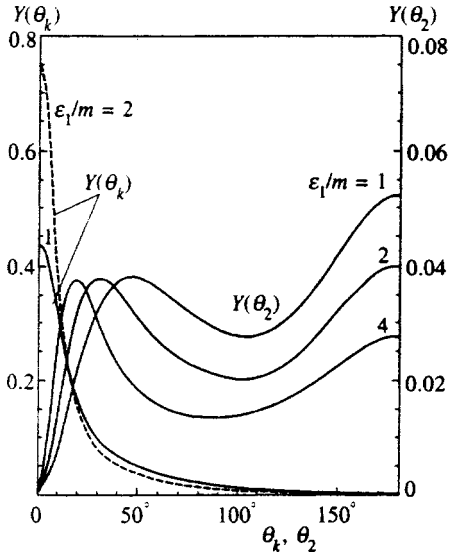

FIG. 4. Angular distributions of photons $Y(\theta_k)=A^{-1}d\sigma_c^{++}/d\Omega_k$ and electrons $Y(\theta_2)=A^{-1}d\sigma_c^{++}/d\Omega_2$ belonging to the central region of the energy spectrum, with $A=\sigma_0Z^3$.

$$\omega=\omega(t_2)=E_0g\pm\sqrt{(E_0g)^2-h^2},$$

$$d\omega=\frac{p_1p_2^2dt_2}{E_0p_2-p_1E_2t_2}, \quad (82)$$

$$g=\frac{\omega_0^2-(p_1t_2)^2}{E_0^2-(p_1t_2)^2},\quad h^2=\frac{\omega_0^4-(p_0p_1t_2)^2}{E_0^2-(p_1t_2)^2},$$

$$p_0^2=E_0^2-m^2. \quad (83)$$

The plus sign in (82) is taken for $\pi/2\leqslant\theta_2\leqslant\pi$, and the minus sign for $0\leqslant\theta_2\leqslant\pi/2$. Substituting (82) into (78), we obtain the angular distribution of the high-energy electrons,

$$\frac{d\sigma_c^{++}}{d\Omega_2}=\frac{\sigma_0Z^3}{16\pi}\frac{m}{p_1}\frac{p_2^2W(x)}{E_0p_2-p_1E_2t_2}. \quad (84)$$

4. DISCUSSION

The angular distributions of photons and electrons belonging to the central region of the energy spectrum are depicted in Fig. 4. We used (76) to construct the photon angular function $Y(\theta_k)=A^{-1}d\sigma_c^{++}/d\Omega_k$, with $A=\sigma_0Z^3$, for two values of ε_1 (the positron's kinetic energy)—m and $2m$. Equation (84) was used to construct the electron angular function $Y(\theta_2)=A^{-1}d\sigma_c^{++}/d\Omega_2$ for three values of ε_1—m, $2m$, and $4m$. The values of ε_1 in units of m label the curves. Comparing the photon and electron angular spectra, we notice that photons preferentially emerge in the forward direction, while electrons are more uniformly distributed in angle. Recall that in the edge region, the probability of ''forward'' photon emission is zero, and electrons are emitted isotropically.

Photon energy spectra (78) are depicted in Fig. 5 for three values of ε_1. Each curve lies within the interval $x_1(\varepsilon_1)\leqslant x\leqslant x_2(\varepsilon_1)$. To establish the behavior of $d\sigma_c^{++}/dx$

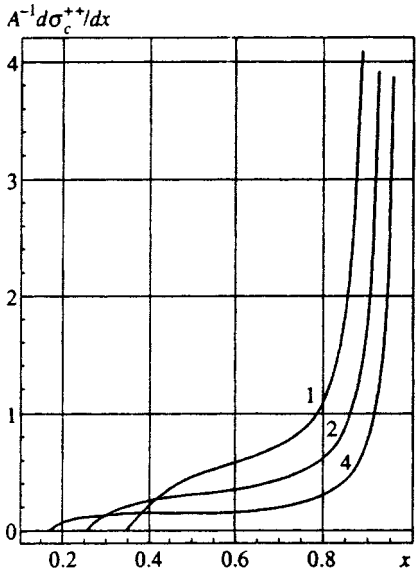

FIG. 5. Energy distribution of photons in the central region of the spectrum; $A=\sigma_0 Z^3$. The curves are labeled with the positron energy in units of the electron mass.

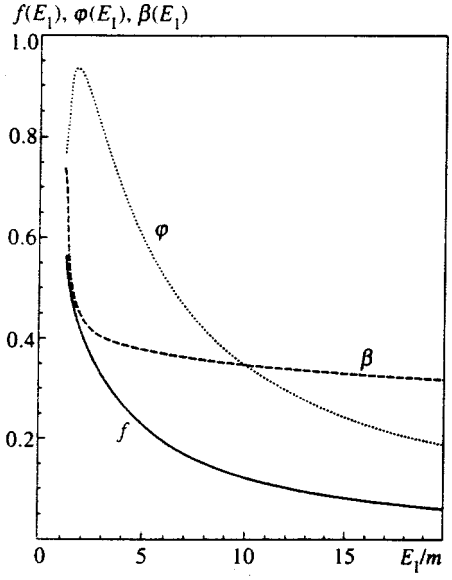

FIG. 6. Dependence on the positron energy E_1 of the total cross sections of double and single ionization and their ratio: $f(E_1)=\sigma_c^{++}/\sigma_0 Z^3$, $\varphi(E_1)=\sigma^{+}/\sigma_0 Z^5$, and $\beta(E_1)=Z^2\sigma_c^{++}/\sigma^{+}$.

outside this interval, we must integrate (71) over photon and electron ejection angles. It can be shown that this method yields

$$\frac{d\sigma_c^{++}}{dx}=A\frac{m}{8\varepsilon_1}W(x)v(x),$$

where the factor $v(x)$ is unity for $x_1<x<x_2$ and decreases by two to three orders of magnitude as we move a distance $\Delta x\sim\alpha Z$ away from this interval.

The total double ionization cross section with one-photon annihilation is the sum of the contribution of the edge and central regions of the energy spectrum (Eqs. (55) and (80)):

$$\sigma^{++}=\sigma_e^{++}+\sigma_c^{++}=\sigma_0 Z^3[B\varphi(E_1)+f(E_1)]. \tag{85}$$

The ratio of the double and single ionization cross sections (see Eqs. (85) and (52)) is

$$R=\frac{\sigma^{++}}{\sigma^{+}}=\frac{B+\beta(E_1)}{Z^2}, \tag{86}$$

$$\beta(E_1)=\frac{f(E_1)}{\varphi(E_1)}. \tag{87}$$

Figure 6 depicts the functions $f(E_1)$, $\varphi(E_1)$, and $\beta(E_1)$. The functions $f(E_1)$ and $\beta(E_1)$ decrease monotonically as E_1 increases, while $\varphi(E_1)$ peaks at $E_1\approx 2m$. We now study the behavior of these functions at low and high energies ε_1 analytically.

For $I\ll\varepsilon_1\ll m$ (the condition $\eta\ll p_1\ll m$ must be met) we have

$$f(E_1)\simeq\frac{8}{27}\frac{m}{p_1},\quad \varphi(E_1)\simeq\frac{4}{3}\frac{p_1}{m},$$

$$\beta(E_1)\simeq\frac{2}{9}\left(\frac{m}{p_1}\right)^2=\frac{1}{9}\frac{m}{\varepsilon_1}. \tag{88}$$

Interestingly, for Z and ε_1 small, R can be greater than unity, i.e., double ionization is more likely to occur than single ionization (for He the ionization potential $I\sim 10^{-4}\,m$, and at $\varepsilon_1=10^{-2}\,m$ the cross-section ratio R is approximately 2.5). This result can be explained entirely by the special role of the central region of the energy spectrum. The emission of two particles (a photon and an electron) allows the momentum transfer to the nucleus in process (4) to be small, which leads to an increase in the cross section. On the other hand, in process (2) a decrease in the positron momentum p_1 leads to an increase in the momentum q transferred to the nucleus. As a result, the one-photon annihilation cross section in the nonrelativistic region decreases together with p_1 ($\sim p_1/m$), while the double ionization cross section increases as p_1 decreases ($\sim m/p_1$).

For $E_1\gg m$ we can obtain the following expansions in reciprocal powers of E_1:

$$f(E_1)\simeq\frac{m}{E_1}\left[1+\frac{m}{4E_1}\left(11\ln\frac{E_1}{m}+6\ln 2-23+\frac{1}{6}\right)\right],$$

$$\varphi(E_1)\simeq\frac{4m}{E_1}\left(1-\frac{4}{3}\frac{m}{E_1}\right),$$

$$\beta(E_1)\simeq\frac{1}{4}\left[1+\frac{1}{4}\frac{m}{E_1}\left(11\ln\frac{E_1}{m}+6\ln 2-\frac{35}{2}\right)\right],$$

$$R(E_1)\xrightarrow[E_1\to\infty]{}\frac{1}{Z^2}\left(B+\frac{1}{4}\right)=\frac{0.340}{Z^2}. \tag{89}$$

We see that the function $\beta(E_1)$ approaches its asymptotic value very slowly: even at $E_1/m=100$ the value of $\beta(E_1)$ differs from its asymptotic value of 1/4 by 10%.

Although we cannot hope to obtain accurate results for large values of Z, let us estimate the total cross section of double ionization with one-photon annihilation at $Z=82$ and $E_1=2m$ and compare the result with the corresponding cross

section of double ionization of the K shell with radiationless annihilation (3) (see Refs. 8 and 9). Since for high-Z atoms the normalization constants N_{p_2} and N_{-p_1} of the electron and positron wave functions are appreciably different from unity, we will keep them in the formulas for the cross section (as was also done in Refs. 8 and 9). Then Eq. (85) becomes

$$\sigma^{++}=N_{p_2}^2 N_{-p_1}^2 \sigma_0 Z^3[B\varphi(E_1)+f(E_1)],$$

$$N_{-p_1}^2=\frac{2\pi\xi_1}{\exp(2\pi\xi_1)-1}, \quad \xi_1=\alpha Z\frac{E_1}{p_1}, \tag{90}$$

where $N_{p_2}^2$ has been defined in (36).

A calculation with this formula yields

$$\sigma^{++}(Z=82,\ E_1=2m)=4.3\times 10^{-29}\ \text{cm}^2.$$

For process (3) under the same conditions we have (see Refs. 8 and 9)[8)]

$$\sigma\simeq 1.8\times 10^{-29}\ \text{cm}^2.$$

Thus, despite the additional small factor α related to the emission of a photon, the cross section of double ionization with one-photon annihilation proves to be greater than the cross section of radiationless annihilation, even for high-Z elements. As Z decreases, this predominance grows, reaching several orders of magnitude for low-Z elements. There are two factors that explain such behavior: the photon exchanged by the electrons is much less virtual in process (4) than in process (3), and the emission of a photon in process (4) significantly reduces the momentum transferred to the nucleus.

We would like to express our gratitude to V. G. Gorshkov and L. N. Labzovskiĭ for their interest in our work and for fruitful discussions.

*)E-mail: Mikhailo@thd.pnpi.spb.ru

1)In this paper we use the relativistic system of units, in which $\hbar=c=1$.

2)The Greek letter α is used to denote both the initial electron state and the fine structure constant, but it is always clear from the context in what sense it is used.

3)The factor $4\pi\alpha$ from the photon propagator and the factor $\sqrt{4\pi\alpha/2\omega}$ from the photon wave function are incorporated into the expression for the cross section.

4)Although a factor αZ precedes the integral in (24), the integral proper is $\sim 1/\alpha Z$, which can be verified by substituting the free Green's function (22) for G_C.

5)The normalization constants $1/\sqrt{2E}$ from the electron wave functions are incorporated into the corresponding bispinors u_p, which are normalized by the condition $\bar{u}_p u_p=m/E$.

6)The expression for the one-photon annihilation cross section in Ref. 18 can be reduced to (49) and (50).

7)We write formula (68) because by virtue of exchange invariance, we can use (68) to derive a formula for the relativistic double photoeffect.

8)The cross section given in Ref. 9 must be doubled.

[1] T. Nagatomo, Y. Nakayama, K. Morimoto, and S. Shimizu, Phys. Rev. Lett. **32**, 1158 (1974).
[2] V. G. Gorshkov, A. I. Mikhaĭlov, and S. G. Sherman, Zh. Éksp. Teor. Fiz. **72**, 32 (1977) [Sov. Phys. JETP **45**, 17 (1977)].
[3] W. R. Johnson, D. J. Buss, and C. O. Carroll, Phys. Rev. **135**, A1232 (1964); W. R. Johnson, Phys. Rev. **159**, 61 (1967).
[4] H. Mazaki, M. Nishi, and S. Shimizu, Phys. Rev. **171**, 408 (1968).
[5] K. W. Broda and W. R. Johnson, Phys. Rev. A **6**, 1693 (1972).
[6] H. S. W. Massey and E. H. S. Burhop, Proc. R. Soc. London, Ser. A **167**, 53 (1938).
[7] S. Shimizu, T. Mukoyama, and Y. Nakayama, Phys. Rev. **173**, 405 (1968).
[8] A. I. Mikhailov and S. G. Porsev, J. Phys. B **25**, 1097 (1992).
[9] A. I. Mikhaĭlov and S. G. Porsev, Zh. Éksp. Teor. Fiz. **105**, 828 (1994) [JETP **78**, 441 (1994)].
[10] E. G. Drukarev and F. F. Karpeshin, J. Phys. B **9**, 399 (1976).
[11] F. W. Byron and C. J. Joachain, Phys. Rev. **164**, 1 (1967).
[12] T. Aberg, Phys. Rev. A **2**, 1726 (1970).
[13] M. Ya. Amusia, E. G. Drukarev, V. G. Gorshkov, and M. P. Kazachkov, J. Phys. B **18**, 1248 (1975).
[14] V. G. Gorshkov, A. I. Mikhaĭlov, and V. S. Polikanov, Nucl. Phys. **55**, 273 (1964).
[15] V. G. Gorshkov, Zh. Éksp. Teor. Fiz. **47**, 352 (1964) [Sov. Phys. JETP **20**, 234 (1965)].
[16] V. G. Gorshkov and V. S. Polikanov, JETP Lett. **9**, 279 (1969).
[17] M. Ya. Amus'ya and A. I. Mikhaĭlov, Zh. Éksp. Teor. Fiz. **111**, 862 (1997) [JETP **84**, 474 (1997)].
[18] A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics*, Wiley, New York (1974).

Translated by Eugene Yankovsky

Gain of a compressed DT fuel statistically ignited by a multispark assembly

A. Caruso and C. Strangio

Associazione Euratom—ENEA, CRE Frascati, 00044 Frascati (Roma), Italia
(Submitted 10 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 805–815 (March 1998)

Most of the current inertial confinement fusion (ICF) schemes are based on the ignition of a high-density DT fuel by a single, high-temperature spherical hot spot (the spark). The spark is self-generated by the implosion process, which is used to bring the fuel to high density. To start ignition the spark has to be dimensioned in such a way that the ion temperature would be greater than 5–7 keV, and that the spark radius would be greater than the α-particle range. A spark with these features is indicated as supercritical. In the scheme based on self-generated spark, ignition can fail to occur when the produced spark strongly deviates from spherical shape, which can make all the surface losses highly relevant. High deformation, or even spark splitting, can occur due to the amplification of initial deviations from spherical shape by hydrodynamic instabilities (or by secular growth) during the implosion process. In principle, ignition can be recovered if the implosion is designed in such a way as to make supercritical at least one of the portions of hot fuel which are produced in this way near stagnation. As a general trend, more compressed final assemblies are required. In this paper we present fuel gain calculations (Gain=Thermonuclear energy/Energy in the compressed fuel) for DT assemblies ignited at the end of an implosion process by a supercritical spark statistically created within a cluster of many subcritical sparks. It is assigned the total number of sparks and the probability of having at least one of them supercritical. As a function of these quantities we calculate, in the framework of an isobaric model, the average thermal energy associated with the spark assembly. The same model is also used to evaluate, by statistical arguments, the areal mass, the burn fraction, and the system's total fuel gain. It is found that the energy distribution function of the sparks is influenced only by a single global parameter, in which the assigned ignition probability and the number of sparks are also represented. Compared to the single central-spark approach, being the final states with allowed inner turbulence, the multispark scheme is characterized by relaxed initial symmetry requirements. For multispark systems we can realistically consider the achievement of fuel gains comparable or greater than those typical of the single-spark approach, when evaluated for currently accepted spark convergence ratios. With regard to the single spark case, higher cold fuel densities are needed, as expected (typically $2\times-3\times$, for the same gain, depending on the energy distribution function). © *1998 American Institute of Physics.* [S1063-7761(98)00303-5]

1. INTRODUCTION

The conventional ICF scheme is based on fuel triggering by a single, self-generated, hot spot (the ignition spark) at the center of a compressed, cold DT fuel assembly. Hot spot ignition is usually preferred to volume ignition, because in the cases of interest for the energetic application of ICF, it provides higher fuel gains at substantially smaller extreme parameters for the compressed fuel assembly (pressure, density, density×fuel size). This is basically due to the circumstance that to get high fuel gains in uniform systems it is necessary to ignite the fuel at comparatively low temperatures. This requires radiation trapping in order to make radiative losses smaller than the thermonuclear energy released to the fuel.

In the central-spark approach, ignition is initiated when the ion temperature in the spark, T_2 ($\approx T_e$, the electron temperature), is greater than 5–7 keV and the spark radius $R_{\rm spark}$ is greater than the range of the α-particle produced in the DT reaction. Under such conditions, in a stagnant spark the thermonuclear energy deposited inside the spark exceeds the radiative and the electronic conductivity losses, and self-heating occurs. A spark with these features is called a supercritical spark.

The burn propagates from the spark to the surrounding fuel, bringing it to temperatures in the range 40–100 keV. Fuel combustion is quenched by hydrodynamic expansion in a time $t_{\rm exp}\approx R_{\rm fuel}/4c_s$, where $R_{\rm fuel}$ is the total radius of the compressed fuel, and c_s is the sound velocity at the burning temperature (much greater than the ignition temperature). The governing number for this process is clearly $\phi=\langle\sigma\nu\rangle_{\rm DT}(\rho/M_{\rm DT})t_{\rm exp}$, where $\langle\sigma\nu\rangle_{\rm DT}$ is the usual Maxwellian-averaged product of the DT reaction cross section times the ion velocity, ρ is the fuel mass density, and $M_{\rm DT}$ is the ionic average mass. The maximum of ϕ, which occurs at a temperature $T\approx 40.6$ keV, is $\phi_{\rm max}\approx\rho R_{\rm fuel}/4.7$ (CGS units are used). It is therefore clear that values of $\rho R_{\rm fuel}$ of several g/cm^2 are needed to obtain high values for

the fractional fuel burn, $f_{\rm burn}$. A widely used formula for $f_{\rm burn}$ is given in the next section. Typically, for $\rho_{\rm fuel}=0.3$, we have $f_{\rm fuel}=0.3$. The value $R_{\rm spark}$ of the spark radius is typically assumed to be of the order of 1/40 of the initial fuel capsule radius $R_{\rm target}$ (i.e., the convergence ratio is $C_{\rm spark}=R_{\rm target}/R_{\rm spark}=40$). Such a large value of $C_{\rm spark}$ is used to maintain the ablation pressure, which is needed to drive the fuel up to the required high implosion velocity, within technically available values (at laboratory energy releases).

The controlled formation of the ignition spark is the main objective of this scheme. Actually, because of the required high convergence ratio, this target design seems sensitive to nonuniformities in the ablation pressure, to irregularities in the target structure, and to hydrodynamic instabilities. For instance, it has been shown[1] how due to the nonuniformities in the initial energy deposition (percent sized low modes, mode number l up 10–20), the thermonuclear gain can drop abruptly to zero. The reason for ignition failure (no gain) is the formation of a highly deformed spark, in which the surface heat losses are too large due to an unfavourable surface-to-volume ratio. Numerical modeling shows that because of hydrodynamic instabilities, the final spark, at stagnation, can be so distorted as to become nearly split into l substructures by the well-known mushroom-shaped jets.[2,3] Preliminary studies[3] of $3D$ perturbed spherical stagnation show that saturation occurs when the perturbation amplitude is of the same order of wavelength. These studies were carried out for the single mode at low harmonic indices (l and m up 6).

It should be noted that low modes (mode indices 5–10) can, in principle, generate a large number of substructures. If the prevailing grain size is of the order of R_s/n (R_s is the radius within which the sparks are enclosed, and n is a typical mode number), approximately n^2 fragments saturated at R_s/n can be allocated in a single spherical layer, while n^3 can be allocated in a volume. At any rate, unbalanced direct irradiation with multibeam modern laser installations may already produce a large number of substructures [$n \propto (\text{number of beams})^{1/2}$].

To the best of our knowledge, multimode $3D$ studies with realistic implosion dynamics and burn are still lacking, as is any acceptable theoretical description of the final turbulence, since it may result for given initial conditions. However, for conventional target designs, it seems reasonable to assume that, unless very high spherical symmetry is achieved in the initial stages of the implosion (by high-quality irradiation and target finish), spark splitting and ignition failure at stagnation can occur.

Let us assume that the initial requirements for symmetry are deliberately relaxed, so that a final turbulent condition is allowed. In this context, an interesting question which must be answered has to do with the possibility of recovering ignition and high gain for this final assembly. Clearly, to find the answer, the relevant key parameter to be determined is the investment in energy required to make supercritical, with assigned probability p, at least one of the hot fuel portions in the turbulent mix (in the center or elsewhere). Once this energy is determined, the assembly parameters needed for high gain can be also found. This information can then be used to identify the general features of implosions which may lead to such final configurations. Since an ignition probability p is introduced, the question may arise about the scheme's usefulness for energy applications, since now a finite ignition failure probability $(1-p)$ is allowed. In energy applications, however, the concept of ignition failure can be introduced without harm if the implosion is designed in such way as to make the ignition probability p sufficiently high. For this kind of design, the well-known reactor loop condition, which relates the target gain to the driver efficiency $\eta_{\rm driver}$, say $\eta_{\rm driver}\times\text{target gain}\approx 10$, needs only be changed to $p\times\eta_{\rm driver}\times\text{target gain}\approx 10$. Ignition failure can arise under reactor conditions for a number of reasons, different from target design (e.g., misalignment, target structural imperfections, lack of reproducibility of the driver pulse, etc.), so that the probability of ignition p should be introduced in any case.

2. A MODEL FOR GAIN CALCULATIONS

The following model is adopted to answer the questions discussed in the previous section.

a) Uniform pressure (P) is assumed at stagnation for the final fuel assembly. Ignition and fuel gain calculations with this assumption have been made by a number authors[4] since 1976.

b) The fuel is divided into two phases, hot and cold. The hot fuel is assumed to be split into N parts at the same temperature. Having the same temperature and pressure, the hot sparks have also the same density ($\rho_{\rm spark}$). The pressure in the cold fuel is assumed to be

$$P=\alpha P_F, \tag{1}$$

where α is the factor by which this part of the fuel is not degenerate, and P_F is the Fermi pressure.[5]

c) The compressed DT fuel assembly is assumed to be a perfect gas. Thus the energy per unit volume is $3P/2$ everywhere, regardless of the degeneracy degree.[6]

d) Lacking better descriptions, simple test-distribution-functions for the turbulence spectrum in energy are adopted (size $\propto$ energy$^{1/3}$). Various dependences are tried for trend sampling. The final gain is computed for the distribution which gives the smaller gains in the explored set.

e) To ignite the assembly, we assume[5] that there must be at least a spark with energy (E) such that

$$E>E_{\rm thr}\approx\frac{91}{\rho_{\rm spark}^2}\,\frac{T_{\rm spark}}{7}\ \text{MJ}, \tag{2}$$

where $T_{\rm spark}$ is in keV, and $\rho_{\rm spark}$ is in g/cm^3. The inequality (2) holds for a set of spheres with a radius $R\geqslant 0.3/\rho_{\rm spark}$ cm and the same density ($\rho_{\rm spark}$) and temperature ($T_{\rm spark}$). To obtain ignition a $T_{\rm spark}=5$ keV is usually considered sufficient. In all the following numerical examples $T_{\rm spark}=7$ keV will be assumed. Unless a critical spark ($E>E_{\rm thr}$) has been created in the assembly, the subcritical sparks ($E<E_{\rm thr}$) are ultimately cooled by thermal conduction, radiative losses or the general assembly expansion.

f) The probability p to have within N sparks at least one with $E\geqslant E_{\rm thr}$ is assigned. Once p and the energy distribution

function for sparks are assigned, the total spark energy ($E_{\rm hot}$) can be calculated in terms of $E_{\rm thr}$, N, and p.

g) The gain calculations require an estimate for the fraction ($f_{\rm burn}$) of mass burned by thermonuclear reactions before the general fuel disassembling occurs. For DT fuel assemblies with spherical symmetry we apply a widely used formula[5] for $f_{\rm burn}$

$$f_{\rm burn}=\frac{\int_0^{R_{\rm fuel}}\rho\ dr}{\int_0^{f_{\rm fuel}}\rho\ dr+7}\ [\text{CGS units}], \tag{3}$$

where the integral is taken from the symmetry center to the radius of the total fuel assembly $R_{\rm fuel}$. In our case we use a variant of Eq. (3), in which the integral is statistically evaluated.

3. THE PROBABILITY OF IGNITION

We assume that, as a result of the implosion, within the isobaric fuel N hot sparks have been formed and distributed in energy E according to $dq=f(E,E_0)dE$, where E_0 is a parameter with dimensions of energy, and dq is the probability for a spark to be between E and $E+dE$. Let $f(E,E_0)=\omega(E/E_0)/E_0$, where ω is a dimensionless function. The probability p to have at least one supercritical spark can easily be calculated in terms of ω as

$$p=1-\left[\int_0^{E_{\rm thr}/E_0}\omega(x)dx\right]^N. \tag{4}$$

This equation can be solved for E_0 as a function of $E_{\rm thr}$, p, and N. After this value of E_0 is inserted in

$$E_{\rm hot}=N\int_0^\infty Ef(E,E_0)dE=NE_0\int_0^\infty x\omega(x)dx, \tag{5}$$

an estimate results for the total energy, $E_{\rm hot}$, which is required in order to obtain, with assigned probability, at least one above-threshold spark. Clearly,

$$E_{\rm hot}=\gamma E_{\rm thr},\quad \gamma=Nh[(1-p)^{1/N}]. \tag{6}$$

where the function h is calculated from Eqs. (4) and (5). We note that, when N is large, the expected p-dependence for γ can be slow.

Since $E\propto R^3$, where R is the radius of the spark, to use ω is tantamount to using the space turbulence spectrum. The distribution functions in E can then be seen in terms of spark radii by introducing a space scale R_0 through $E=(R/R_0)^3E_0$. The function $f(E,E_0)$ will then correspond to the function

$$g(R,R_0)=R_0f(E,E_0)\ \frac{dE}{dR}. \tag{7}$$

We now consider a few examples of the distribution functions in E, which are chosen to illustrate some relevant points in view of their mathematical simplicity. We adopt three functions, a square-box, an exponential function, and a simple combination of power-laws in energy, which qualitatively represent quite different energy (radii) distributions. The square-box dependence is

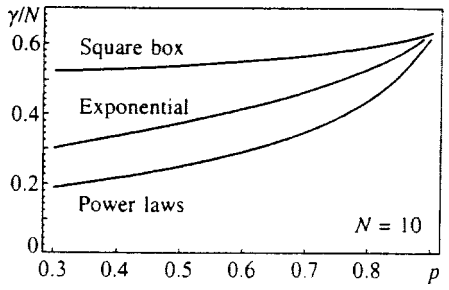

FIG. 1. Dependence of the parameter γ/N on p. The calculations are presented for the distribution functions given by Eqs. (8), (9), (11), and (13).

$$f(E,E_0)=\frac{1}{E_0}\quad \text{for}\quad E\leqslant E_0, \tag{8}$$

$$f(E,E_0)=0\quad \text{for}\quad E>E_0, \tag{9}$$

and the corresponding normalized energy is

$$\gamma=\frac{N}{2(1-p)^{1/N}}. \tag{10}$$

The exponential dependence

$$f(E,E_0)=\frac{\exp(-E/E_0)}{E_0} \tag{11}$$

corresponds to

$$\gamma=-\frac{N}{\ln[1-(1-p)^{1/N}]}. \tag{12}$$

and for the power-law case we have

$$f(E,E_0)=\frac{2}{3}\left[1-\left(\frac{E}{E_0}\right)^{1/3}\right]\left(\frac{E}{E_0}\right)^{-2/3}\frac{1}{E_0}\quad \text{for}\quad E\leqslant E_0,$$

$$f(E,E_0)=0\quad \text{for}\quad E>E_0, \tag{13}$$

$$\gamma=\frac{N}{10[1-\sqrt{1-(1-p)^{1/N}}]^3}. \tag{14}$$

The dimensionless hot energy γ (in units of N) is represented as a function of p in Fig. 1, for $N=10$. It is worth noting how already for this value of N both the ''exponential'' and the ''power-law'' functions give better results than the ''square-box'' function. As is obvious from Eqs. (10), (12), and (14), the dependence on p becomes less important for increasing N. Note also that $E_{\rm hot}$ is substantially smaller than $NE_{\rm thr}$.

4. FUEL ASSEMBLY DIMENSIONS

Let us assume that an energy $E_{\rm fuel}$ has been transferred to the compressed fuel. The fraction

$$\nu=E_{\rm hot}/E_{\rm fuel} \tag{15}$$

is used to parametrize our calculations. If the parameter ν is given, by using Eq. (6) we can write the allowed $E_{\rm thr}$ as follows:

$$E_{\rm thr}=\frac{\nu}{\gamma}E_{\rm fuel}, \tag{16}$$

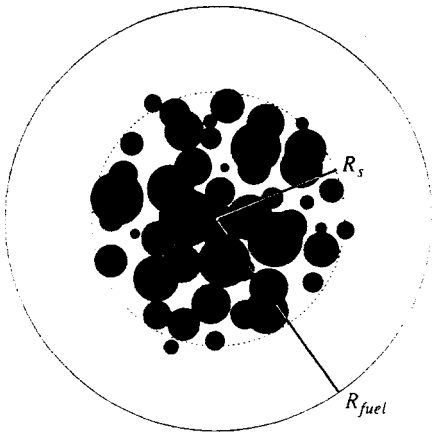

FIG. 2. The numerical examples reported in the paper refer to sparks randomly distributed within a sphere of radius R_s smaller or equal to the fuel radius $R_{\rm fuel}$.

so that the allowed hot fuel density, $\rho_{\rm spark}$, can be computed from Eq. (2). Since the temperature is given (say, $T_{\rm spark}=7$ keV), we can find the hot fuel pressure. Because the system is isobaric, the pressure (P) is the same everywhere in the fuel. Equation (1) can then be used to determine the cold fuel density, $\rho_{\rm cold}$. The energy density throughout the fuel is $3P/2$, so that when we assume the fluid is entirely enclosed within a sphere of radius $R_{\rm fuel}$ (see Fig. 2), the system size is

$$R_{\rm fuel}=\left(\frac{E_{\rm fuel}}{2\pi P}\right)^{1/3}. \tag{17}$$

Similar considerations provide the total fuel mass as a sum of the cold and hot part contributions:

$$M_{\rm fuel}=\frac{2}{3}\left[\rho_{\rm cold}(1-\nu)+\rho_{\rm spark}\nu\right]\frac{E_{\rm fuel}}{P}. \tag{18}$$

Since the critical spark volume is $E_{\rm thr}/(3P/2)$, from the previous formulas the critical spark radius is found to be

$$R_{\rm thr}=(\nu/\gamma)^{1/3}R_{\rm fuel}. \tag{19}$$

We note that the distribution function in energy (or R) enters in these estimates only through the single dimensionless parameter γ introduced in the initial $E_{\rm thr}$ evaluation [through Eq. (16)].

The thermonuclear gain for this fuel assembly is then

$$G=\varepsilon_{\rm tn}\frac{M_{\rm fuel}}{E_{\rm fuel}}f_{\rm burn}, \tag{20}$$

where $\varepsilon_{\rm tn}$ is the thermonuclear energy released by burning the unit mass, and $f_{\rm burn}$ is the fraction of burned mass, now to be evaluated by a statistical version of Eq. (3).

5. THE BURN FRACTION $f_{\rm burn}$

For the situation analyzed in this paper, the integral appearing in the Eq. (3) for $f_{\rm burn}$ is

$$\int_0^{R_{\rm fuel}}\rho\ dr=\rho_{\rm spark}L_{\rm spark}+\rho_{\rm cold}(R_{\rm fuel}-L_{\rm spark}), \tag{21}$$

where $L_{\rm spark}$ is obtained by adding the portions of the radius $R_{\rm fuel}$ embedded in the randomly distributed sparks (e.g., the dashed segments in Fig. 2).

Let us assume that a line starting from the system's center meets a spherical spark of radius $R(\propto E^{1,3})$. It is easily seen that, on the average, the line segment within the sphere is $(4/3)R(E)$. If, along the radial coordinate r in the fuel, the space-energy distribution is then assumed to be $dN=4\pi r^2 n(r)f(E,E_0)dr\ dE$, we obtain

$$L_{\rm spark}=\int_0^{R_{\rm fuel}}dr\int_0^{\infty}dE\ n(r)f(E,E_0)\pi R^2(E)\cdot\frac{4}{3}R(E). \tag{22}$$

However, $(4/3)\pi R^3=2E/(3P)$, where P is the constant presure of the system, so that through Eqs. (5), (17), and (19) it follows that

$$L_{\rm spark}=R_{\rm fuel}\nu\left(\frac{R_{\rm fuel}}{R_{\rm turb}}\right)^2,\quad R_{\rm turb}^2=N\left[\frac{4\pi}{3}\int_0^{R_{\rm fuel}}n(r)dr\right]^{-1}, \tag{23}$$

without any additional assumption on the distribution function $f(E,E_0)$. As we have seen in the previous section, the system dimensioning depends on $f(E,E_0)$ only through γ. Thus the entire model depends only on the integral parameters γ and $R_{\rm turb}$.

If the sparks are distributed uniformly in the volume enclosed within a radius $R_s\leqslant R_{\rm fuel}$ (see Fig. 2), through Eq. (23) we find $R_{\rm turb}=R_s$ and

$$L_{\rm spark}=R_{\rm fuel}\ \nu(R_{\rm fuel}/R_s)^2. \tag{24}$$

In this case, a more convenient parameter, the dilution of the sparks (d), can be used instead of R_s:

$$d=\frac{4\pi R_s^3}{3}\left(\frac{2E}{3P}\right)^{-1}=\frac{1}{\nu}\left(\frac{R_s}{R_{\rm fuel}}\right)^3, \tag{25}$$

where Eqs. (15) and (17) are used. By definition $d\geqslant 1$. After d is assigned in the parametrization, we require $\nu d=(R_s/R_{\rm fuel})^3\leqslant 1$, so that $\nu\leqslant 1/d$. In terms of d, $L_{\rm spark}=R_{\rm fuel}(\nu/d^2)^{1/3}$, so that finally

$$\int_0^{R_{\rm fuel}}\rho\ dr=\left[\rho_{\rm cold}-(\rho_{\rm cold}-\rho_{\rm spark})\left(\frac{\nu}{d^2}\right)^{1/3}\right]R_{\rm fuel}. \tag{26}$$

Since $\rho_{\rm cold}>\rho_{\rm spark}$, the integral (and $f_{\rm burn}$) increases with d.

The conventional calculations for single, central spark[5] can be formally recovered from this model by setting $\gamma=1$ and $d=1$. From Eqs. (15), (16), (19), (24), and (25) it results in $E_{\rm hot}=E_{\rm thr}$, $L_{\rm spark}=R_s=R_{\rm thr}$, and the statement holds.

6. FUEL GAIN CALCULATIONS

In the following we present some gain evaluations for assemblies having the sparks dispersed uniformly within a radius $R_s\leqslant R_{\rm fuel}$ [see Fig. 2, and Eqs. (24)–(26)]. The uniform distribution in the entire fuel is obviously included as a special case ($\nu=1/d$). The relevant formula for $f_{\rm burn}$ is that obtained by using Eqs. (3) and (26).

The numerical results presented in this section refer to $E_{\rm fuel}=400$ kJ; this value is taken just for illustrative purposes. Here, we are not interested in the value of the gain

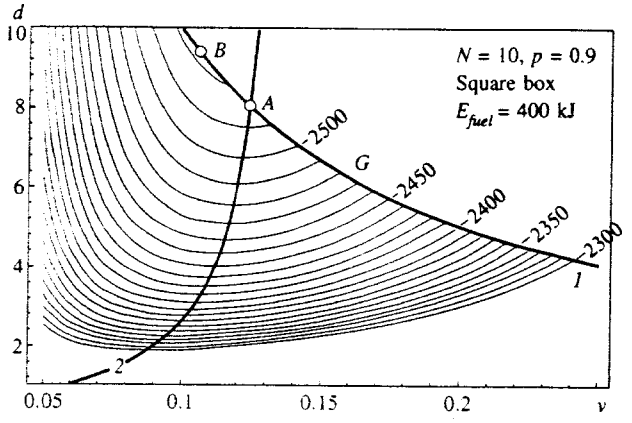

FIG. 3. Gain isolevels in the ν, d plane. The parameter $\nu = E_{\text{hot}}/E_{\text{fuel}}$ represents the total energy given to the sparks, measured in units of the total fuel energy E_{fuel}. The dilution parameter d represents the ratio of the volume in which the sparks are distributed (radius R_s) to the overall sparks volume. Between the curve *1* and the line $d=1$ lie the physically interesting points, those for which $R_s \leqslant R_{\text{fuel}}$. On the curve *2* lie the points of maximum gain at assigned d. The maximum possible gain is achieved on B, where *1* is tangent to an isolevel.

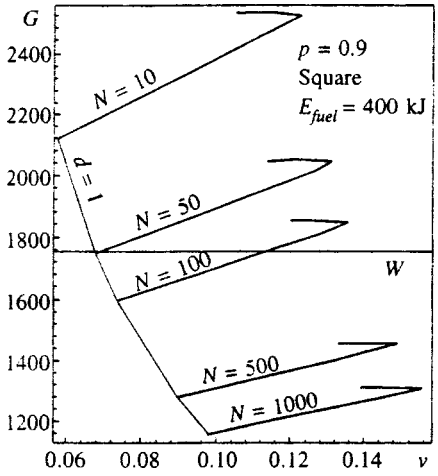

FIG. 4. The maximum gain (evaluated along line *2* in Fig. 3, and then from A to B) is reported for different values of N. The starting value is $d=1$.

which can be achieved, but in its relative change as a function of the number of sparks N. In all the calculations, the ignition temperature was assumed to be 7 keV [see Eq. (2)].

All the graphs presented refer to an ignition probability $p=0.9$ (unless the lowest values of N are considered, the results are largely insensitive to the value of this parameter; see Fig. 1). The ratio between the fuel pressure and the Fermi pressure in the cold fuel was assumed to be[8] $\alpha=2$. The typical behavior of the gain G in terms of ν and d is illustrated in Fig. 3, where the isolevels for G are shown. The behavior is always qualitatively the same, regardless of which values of N and p are considered, or the distribution function in energy is adopted.

Curve *1* represents the boundary for the region where $\nu \leqslant 1/d$ ($R_s \leqslant R_{\text{fuel}}$; see Sec. 5). The points on curve *1* represent uniform distributions of the sparks throughout the fuel volume ($R_s = R_{\text{fuel}}$). The curve *2* represents the loci where a maximum of G occurs for assigned d. On this curve, in the region below the point A, $(\partial G/\partial \nu)_d = 0$. An additional modest increase in G can be obtained by moving along curve *1* from A to the point B, where the line *1* is tangent to an isolevel [on this path $(\partial G/\partial \nu)_d \geqslant 0$]. The gain evaluated in B is the highest possible. The coordinates of B are given in Table I for the ''square-box'' distribution function (and $p=0.9$). Lying on curve *1*, these B points correspond to a uniform spark distribution in the fuel (i.e., $R_s = R_{\text{fuel}}$).

The maxima corresponding to different values of N are shown in Fig. 4 for the ''square-box'' distribution (following curve *2* up to the point A, and then curve *1* from A to B). Along each curve the value of the parameter d changes from 1 to the corresponding B points, which are listed in Table I. The sudden change in direction of each curve occurs at the transit through the A points (see Fig. 3). In Figs. 5a and 5b calculations for G and ρ_{cold} are presented (at $N=50$) for the ''square-box'' distribution. To carry out comparisons, the curve for the single-spark gain is also shown in Fig. 5a. The maximum for a single spark ($G_{ss}=2706$) occurs at $\nu = 0.0483$, and the associated cold fuel density is $\rho = 910\ \text{g/cm}^3$. This density value is about 0.5 of that associated with the maximum on curve *5* in Fig. 5a. Clearly, substantially more favorable results for a multispark follow from the calculations relative to the other distribution functions (see Fig. 1). For the single-spark approach, the site of the maximum gain is unlikely to be a working point since it corresponds to spark convergence ratios that are not practical. This can be deduced from prescriptions derived from some high-gain capsule designs, where at 3/4 of the initial target radius the in-flight aspect ratio (IFAR) is set equal to 30 and the total thickness (ablator+DT) is assumed to be equal to that of the initial, solid-state density DT layer.[7,8] Thus the initial shell radius is

$$R_0 = \left(\frac{\text{IFAR}}{3\pi} \frac{M_{\text{fuel}}}{\rho_0} \right)^{1/3}. \tag{27}$$

where ρ_0 is the DT density at the solid state. The spark convergence ratio is evaluated as $C_{\text{spark}} = R_0/R_{\text{thr}}$. The single-spark maximum gain corresponds to $C_{\text{spark}}=99$, a value much greater than those currently accepted[9] (≈ 40). The working value ($C_{\text{spark}}=40$) is recovered at $\nu=0.308$, where the gain is $G_{ssw}=1741$. This working point was labelled W in Fig. 5a. The gain at W is smaller than the maximum gain obtained with multi-spark systems. For a multispark, the geometric parameter corresponding to C_{spark} may be $C_{\text{hot}} = R_0/R_s$, where R_s is the radius of the turbulent region (see Sec. 5). For the highest gain $R_s = R_{\text{fuel}}$, and $C_{\text{hot}} = C_{\text{fuel}} = R_0/R_{\text{fuel}}$. Under these conditions, the highest gain for the case of Fig. 5a corresponds to $C_{\text{fuel}}=45$, whereas

TABLE I.

N	ν	d	G
10	0.1061	9.423	2523
50	0.1148	8.710	2042
100	0.1197	8.353	1845
500	0.1329	7.526	1445
1000	0.1394	7.174	1297

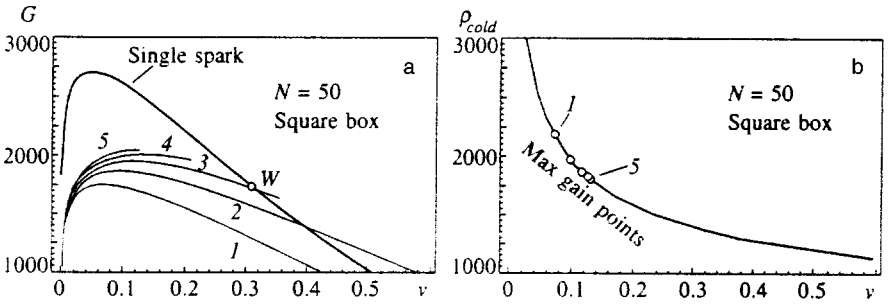

FIG. 5. a) Fuel gain for the case of 50 sparks distributed in energy according to the ''square-box'' distribution; $p=0.9$, $E_{\text{fuel}}=400$ kJ; *1*—$d=1$, *2*—$d=1.7$, *3*—$d=2.8$, *4*—$d=4.8$, *5*—$d=8$. b) Cold fuel density plotted as a function of v.

$C_{\text{fuel}}=39$ for the ''power-law'' case. To obtain a gain G_{ssw} by multispark configurations, more relaxed conditions are sufficient, namely those corresponding to $C_{\text{fuel}}\approx 38$ for the case of Fig. 5a, and $C_{\text{fuel}}\approx 30$ for a ''power-law'' distribution. These conditions should be used in the implosion design, taking into account that in a multispark assembly the overall shape of the final fuel is expected to play a secondary role in determining thermonuclear performances of the system.

7. CONCLUSIONS

In the conventional ICF approach a central issue is the formation of a spherical triggering spark at the center of the fuel, at the end of the implosion process. Highly deformed sparks do not ignite[1] and in experiments the measured number of neutrons was found to deviate substantially from that computed in $1D$ simulations. This deviation increases with the implosion convergence ratio. The usual interpretation attributes these results to a mixing process.

In this paper we assumed that as a final result of the implosion, a multispark system has been generated. It was found that high thermonuclear fuel gain can be still obtained from compressed fuel assemblies, in which a large number of sparks ($N=10-100$) with a statistically generated energy (size) spectrum is created. By requiring to have, with an assigned probability, at least one spark large enough to ignite the fuel we have computed the assembly dimensioning and by using statistical arguments we have computed the burn fraction and the fuel gain. It was found that the results may depend on the values of two integral parameters, which depend on the energy-space distribution function of the sparks.

Although requiring more compressed fuels (typically $2\times-3\times$), the multispark approach may prove to be interesting because (in principle) it is based on the hypothesis of *ab initio* relaxed implosion symmetry requirements. Final mixing processes, which are currently considered to be adverse in ICF, could turn out to be not so dangerous, since high gain can still be obtained. Based on these considerations, the domain of implosion designs can be expanded so as to make it relevant for thermonuclear energy.

[1] R. L. McCrory and C. P. Verdon, in *Proc. of the Intern. School Piero Caldirola on Inertial Confinement Fusion*, ed. by A. Caruso and E. Sindoni, Editrice Compositori (1988), p. 83.
[2] H. Sakagami and K. Nishihara, Phys. Fluids B **2**, 2715 (1990).
[3] H. Sakagami and K. Nishihara, Phys. Rev. Lett. **42**, 432 (1990).
[4] S. Yu. Gus'kov, O. N. Krokhin, and V. B. Rozanov, Nucl. Fusion **16**, 957 (1976).
[5] M. D. Rosen, J. D. Lindl, and A. R. Thiessen, LLNL Annual Report (1983).
[6] L. D. Landau and E. M. Lifshitz, *Theoretical Physics, Vol. 5*, Pergamon Press (1980), pp. 162–166.
[7] A. Caruso, in *Proc. of the Intern. School Piero Caldirola on Inertial Confinement Fusion*, ed. by A. Caruso and E. Sindoni, Editrice Compositori (1988), p. 139.
[8] J. Lindl, Phys. Plasmas **2**, 3933 (1995).
[9] J. Lindl, in *Proc. of the Intern. School Piero, Caldirola on Inertial Confinement Fusion*, ed. by A. Caruso and E. Sindoni, Editrice Compositori (1988), p. 595.

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Superfluorescence of photonic paint

A. Yu. Zyuzin

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
Max-Planck-Institut für Festkorperforschung, BP166, F-38042 Grenoble, France
(Submitted 16 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 816–825 (March 1998)

We consider the cooperative decay of incoherently pumped atoms in a disordered medium, where light undergoes multiple scattering. It is shown that the cooperation number, which determines the duration and amplitude of superfluorescent impulses, is given by the number of atoms along a diffusive trajectory of the light propagating through the medium. We also consider the problem of reflection of a probe wave during cooperative emission.
 [S1063-7761(98)00403-X]

1. INTRODUCTION

There is growing interest in active photonic paints. These are media in which light undergoes multiple random scattering, resulting in the diffusive propagation of radiation, while interacting with atoms that can be pumped to obtain a positive population difference. The reflection and transmission of electromagnetic waves through such a cavity has been extensively studied over the past decade. The speckle pattern resulting from scattering has an average enhancement in the direction opposite the direction of the incident radiation[1] (a comprehensive review of other statistical properties of the speckle of reflected and transmitted waves is given in Ref. 2).

Feedback provided by scattering in such a random cavity can serve to set up laser oscillations.[3] Laser action in powdered laser materials,[4,5] laser dye solutions with scattering nanoparticles,[6] and dye-doped microdroplets containing Intralipid as a scatterer[7] has recently been reported. These experiments concentrated mostly on temporal and spatial properties of emission.

Recently, the proposed[8] enhancement of the weak localization peak in backscattering from an amplifying photonic paint was observed.[9]

The relevant question concerning recent observations of generation of light in active photonic paints[4,5] is to what extent this phenomenon is reminiscent of superfluorescence[10,11] (i.e., the cooperative decay of an incoherently pumped system of dipole transitions, started by initial noise or an external electromagnetic field), which has usually been studied in systems without scattering.

Here we consider the cooperative decay of incoherently pumped atoms in a random cavity, which is a slab of thickness L ($L \gg l$, where l is the mean free path of radiation). This geometry is often used in experiments. The time that light spends in this cavity is of the order of L^2/D, where D is the diffusion constant. This time is to be compared with the energy exchange time between atoms and field. We show that if the latter is greater than L^2/D, then after some delay, the system will generate a superfluorescent pulse of hyperbolic secant form.

The duration of the superfluorescent pulse is $\tau_{\rm rad} N_C^{-1}$,[10] where $\tau_{\rm rad}$ is the radiative decay time of a single atom and N_C is the cooperation number, i.e., the number of atoms that take part in cooperative decay. We find that in disordered systems, this number is $N_C \propto \rho \lambda^2 L^2/l$ (λ is the wavelength of the radiation and ρ is the density of active atoms, such that $\rho\lambda^3 \gg 1$), i.e., it is equal to the number of atoms in a tube with cross section λ^2 and length of the order of L^2/l, which is the length of a diffusive trajectory of radiation.

The intensity of radiation of cooperating atoms at the maximum of the superfluorescent pulse is $\propto N_C^2/\tau_{\rm rad}$.[10] We show that the diffusive slab radiates at maximum as a system of V/N_C independent groups of cooperating atoms, and at the peak of the superfluorescent pulse, the intensity emitted by the slab is $\propto (N_C^2/\tau_{\rm rad})(V/N_C)$; V is the volume of the slab.

The maximum cooperation number for given $\tau_{\rm rad}$ and density of active atoms is determined by the condition that the time of energy exchange between atoms and the field equal the time that light spends in the cavity. From this condition, we find that the maximum cooperation number in the random cavity is $N_C^{\max} \propto \lambda \sqrt{c\rho\tau_{\rm rad}}$, where c is the speed of radiation in the slab.

These results are valid in the case of weak dephasing processes and long relaxation of population difference. Below we take into account the effect of dephasing on superfluorescence.

In the limit of large radiation escape time L^2/D, atoms exchange energy with the field many times, so stimulated emission becomes important and the system exhibits oscillatory behavior.

We also consider the reflection of the probe wave during decay of the pumped system.

2. BASIC EQUATIONS

We model a random medium in the following way. The dielectric function $\varepsilon(\mathbf{r})$ of the medium, which contains active atoms, is a random function of position, such that $\langle \varepsilon(\mathbf{r}) \rangle = 1$. Scattering of light is due to fluctuations of the dielectric function with white noise-like variance

$$\langle \delta\varepsilon(\mathbf{r})\,\delta\varepsilon(\mathbf{r}')\rangle = \frac{\lambda^4}{4\pi^3 l}\,\delta(\mathbf{r}-\mathbf{r}').$$

We consider the case of a weakly disordered system $l \gg \lambda$, with dimensions larger than the mean free path, so propagation of the field can be described as a diffusion process with diffusion constant $D=cl/3$; c is the speed of light in the medium.

The coupling between the polarization density

$$\frac{1}{2}\,[e^{i\omega t}P(\mathbf{r},t)+e^{-i\omega t}P^*(\mathbf{r},t)]$$

averaged over scales smaller than λ, the population difference density $\Delta N(\mathbf{r},t)$, and the field

$$\frac{1}{2}\,[e^{i\omega t}E(\mathbf{r},t)+e^{-i\omega t}E^*(\mathbf{r},t)]$$

can be described by the classical Maxwell–Bloch equations. In this approach, amplified spontaneous emission noise is neglected, which is a good approximation for superfluorescence;[12] $P(\mathbf{r},t)$ and $E(\mathbf{r},t)$ are slowly time-varying complex quantities, which we consider to be scalars; ω is the atomic frequency.

The first two Maxwell–Bloch equations have the form[13]

$$\left[\frac{d}{dt}+\gamma\right]P(\mathbf{r},t)=\frac{i|\mu|^2}{\hbar}\,\Delta N(\mathbf{r},t)E(\mathbf{r},t), \tag{1}$$

$$\frac{d}{dt}\,\Delta N(\mathbf{r},t)=-\frac{i}{2\hbar}\,[P^*(\mathbf{r},t)E(\mathbf{r},t)-P(\mathbf{r},t)E^*(\mathbf{r},t)]. \tag{2}$$

Here γ is the inverse dephasing time and μ is the electric dipole moment. It is assumed that the population inversion relaxation time is longer than the delay time of the superfluorescent pulse. We also neglect inhomogeneous broadening.

The quantities $\Delta N(\mathbf{r},t)$ and $P(\mathbf{r},t)/|\mu|$ are components of the local Bloch vector averaged over scales smaller than the wavelength of the radiation. The rate at which its length decreases, according to (1) and (2), is determined by γ^{-1}.

The wave equation for the slow time variations of the field component $E(r,t)$ has the form

$$i\,\frac{dE(\mathbf{r},t)}{dt}-\left[-\frac{c^2}{2\omega}\,\Delta-\frac{\omega\varepsilon(\mathbf{r})}{2}\right]E(\mathbf{r},t)=2\pi\omega P(\mathbf{r},t). \tag{3}$$

Although $E(\mathbf{r},t)$ and $P(\mathbf{r},t)$ vary slowly in time, they still contain random spatial phases, which result from random interference between waves coming to the point $\mathbf{r}$ via different diffusive trajectories. To get rid of these phase factors, it is convenient to consider the diffusion propagator $D(\mathbf{r};t_1,t_2)$, which determines the correlation function of the polarization density and field:

$$\langle E(\mathbf{r},t_1)E^*(\mathbf{r},t_2)\rangle=4\pi k^3\omega D(\mathbf{r};t_1,t_2). \tag{4}$$

Correlation functions involving the polarization density can be obtained by using Eq. (1).

To obtain the equation for the diffusion propagator, it is convenient to eliminate the polarization density from Eqs. (1) and (3). Then the usual diagram technique[14] makes it possible to calculate the average of the product $E(\mathbf{r},t_1)E^*(\mathbf{r},t_2)$. The corresponding diagrams are shown in Fig. 1.

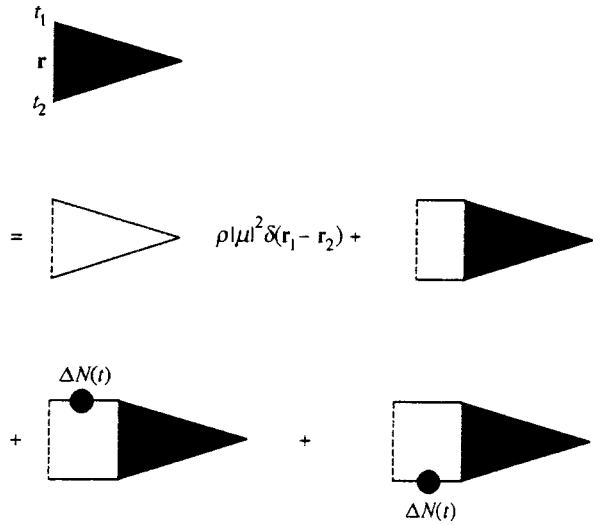

FIG. 1. Diagram for the diffusion ladder, which is denoted by shaded triangles. Solid lines correspond to the average Green function of the field, dashed lines denote scattering. Shaded circles denote interaction between field and atoms. The average Green function satisfies the equation $idE(\mathbf{r},t)/dt-\{-c^2\Delta/2\omega-\omega/2+ic,2l\}E(\mathbf{r},t)=\delta(\mathbf{r}-\mathbf{r}')\,\delta(t)$.

Considering the evolution of the Bloch vector from time $t=0$, at which a positive population difference is created, we obtain for the diffusion propagator

$$\left[\frac{d}{dt_1}+\frac{d}{dt_2}-D\nabla^2\right]D(\mathbf{r};t_1,t_2)=f(\mathbf{r};t_1,t_2)+\frac{1}{\rho\tau_0^2}\int_0^{t_1}dt\,\exp[-\gamma(t_1-t)]\Delta N(\mathbf{r},t)D(\mathbf{r};t,t_2)+\frac{1}{\rho\tau_0^2}\int_0^{t_2}dt\,\exp[-\gamma(t_2-t)]\Delta N(\mathbf{r},t)D(\mathbf{r};t_1,t). \tag{5}$$

Here

$$\tau_0\equiv\sqrt{\frac{\hbar}{2\pi\rho\omega|\mu|^2}}=\sqrt{\frac{4\pi\tau_{\rm rad}}{3\omega\rho\lambda^3}}$$

is the characteristic time of energy exchange between the field and the atomic system,[11] ρ is the density of active atoms, and

$$\tau_{\rm rad}^{-1}=\frac{8\pi^2|\mu|^2}{3\hbar\lambda^3}$$

is the radiative decay time of a single atom.

The function $f(\mathbf{r}:t_1,t_2)$ depends on the initial conditions. Here we choose the initial condition such that $\langle P(\mathbf{r},t=0)\rangle=0$ and $\langle P(\mathbf{r},t=0)P^*(\mathbf{r}',t=0)\rangle=\rho|\mu|^2\delta(\mathbf{r}-\mathbf{r}')$. This initial condition corresponds to an initial incoherent state. In this case

$$f=\rho|\mu|^2\,\exp[-\gamma(t_1+t_2)] \tag{6}$$

for times greater than the mean free time of radiation l/c.

The equation for the mean population inversion density can be obtained by using Eqs. (1), (2), and (5):

$$\frac{d\Delta N(\mathbf{r},t)}{dt}=-\frac{(2\pi)^3}{\hbar\rho\lambda^3\tau_0^2}\int_0^t dt'\ \exp[-\gamma(t-t')]\times\Delta N(\mathbf{r},t')[D(\mathbf{r};t',t)+D(\mathbf{r};t,t')]. \quad (7)$$

For the population difference we choose $\Delta N(\mathbf{r},t=0)=\Delta N>0$ as the initial condition ($\Delta N=\rho$).

The usual boundary conditions for the diffusion propagator are $D(\mathbf{r};t_1,t_2)$ on an open surface and $\mathbf{n}\cdot\nabla D(\mathbf{r};t_1,t_2)=0$ on a reflecting surface; $\mathbf{n}$ is the normal to the reflecting surface.

The diffusion approach is justified if the time of energy exchange between atoms and field is greater than the mean free time of radiation, $\tau_0\gg l/c$.

3. COOPERATIVE DECAY IN PHOTONIC PAINT

Below we consider a slab of thickness L ($L\gg l$). Let z be the coordinate across the slab $L\geqslant z\geqslant 0$. It is convenient to study the solution of Eq. (5) in the form

$$D(\mathbf{r};t_1,t_2)=\sqrt{\frac{L}{2}}\sum_{n=1}^{\infty}\Psi_n(z)D_n(t_1,t_2), \quad (8)$$

where

$$\Psi_n(z)=\sqrt{\frac{2}{L}}\sin\frac{\pi nz}{L}$$

is an eigenfunction of the diffusion equation with boundary condition $\Psi_n(z)=0$ at the free boundaries $z=0, L$.

Let us consider the initial evolution of the diffusion propagator, when the population difference does not depend on time. For the coefficients in (8), we obtain from (5)

$$\left[\frac{d}{dt_1}+\frac{d}{dt_2}+\omega_n\right]D_n(t_1,t_2)=\sqrt{\frac{2}{L}}\,\rho|\mu|^2\int dz\Psi_n(z)\exp[-\gamma(t_1+t_2)]+\frac{\Delta N}{\rho\tau_0^2}\int_0^{t_1}dt\times\exp[-\gamma(t_1-t)]D_n(t,t_2)+\frac{\Delta N}{\rho\tau_0^2}\int_0^{t_2}dt\times\exp[-\gamma(t_2-t)]D_n(t_1,t). \quad (9)$$

Here $\omega_n=D\pi^2n^2/L^2$ is an eigenvalue of the diffusion equation. Solving Eq. (9) via the Laplace transform with initial conditions $D_n(0,t_2)=D_n(t_1,0)=0$ (the field vanishes at $t=0$), we obtain

$$D_n(t,t)\propto\exp\left\{\left[\sqrt{\left(\frac{\omega_n}{2}-\gamma\right)^2+\frac{4\Delta N}{\rho\tau_0^2}}-\frac{\omega_n}{2}-\gamma\right]t\right\}. \quad (10)$$

The critical value of positive inversion density ΔN_n, above which the growth rate of a particular diffusion mode

$$Z_n=\sqrt{\left(\frac{\omega_n}{2}-\gamma\right)^2+\frac{4\Delta N}{\rho\tau_0^2}}-\frac{\omega_n}{2}-\gamma$$

becomes positive, is $\Delta N_n=(\omega_n\gamma/2)\rho\tau_0^2$.[3] More detailed calculations of (10) are given in the next section.

To proceed further in solving Eqs. (5) and (7), we make two approximations.

1) Below we consider the case of fast escape of radiation from the system, where $\omega_1\gg d/dt_1, d/dt_2, \gamma$ (or, according to (10), $\tau_0\omega_1>1$ for weak dephasing), so we can neglect the time derivative in Eq. (5). In the language of superfluorescence, this situation corresponds to the case in which there is no energy exchange between the emitted field and atomic subsystem.[12] The field serves only to develop correlation between atoms.

2) We consider only the most unstable mode $D_1(t_1,t_2)$. At $t=0$, the off-diagonal elements of

$$\Delta N_{nm}(t)\equiv\int dz\Delta N(z,t)\Psi_n(z)\Psi_m(z)$$

are zero by definition, and interaction between modes is irrelevant for most of the time of decay. We therefore assume that the interaction of the first diffusion mode $D_1(t_1,t_2)$ with higher modes does not qualitatively change the description of cooperative decay.

Under these assumptions the equation for the diffusion propagator has the form

$$D_1(t_1,t_2)=\frac{4\rho|\mu|^2}{\pi\omega_1}\exp[-\gamma(t_1+t_2)]+\frac{1}{\rho\omega_1\tau_0^2}\int_0^{t_1}dt\times\exp[-\gamma(t_1-t)]\Delta N_{11}(t)D_1(t,t_2)+\frac{1}{\rho\omega_1\tau_0^2}\times\int_0^{t_2}dt\ \exp[-\gamma(t_2-t)]\Delta N_{11}(t)D_1(t_1,t), \quad (11)$$

and for the population difference

$$\frac{d}{dt}\Delta N_{11}(t)=-\frac{k^3}{\hbar}\left[\frac{8}{3\pi}\,\omega_1D_1(t,t)-\rho|\mu|^2\times\exp(-2\gamma t)\right]. \quad (12)$$

Introducing

$$D(t_1,t_2)\equiv\exp[-\gamma(t_1+t_2)]G(\chi(t_1),\chi(t_2)) \quad (13)$$

in (11), where

$$\chi(t)=\frac{1}{\rho\omega_1\tau_0^2}\int_0^t dt\Delta N_{11}(t), \quad (14)$$

we obtain

$$G(\chi_1,\chi_2)=\frac{4\rho|\mu|^2}{\pi\omega_1}+\int_0^{\chi_1}d\chi G(\chi,\chi_2)+\int_0^{\chi_2}d\chi G(\chi_1,\chi). \quad (15)$$

Equation (15) can be solved by a Laplace transform:

$$G(\chi_1,\chi_2)=\frac{4\rho|\mu|^2}{\pi\omega_1}\int_{-i\infty+C}^{i\infty+C}\frac{dz_1}{2\pi i}\frac{dz_2}{2\pi i}\frac{\exp(z_1\chi_1+z_2\chi_2)}{z_1z_2-z_1-z_2}. \quad (16)$$

The asymptotic form of (16) for $\chi_1=\chi_2\equiv\chi>1$ is

$$G(\chi,\chi)\approx\frac{2\rho|\mu|^2}{\pi\omega_1}\frac{\exp(4\chi)}{\sqrt{\pi\chi}}. \tag{17}$$

The equation for the population difference ($\chi>1$) is

$$\frac{d^2}{dt^2}\chi=-\frac{8k^3}{3\pi\hbar\rho\tau_0^2}G(\chi,\chi)\exp(-2\gamma t). \tag{18}$$

Taking into account only exponential factors, we obtain the solution of Eq. (18):

$$\Delta N_{11}(t)=\delta N(0)\tanh\left\{\frac{2\delta N(0)(t_0-t)}{\rho\tau_0^2\omega_1}\right\}+\frac{\gamma\rho\tau_0^2\omega_1}{2}. \tag{19}$$

Here we introduce $\delta N_{11}(0)\equiv\Delta N-\gamma\omega_1\rho\tau_0^2/2$; ΔN is the population difference at the beginning of exponential growth of radiative intensity, when deviation from the initial population difference is small ($\Delta N=\rho$).

The delay time in (19) is

$$t_0=\frac{\omega_1\rho\tau_0^2}{2\delta N(0)}\ln\left[\frac{\delta N(0)}{\rho}\sqrt{\rho\lambda^3}\frac{L}{\sqrt{l\lambda}}\right].$$

In deriving this expression we took into account the relation between the time of energy exchange between atoms and field and $|\mu|^2$, which enters into the initial condition for polarization density.

The radiative intensity is proportional to $d\Delta N_{11}(t)/dt$, and is emitted as a hyperbolic secant pulse. The result (19) coincides with that of the Markov theory of superfluorescence in a system without scattering.[10,11] The difference is in the definition of the cooperation number.

It follows from Eq. (19) that in the case of weak dephasing, the duration of a superfluorescent pulse is $\tau_0^2\omega_1/4\equiv\tau_{\rm rad}N_C^{-1}$, where $\tau_{\rm rad}$ is the time of radiative decay of a single atom and

$$N_C=\frac{18\rho\lambda^2L^2}{\pi^2 l}$$

is the cooperation number, i.e., the number of atoms that take part in the cooperative decay. This is equal to the number of atoms in a tube of cross section λ^2 with the length of the diffusive trajectory L^2/l. The maximum of the cooperation number is determined by the condition $\tau_0\omega_1=1$, whereupon $N_C^{\max}=2\lambda\sqrt{6c\rho\tau_{\rm rad}}$. Under this condition, atoms can exchange energy with the field only once, i.e., stimulated emission can be neglected. We note that for a given density ρ, decay time $\tau_{\rm rad}$ and velocity, the maximum cooperation number in a disordered system is smaller than in a pencil-shaped system without scattering.[15]

The maximum emitted radiation is

$$V\frac{d}{dt}\Delta N_{11}(t=t_0)$$

(V is the volume of the slab). It can also be written $\mathcal{N}N_C^2/\tau_{\rm rad}$, where $\mathcal{N}=V\rho/N_C$ is the number of cooperative regions in the slab.

Cooperative decay in a diffusive medium can therefore be interpreted as the independent cooperative decay of $\mathcal{N}=V\rho/N_C$ systems, each consisting of N_C atoms.

Dephasing processes increase the duration of a pulse by a factor $\rho(\Delta N-\gamma\omega_1\rho\tau_0^2/2)^{-1}$, and reduce the peak intensity by the square of this factor. Note that this result coincides with that for a system without scattering.[16]

If $\tau_0\omega_1\ll1$, atoms exchange energy with the field many times. In this case we expect spiking of intensity. The frequency of spiking can be estimated[13] from Eq. (10) as

$$\sqrt{\left|\frac{4\Delta N}{\rho\tau_0^2}\right|-\left(\frac{\omega_1}{2}-\gamma\right)^2}.$$

To obtain this expression we insert a negative value of the population inversion[13] into (10) (this situation will occur after the pumped atoms exchange energy with the field).

4. AMPLIFICATION IN THE BACKWARD DIRECTION

Correlation between pumped atoms can also be due to the external field, which stimulates emission in the forward direction in a system without scattering.[13] In a disordered system one might expect enhancement of emission in the backward direction.

Here we consider the reflection of a weak probe plane wave with frequency ω during the development of superfluorescent emission. The amplitude of the probe is low, so the effect of the external field on emission can be neglected. We can also neglect interference between the external and emitted fields, because the initial state of polarization is incoherent. Below we consider in detail the linear stage of decay when the inversion density is high enough to produce only the lowest diffusion mode instability, $\Delta N=\Delta N_1(1+\delta)$, $\delta\ll1$. This situation resembles the experimental setup of Ref. 9.

It is convenient to calculate the albedo, which is the ratio between the intensities of the reflected and incident fields. The time-dependent albedo can be expressed as[17]

$$\alpha(\mathbf{q},t)=\frac{c}{4\pi l^2}\int_0^\infty dz\,dz'\,\exp\left(-\frac{z+z'}{l}\right)\times\int d\boldsymbol{\rho}[1+\cos(\mathbf{q},\boldsymbol{\rho})]D(z,z',\boldsymbol{\rho};t). \tag{20}$$

Here $\mathbf{q}$ is the sum of the incident and outgoing wave vectors, and ρ is the position in the plane. The diffusion propagator (20) obeys Eq. (5) with the substitution of $\delta(\mathbf{r}-\mathbf{r}')$ for $f(\mathbf{r};t_1,t_2)$. We also assume that the incident wave is close to the normal to the surface.

The first term describes diffusion scattering and the second term describes the interference part, which is strongly peaked in the backward direction. The physical mechanism of the interference contribution is exhaustively discussed in the literature: see, for example, Refs. 1 and Ref. 8, and references therein.

The diffusion propagator can be represented as

$$D(\mathbf{r},\mathbf{r};t_1,t_2)=\sum_n \Psi_n(z)\Psi_n(z') \times\exp[i\mathbf{q}(\boldsymbol{\rho}-\boldsymbol{\rho}')]\hat{D}_n(\mathbf{q},t_1,t_2). \quad (21)$$

The Laplace transform of Eq. (5) for time-independent $\Delta N>0$ yields

$$\hat{D}_n(\mathbf{q},t_1,t_2)=\int_{-i\infty+C}^{i\infty+C}\frac{dp_1dp_2}{(2\pi i)^2}\frac{\exp(p_1t_1+p_2t_2)}{p_1p_2} \times\left\{p_1+p_2+\Omega_n(\mathbf{q})-\frac{\Delta N}{\rho\tau_0}\left(\frac{1}{p_1+\gamma}+\frac{1}{p_2+\gamma}\right)\right\}^{-1}, \quad (22)$$

where the real part of the integration contour passes to the left of all singularities, and $\Omega_n(\mathbf{q})=Dq^2+\omega_n$ is the eigenvalue of the diffusion equation for the slab geometry.

At $t_1=t_2$, integrating over the difference p_1-p_2 in (22), we obtain for the first mode at $t>\omega_1^{-1}$

$$\hat{D}_1(\mathbf{q},t,t)=\frac{(2\gamma)^{3/2}}{\sqrt{\omega_1(\omega_1+2\gamma)}}\int_{-i\infty+C}^{i\infty+C}\frac{dp}{2\pi i}\frac{\exp(pt)}{p\sqrt{p-Z_1(q)}} \times\left[p+\sqrt{\frac{\gamma(\omega_1+2\gamma)}{2\omega_1}}\sqrt{p-Z_1(q)}\right]^{-1}. \quad (23)$$

Here we introduce

$$Z_1(q)=\frac{2\gamma\omega_1}{2\gamma+\omega_1}\left(\delta-\frac{Dq^2}{\omega_1}\right),$$

which is the growth rate of $\hat{D}_1(\mathbf{q},t,t)$.

For moderate times

$$\frac{t\omega_1Z_1^2}{\gamma(2\gamma+\omega_1)}<1,$$

we obtain from (23)

$$\hat{D}_1(\mathbf{q},t,t)=\frac{2}{\omega_1}\frac{\exp[Z_1(q)t]-1}{\delta-Dq^2/\omega_1}. \quad (24)$$

This expression is valid for either sign of Z_1, i.e., above as well as below threshold.

Taking into account that

$$\Psi_1(z)=\sqrt{\frac{2}{L}}\sin\frac{\pi z}{L},$$

we obtain the singular contribution to the albedo from the first mode:

$$\delta\alpha(\mathbf{q},t)=\frac{3l}{\pi L}\left\{\frac{\exp[Z_1(0)t]-1}{\delta}+\frac{\exp[Z_1(q)t]-1}{\delta-Dq^2/\omega_1}\right\}. \quad (25)$$

Below threshold the albedo is saturated. The peak at large times has a Laplacian form $\propto(|\delta|+Dq^2/\omega_1)^{-1}$. At threshold and above, the peak narrows with time. Exactly at threshold the albedo is linear with time, and above threshold the albedo grows exponentially.

5. CONCLUSION

To summarize, superfluorescent emission of active photonic paint develops due to the cooperation of atoms along a diffusive trajectory through a system with cross-sectional dimensions of the order of a wavelength. The pulse therefore becomes narrower with decreasing mean free path of radiation until the cooperation number reaches its maximum value. The maximum cooperation number does not depend on disorder.

An external field enhances emission in the backward direction. The peak sharpens in coherent backscattering during cooperative decay in a disordered system.

I thank A. V. Gol'tsev for useful suggestions. This work was supported by the Russian Fund for Fundamental Research under Grant number 97-02-18078.

[1] Y. Kuga and A. Ishimaru, J. Opt. Soc. Am. A **8**, 831 (1984); M. P. van Albada and A. Lagendijk, Phys. Rev. Lett. **55**, 2692 (1985); P. E. Wolf and G. Maret, Phys. Rev. Lett. **55**, 2696 (1985).
[2] A. Z. Genack, in *Scattering and Localization of Classical Waves in Random Media*, P. Sheng (ed.), World Scientific, Singapore (1990).
[3] V. S. Letokhov, Zh. Eksp. Teor. Fiz. **53**, 1442 (1968) [Sov. Phys. JETP **26**, 835 (1968)].
[4] N. E. Ter-Gabrielyan, V. M. Markushev, V. R. Belan, C. M. Briskina, O. V. Dimitrova, V. F. Zolin, and A. V. Lavrov, Kvant. Elektr. **18**, 313 (1991) [Sov. J. Quantum Electron. **21**, 281 (1991)].
[5] C. Gouedard, D. Husson, C. Sauteret, F. Auzel, and A. Migus, J. Opt. Soc. Am. B **10**, 2358 (1993).
[6] N. M. Lawandy, R. M. Balachandran, A. S. L. Gomes, and E. Sauvain, Nature (London) **368**, 436 (1994).
[7] H. Taniguchi, S. Tanosaki, K. Tsujita, and H. Inaba, IEEE J. Quantum Electron. **32**, 1864 (1996).
[8] A. Zyuzin, Europhys. Lett. **26**, 517 (1994).
[9] D. S. Wiersma, M. P. van Albada, and A. Langendijk, Phys. Rev. Lett. **75**, 1739 (1995).
[10] R. H. Dicke, Phys. Rev. **93**, 99 (1954).
[11] R. Bonifacio and L. A. Lugiato, Phys. Rev. A **11**, 1507 (1975).
[12] R. Bonifacio, P. Schwendimann, and F. Haake, Phys. Rev. A **4**, 302, 854 (1971).
[13] A. E. Siegman, *Lasers*, Oxford University Press, New York (1986).
[14] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics*, Interscience, New York (1963).
[15] F. T. Arecchi and E. Courtens, Phys. Rev. A **2**, 1730 (1970).
[16] A. V. Andreev, Yu. A. Il'inskii, and R. M. Khokhlov, Zh. Éksp. Teor. Fiz. **73**, 1296 (1977) [Sov. Phys. JETP **46**, 682 (1977)].
[17] E. Akkermans, P. E. Wolf, and R. Maynard, Phys. Rev. Lett. **56**, 1471 (1986).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Study of the rates of collisional decay of population, orientation, and alignment by stimulated photon echo in a molecular gas

L. S. Vasilenko, N. N. Rubtsova,*) and E. B. Khvorostov

Institute of Semiconductor Physics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 13 September 1997)
Zh. Éksp. Teor. Fiz. **113**, 826–833 (March 1998)

A stimulated photon echo technique with specially selected linear polarizations of the coherent resonant driver pulses is used to study depolarizing collisions in the molecular gas SF_6 and in mixtures of it with buffer He and Xe. The collisional decay rates of the population, orientation, and alignment in an ensemble of gas particles are determined for the first time in a single experiment. These relaxation rates are measured as a function of the longitudinal translational velocities of the resonant particles. To within the experimental accuracy, no significant dependence of the collisional decay rates on the translational velocities of the particles was observed. This result confirms the conventional theoretical approach to depolarizing collisions. In pure SF_6 the decay rates for the orientation and alignment were lower than the relaxation constant for collisions involving a change in the longitudinal velocity (elastic collisions) that is known from experimental observations of the ordinary photon echo. This means that only some of the elastic collisions participate in destroying the multipole moments of the levels. Evidence is found that the relaxation of the multipole moments created by polarized radiation in a resonant medium of molecular SF_6 gas depends on j, the total angular momentum of the level. © *1998 American Institute of Physics.* [S1063-7761(98)00503-4]

1. INTRODUCTION

Since the first experiments on optical pumping,[1] it has been well known that polarized light can create a distinctive nonequilibrium distribution of the population over the magnetic sublevels of an excited and (or) ground state.[2] This sort of polarized state decays through depolarizing collisions which are very sensitive to asymmetries in the interaction potential.[3]

Coherent transient phenomena, such as the photon echo and modifications of it, are widely used to study relaxation processes in various media,[4,5] including gases.[6–8]

The ordinary photon echo develops in a resonant medium at a time $2T_{12}$ following passage of two short light pulses separated by a time delay T_{12}. The photon echo is a form of spontaneous coherent emission from an ensemble of particles which retains a phase memory of the first exciting pulse in the nondiagonal component of the optical coherence matrix. Thus, the ordinary photon echo has been used successfully for measuring the relaxation constant $\gamma_{ab}^{(1)}$ (the rate of decay of the polarization of the medium).[9] It was predicted that this method might make it possible to measure other relaxation characteristics γ^{κ} $(\kappa \neq 1)$ of the optical coherent matrix by creating a photon echo in an external magnetic field.[10]

Polarization effects during photon echo formation can be used to identify a transition for small j or to identify the type of transition for $j \gg 1$.[11] The polarization features of echo-response generation in standing waves can aid in measuring the spontaneous emission probability for a resonant transition, the relaxation rates for all the components of the optical coherence matrix,[12] and the magnitude of the collisional shift in a spectrum line.[13]

Stimulated photon echo techniques can be employed in measuring some relaxation constants of the quantum levels themselves (given their degeneracy over the magnetic sublevels), such as the relaxation of population, orientation, and alignment without applying an external magnetic field.[14] A stimulated photon echo signal appears at a time $2T_{12}+T_{23}$, after which the resonant medium is successively excited by three light pulses with time delays T_{12} between the first and second and T_{23} between the second and third. A phase memory of the first and second driver pulses is retained, throughout the time delay T_{23}, in the diagonal components of the coherence matrices $\rho_{\mu\mu'}^{(aa)}$ and $\rho_{mm'}^{(bb)}$. Here m and μ are the projections of the total angular momenta of the upper (j_a) and lower (j_b) levels, respectively.

Following Evseev *et al.*,[14] we consider the excitation of a resonant medium by a sequence of three linearly polarized pulses propagating along the y axis. Let the polarization of the third pulse coincide with the direction of the z axis, while the polarizations of the first and second pulses are rotated by angles ψ_1 and ψ_2, respectively. In order to determine the electric amplitude of the simulated photon echo, we use an expansion of the coherence matrix $\rho_{m\mu}^{(ba)}$ and the collision matrix $\Gamma_{m\mu}^{m'\mu'}(\mathbf{v})$ in terms of irreducible tensor operators.[10] The first three terms in the expansion of the diagonal component of $\rho_{m\mu}^{(ba)}$ are the total population of the level (zeroth term), the orientation of the level (first term), and the alignment of the level (second term). The appearance of the first or second terms corresponds to the formation of a macro-

scopic magnetic moment or macroscopic electric quadrupole moment in a subensemble of resonant particles. The expansion of the collision matrix $\Gamma^{m'\mu'}_{m\mu}(\mathbf{v})$ gives the corresponding relaxation rates: $\gamma^{(0)}$ for the population, $\gamma^{(1)}$ for the orientation, and $\gamma^{(2)}$ for the alignment of the levels. For driver pulses with small areas ($\theta_i \ll 1$, $i=1,2,3$), these theoretical calculations of the stimulated photon echo amplitude yield Eqs. (1a)–(1c) for $j \leftrightarrow j$ transitions and Eqs. (2a)–(2c) for $j \leftrightarrow j+1$ transitions for $j \gg 1$ (typical for experiments in SF_6 gas):[14]

$$\mathscr{E}_x \propto \frac{1}{15j} \sin(\psi_1+\psi_2) A_2(T_{23}), \tag{1a}$$

$$\mathscr{E}_y = 0, \tag{1b}$$

$$\mathscr{E}_z \propto \frac{1}{9j}\left[\cos(\psi_1-\psi_2)A_0(T_{23}) + \frac{2}{5}(2\cos\psi_1\cos\psi_2 - \sin\psi_1\sin\psi_2)A_2(T_{23})\right], \tag{1c}$$

and

$$\mathscr{E}_x \propto \frac{1}{60j}[\sin(\psi_1+\psi_2)A_2(T_{23}) - 5\sin(\psi_1 - \psi_2)A_1(T_{23})], \tag{2a}$$

$$\mathscr{E}_y = 0, \tag{2b}$$

$$\mathscr{E}_z \propto \frac{1}{9j}\left[\cos(\psi_1-\psi_2)A_0(T_{23}) + \frac{1}{10}(2\cos\psi_1\cos\psi_2 - \sin\psi_1\sin\psi_2)A_2(T_{23})\right], \tag{2c}$$

where

$$A_\kappa(T_{23}) = \exp(-\gamma_a^{(\kappa)}T_{23}) + \exp(-\gamma_b^{(\kappa)}T_{23}), \quad \kappa = 0,1,2. \tag{3}$$

Equations (3) include the six relaxation rates $\gamma_i^{(\kappa)}$ ($\kappa = 0, 1, 2$, $i = a, b$):

$$\gamma_i^{(\kappa)} = \gamma_i^{(0)} + \Gamma_i^{(\kappa)}, \quad i = a,b. \tag{4}$$

Here the $\gamma_i^{(0)}$ ($i=a, b$) are the relaxation rates owing to inelastic collisions and the $\Gamma_i^{(\kappa)}$ ($i=a, b$) are the relaxation rates owing to elastic depolarizing collisions. $\Gamma_i^{(0)}=0$ ($i=a, b$) since elastic collisions cannot change the level populations. Collisions with a phase disruption can be neglected in the molecular gas SF_6.

In deriving Eqs. (1) and (2) it has been assumed that there is no dependence of the relaxation rates on the translational velocity of a resonant particle. This important condition was verified in our experiment (see below).

For infrared vibrational–rotational transitions the decay rates of the upper level a and lower level b are usually very close to one another, so the equations can be simplified. (Instead of six decay rates, we have only three.) A further simplification can be achieved by specially selecting the angles ψ_1 and ψ_2 so that only two of the decay rates have an effect on the amplitude of the photon echo.

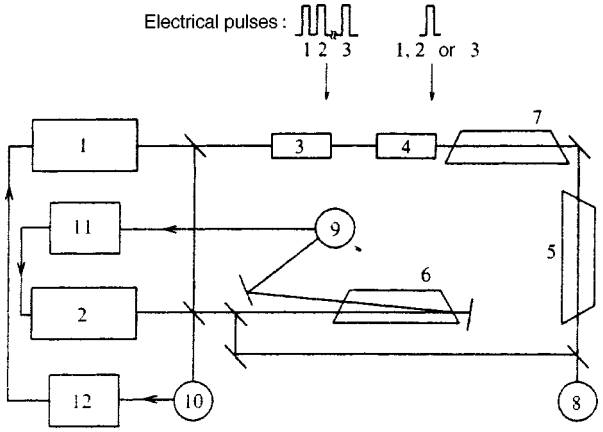

FIG. 1. The experimental setup: 1, 2 cw CO_2 lasers; 3, 4 electrooptical gates (the second gate has no polarizer at its output); 5–7 gas cells; 8–10 CdHgTe IR photodetectors; 11 frequency stabilization system; 12 frequency monitor.

2. EXPERIMENT

The characteristic feature of our experiment, the setup of which is sketched in Fig. 1, is the use of a cw laser 1 with a high quality output as an excitation source. The required sequence of linearly polarized driver pulses was shaped from the cw light by an electrooptical gate 3. If necessary, a second electrooptical crystal 4 could rotate the polarization of a given light pulse by ninety degrees.

An important feature of our experiments is the ability to excite a subensemble of resonant gas particles with definite longitudinal translational velocities. The width of such an ensemble of longitudinal velocities is determined by the pulse parameters (duration and intensity), while the central velocity can be chosen by tuning the frequency of the driver radiation over the Doppler profile of the transition being studied. This makes it possible to investigate collisional relaxation as a function of the translational velocities by studying coherent transient effects and extracting information on the relaxation mechanisms and on the steepness of the interaction potential from these data.[9,16]

In these experiments we use a cw CO_2 laser 2, whose frequency is stabilized at the center of the SF_6 transition, as a reference light source. The frequency of the driver laser 1 was coupled to that of the reference laser 2 with a fixed difference Ω (which determined the central velocity of the molecular subensemble which forms the coherent response: $v_y = \Omega/k$).

In an earlier paper,[17] the magnitude of the difference $\Gamma^{(2)} = \gamma^{(2)} - \gamma^{(0)}$ between the rates of relaxation of the alignment and populations of levels in the Q(38) transition of the ν_3 vibrational mode of gaseous SF_6 was measured. The rate of relaxation of the alignment $\gamma^{(2)}$ was obtained from the value of $\Gamma^{(2)}$ and from a population decay rate $\gamma^{(0)}$ given by other experiments.[18]

In this paper, experimental studies of the stimulated photon echo were conducted on the P(33) $\mathbf{A}_2^1$ vibrational–rotational line of the ν_3 mode of gaseous SF_6. Experiments were done with four different polarizations of the driver pulses. The polarizations of the first, second, and third pulses were rotated by an angle of 90° relative to the polarization of

the other two optical pulses, respectively, in cases (100), (010), and (001), respectively. The fourth case represented a situation in which all three pulses had the same linear polarization (case (000)).

In this way we obtained four different kinetic curves for the dependence of the stimulated photon echo signal on the delay time between the second and third driver pulses.

In cases (100) and (010) only the component $\mathscr{E}_x$ of the stimulated photon echo signal amplitude exists and only two decay rates, $\gamma^{(1)}$ and $\gamma^{(2)}$, contribute to this amplitude:

$$\mathscr{E}_{100} \propto 5\exp(-\gamma^{(1)}T_{23}) - \exp(-\gamma^{(2)}T_{23}), \tag{5}$$

and

$$\mathscr{E}_{010} \propto 5\exp(-\gamma^{(1)}T_{23}) + \exp(-\gamma^{(2)}T_{23}). \tag{6}$$

Equations (5) and (6) were obtained from Eq. (2a) with $\psi_1 = 90°$, $\psi_2 = 0$ and $\psi_1 = 0$, $\psi_2 = 90°$, respectively.

Since the stimulated photon echo signals were extremely weak, we used part of the emission from the reference laser 2 for optical heterodyning of the detected stimulated photon echo signal. Thus, the detector 8 recorded signals $\mathscr{S}_{100} \propto \mathscr{E}_{100}$ and $\mathscr{S}_{010} \propto \mathscr{E}_{010}$.

Similarly, in cases (001) and (000) we have only the component $\mathscr{E}_z$ of the stimulated photon echo signal, damped with relaxation constants $\gamma^{(0)}$ and $\gamma^{(2)}$ (recall that the z axis is assumed to coincide in direction with the polarization of the third driver pulse):

$$\mathscr{E}_{001} \propto 10\exp(-\gamma^{(0)}T_{23}) - \exp(-\gamma^{(2)}T_{23}), \tag{7}$$

and

$$\mathscr{E}_{000} \propto 5\exp(-\gamma^{(0)}T_{23}) + \exp(-\gamma^{(2)}T_{23}). \tag{8}$$

Equations (7) and (8) were obtained from Eq. (2c) with $\psi_1 = \psi_2 = 90°$ and $\psi_1 = \psi_2 = 0$, respectively.

Optical heterodyning was not used in cases (001) and (000), so we recorded signals $\mathscr{S}_{001} \propto \mathscr{E}_{001}^2$ and $\mathscr{S}_{000} \propto \mathscr{E}_{000}^2$.

3. RESULTS AND DISCUSSION

It is clear that $\gamma^{(1)}$ and $\gamma^{(2)}$ can be obtained by simply taking the logarithm of the sum or difference of the signal amplitudes $\mathscr{S}_{010}$ and $\mathscr{S}_{100}$ detected by the detector 8 as the polarizations of the first and second pulses are rotated. The resulting two sets of data, $\ln(\mathscr{S}_{010} + \mathscr{S}_{100})$ and $\ln(\mathscr{S}_{010} - \mathscr{S}_{100})$, for different values of T_{23} were approximated by linear functions of the form

$$\mathscr{T}(T_{23}) = \gamma^{(\kappa)}T_{23} + \text{const}, \quad \kappa = 1,2.$$

which were constructed by least squares. The results for pure SF_6 at gas pressures $P = (0.9\text{–}5.4)\cdot 10^{-3}$ Torr are shown in Fig. 2.

The other pair of relaxation rates, $\gamma^{(0)}$ and $\gamma^{(2)}$, can be obtained by mathematical analysis of the signals detected by the detector 8 in cases (001) and (000): the data were transformed into sets of the form $\ln(\sqrt{\mathscr{S}_{000}} + \sqrt{\mathscr{S}_{001}})$ and $\ln(\sqrt{\mathscr{S}_{000}} - 2\sqrt{\mathscr{S}_{001}})$ for different values of T_{23}, which were approximated by least squares as functions of the same form:

$$\mathscr{T}(T_{23}) = \gamma^{(\kappa)}T_{23} + \text{const}, \quad \kappa = 0.2.$$

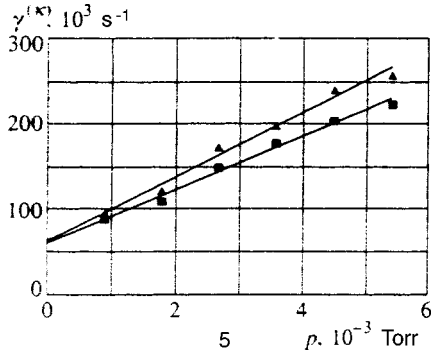

FIG. 2. Rates of relaxation of orientation $\gamma^{(1)}$ (squares) and alignment $\gamma^{(2)}$ (triangles) in SF_6 gas. The nonzero relaxation rates at zero pressure are explained by a transit-time effect.

The results for pure SF_6 are shown in Fig. 3 for the same gas pressures.

In pure SF_6 the collisional decay rate for the alignment was somewhat higher than the decay rates for the orientation and population. Our measurements yielded the following results (relaxation rate per unit pressure of SF_6): $\gamma^{(0)}/p = (28 \pm 3)\cdot 10^6$ s^{-1}Torr^{-1}, $\gamma^{(1)}/p = (32 \pm 3)\cdot 10^6$ s^{-1}Torr^{-1}, and $\gamma^{(2)}/p = (38 \pm 3)\cdot 10^6$ s^{-1}Torr^{-1}. Our value of $\gamma^{(0)}$ is comparable to that in Ref. 18, $\gamma^{(0)}/p = (30 \pm 5)\cdot 10^6$ s^{-1}Torr^{-1}. It is interesting that the elastic collision rates for depolarization (see Eq. (4)), $\Gamma^{(1)}/p \simeq 4\cdot 10^6$ s^{-1}Torr^{-1} and $\Gamma^{(2)}/p \simeq 10\cdot 10^6$ s^{-1}Torr^{-1}, were smaller than the relaxation rate owing to elastic collisions (collisions involving a change in the longitudinal translational velocity of the resonant particles) measured by the ordinary photon echo technique,[19] $\Gamma_{vcc}/p \sim 17\cdot 10^6$ s^{-1}Torr^{-1}. Of course, Γ_{vcc} was measured for a nondiagonal component of the matrix $\rho^{(ab)}$, while we obtained $\Gamma^{(1)}$ and $\Gamma^{(2)}$ for its diagonal components, but the physical significance of Γ_{vcc} (it is determined by elastic collisions, i.e., those in which only the translational velocity of the particles changes.) allows us to regard this constant as the same for all four components of the matrix $\rho^{(ab)}$. Thus, it is entirely appropriate to compare $\Gamma^{(1)}$ and $\Gamma^{(2)}$ with Γ_{vcc}. The inequality $\Gamma^{(1)} < \Gamma^{(2)} < \Gamma_{vcc}$ means that only some of the elastic collisions participate in the breakup of the multipole moments of the levels.. It is also evident that the magnetic moment induced by the radiation in a subensemble of reso-

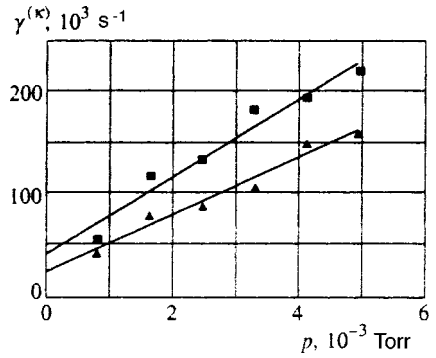

FIG. 3. Rates of relaxation of orientation $\gamma^{(0)}$ (triangles) and alignment $\gamma^{(2)}$ (squares) in SF_6 gas. The nonzero relaxation rates at zero pressure are explained by a transit-time effect.

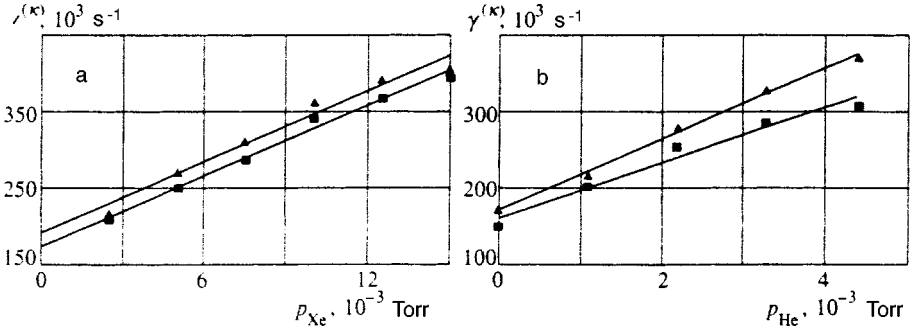

FIG. 4. Rates of relaxation of orientation $\gamma^{(1)}$ (squares) and alignment $\gamma^{(2)}$ (triangles) in SF_6+Xe (a) and SF_6+He (b) gas mixtures for an SF_6 pressure of $2.7\cdot10^{-3}$ Torr. Relaxation at zero pressures of Xe and He is determined by the transit-time effect and SF_6–SF_6 collisions.

nant particles was longer-lived than the electric quadrupole moment ($\Gamma^{(1)}<\Gamma^{(2)}$). It is interesting that the value of $\Gamma^{(2)}$ determined here for the P(33) A_2^1 level turned out to be noticeably smaller than the previously measured[17] value of the same quantity, $\Gamma^{(2)}$, for the Q(38) line. This result indicates that the relaxation of the multipole moments of these levels depends on j.

Measurements of $\gamma^{(1)}$ and $\gamma^{(2)}$ in gaseous mixtures of SF_6 with heavy (Xe) and light (He) buffers are shown in Fig. 4 and correspond to the following values of the relaxation rates (per unit pressure of buffer gas): $\gamma^{(1)}_{SF_6-Xe}/p_{Xe}=(16\pm2)\cdot10^6$ s^{-1}Torr^{-1}, $\gamma^{(2)}_{SF_6-Xe}/p_{Xe}=(16\pm2)\cdot10^6$ s^{-1}Torr^{-1}, $\gamma^{(1)}_{SF_6-He}/p_{He}=(36\pm4)\cdot10^6$ s^{-1}Torr^{-1}, and $\gamma^{(2)}_{SF_6-He}/p_{He}=(47\pm4)\cdot10^6$ s^{-1}Torr^{-1}.

For collisions of SF_6 molecules with the heavy buffer Xe, it appeared that $\gamma^{(1)}\approx\gamma^{(2)}$. In mixtures with He, we have $\gamma^{(1)}<\gamma^{(2)}$, as in pure SF_6.

The dependence of the constants $\gamma^{(\kappa)}$ on the translational velocities were measured for several values of the frequency detuning between the driver laser and the center of the SF_6 line over the range $\Delta\nu=0-9$ MHz. The value 9 MHz corresponds to $\sim1/2$ of the Doppler width ku in gaseous SF_6 at room temperature. The coherent responses were produced by a subensemble of particles with a width of $\sim0.1ku$. Both in pure SF_6 and in mixtures with light (He) and heavy (Xe) buffers, the decay rates of the polarization moments depend only weakly on the translational velocities of the colliding particles. For example, Fig. 5 shows measurement data on $\gamma^{(1)}$ and $\gamma^{(2)}$ in an SF_6+Xe mixture. It is clear that the slight enhancement in relaxation as the detuning from the absorption line center is increased does not exceed the experimental error. This is consistent with the approximation used in the theory.[11]

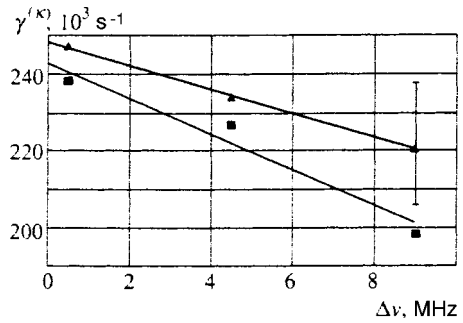

FIG. 5. Rates of relaxation of orientation $\gamma^{(1)}$ (squares) and alignment $\gamma^{(2)}$ (triangles) in a gaseous mixture of SF_6 ($p=2.5\cdot10^{-3}$ Torr) with Xe ($p_{Xe}=7.2\cdot10^{-3}$ Torr as functions of the frequency detuning $\Delta\nu$ between the driver radiation and the center of the P(33) A_2^1 line of the ν_3 mode. Here $\Delta\nu=9$ MHz corresponds to a longitudinal velocity for the molecules of $v_x\approx90$ m/s ($\sim1/2$ the Dopper width).

It is well known that the P(33) A_2^1 line of the ν_3 vibrational mode and the wings of ~40 lines in the series of ''hot'' vibrational–rotational bands lie in the same spectral region. Thus, in general, we must use a set of expressions similar to Eqs. (1) and (2), but with different coefficients that depend on j_a and j_b,[14] to obtain valid formulas for the amplitude of the stimulated photon echo instead of Eqs. (5)–(8). The qualitative behavior of these amplitudes (sum and different of the two exponents) would be the same, so, on the whole, the method for determining the relaxation constants does not raise doubts. However, the absence of an isolated spectrum line could lead us to some quantitative error in calculating the constants $\gamma^{(\kappa)}$, especially if the existence of a substantial dependence of the relaxation constants for the polarization moments on j is confirmed.

4. CONCLUSION

The basic results of this study of depolarizing collisions in the molecular gas SF_6 and its mixtures with buffer He and Xe are the following:

1. The dependence of the relaxation rates for the orientation and alignment on the translational velocities of the resonant particles is negligible in all three cases.

2. Only some of the collisions, which do not destroy the total population of the level, will destroy the multipole moments formed in the magnetic sublevels, and in pure gaseous SF_6 and in SF_6+He mixtures it is found that the magnetic moment (orientation) is longer lived than the electric quadrupole moment (alignment), while in SF_6+Xe mixtures both moments decay at the same rate.

3. We have found evidence that the relaxation rate for the alignment depends on j, the total angular momentum.

We thank the Russian Fund for Fundamental Research for financial support of our scientific group (Grants No. 95-02-04603 and No. 97-02-18496).

*)E-mail: rubtsova@isp.nsc.ru

[1] W. Hanle, Z. Phys. **30**, 93 (1924).
[2] A. Kastler, J. Phys. Radium **11**, 255 (1950).
[3] W. Happer, Rev. Mod. Phys. **44**, 169 (1972).

[4] R. L. Shoemaker, *Laser and Coherence Spectroscopy*, J. I. Steinfeld, Ed., Academic Press, N.Y. (1978).
[5] É. A. Manykin and V. V. Samartsev, *Optical Echo Spectroscpy [in Russian]*, Nauka, Moscow (1984).
[6] A. I. Alekseev and A. M. Basharov, Izv. Akad. Nauk SSSR, Ser. Fiz. **46**, 557 (1982).
[7] M. A. Gubin, I. V. Evseev, and V. A. Reshetov, FIAN Preprint No. 214, Moscow (1984).
[8] M. A. Gubin, I. V. Evseev, and V. M. Ermachenko, FIAN Preprint No. 7, Moscow (1985).
[9] L. S. Vasilenko, N. N. Rubtsova, and V. P. Chebotaev, JETP Lett. **38**, 474 (1983).
[10] D. S. Bakaev, I. V. Evseev, and V. M. Ermachenko, Zh. Éksp. Teor. Fiz. **76**, 1212 (1979) [Sov. Phys. JETP **49**, 615 (1979)].
[11] I. V. Evseev, V. M. Ermachenko, and V. S. Samarchenko, *Depolarizing Collisions in Quantum Electrodynamics [in Russian]*, Nauka, Moscow (1992).
[12] A. I. Alekseev, A. M. Basharov, and V. N. Beloborodov, Zh. Éksp. Teor. Fiz. **79**, 787 (1980) [Sov. Phys. JETP **52**, 401 (1980)].
[13] A. I. Alekseev and A. M. Basharov, Opt. Spektrosk. **52**, 220 (1982) [Opt. Spectrosc. (USSR) **52**, 132 (1982)].
[14] I. V. Evseev, V. M. Ermachenko, and V. A. Reshetov, Zh. Éksp. Teor. Fiz. **78**, 2213 (1980) [Sov. Phys. JETP **51**, 1108 (1980)].
[15] L. S. Vasilenko and N. N. Rubtsova, Kvant. Élektronika. **9**, 2243 (1982) [Sov. J. Quantum Electron. **12**, 1457 (1982)].
[16] L. S. Vasilenko, N. N. Rubtsova, and E. V. Khvorostov, Laser Phys. **6**, 165 (1996).
[17] N. S. Belousov, L. S. Vasilenko, I. D. Matveenko, and N. N. Rubtsova, Opt. Spektrosk. **63**, 34 (1987) [Opt. Spectrosc. (USSR) **34**, 19 (1987)].
[18] L. S. Vasilenko and N. N. Rubtsova, Opt. Spektrosk. **58**, 697 (1985) [Opt. Spectrosc. **58**, 422 (1985)].
[19] B. Comaskey, R. E. Scotti, and R. L. Shoemaker, Opt. Lett. **6**, 45 (1981).

Translated by D. H. McNeill

Two-step excitation of an auto-ionized state of the Ba atom associated with two-photon excitation of an intermediate state

A. Yu. Elizarov*)

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted 23 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 834–840 (March 1998)

An experimental and theoretical study is performed of the angular photoelectron distribution for three-photon ionization of Ba atoms through the 2ω-excited intermediate state $6p^2(^1S_0)$ and the auto-ionized state $6p8s(^3P_1)$. Rotation of the polarization plane of dye-laser radiation allowed us to investigate the photoelectron angular distribution. Electrons were counted with the help of a time-of-flight electron spectrometer. The density-matrix formalism is used to obtain expressions for the angular dependence of the differential ionization probability. Possible experiments are discussed. © *1998 American Institute of Physics.* [S1063-7761(98)00603-9]

1. INTRODUCTION

The study of the photoelectron angular distribution for stepwise ionization of atoms by polarized laser radiation from an excited oriented state is a branch of polarization laser spectroscopy and has been pursued experimentally and theoretically for more than 20 years. This method allows one to obtain unique information about various (including auto-ionized) states of atoms. The use of tunable lasers and synchrotron radiation opens up new experimental possibilities, for example, the possibility of studying auto-ionized states lying in the VUV region of the spectrum.[1–3]

Auto-ionized states can be excited by laser radiation in the visible region of the spectrum with the help of multiphoton resonant or nonresonant excitation processes. An additional advantage of two-photon processes over two-step ones arises for intermediate states with nonzero total angular momentum, when the orientation of atoms in intermediate states can be destroyed as a result of processes of radiation effusion and collisional depolarization. Thus, for example, depolarization of atoms in the intermediate state $6s6p(^1P_1)$ during stepwise excitation by polarized radiation of Ba atoms was observed for atomic densities in the beam greater than 10^{11} cm^{-3} (Ref. 4).

The influence of depolarization processes is greatly diminished if a two-photon process is used to excite the intermediate state. In this case radiative relaxation to the ground state is forbidden by the selection rules, as a result of which radiation effusion and collisional depolarization have practically no effect on the orientation of the atoms in the intermediate state for beam atomic densities up to 10^{13} cm^{-3} (Ref. 5), which provides experimental advantages in high-resolution studies of angular distributions.

The excitation of auto-ionized states by polarized radiation was first described theoretically in Ref. 6, where the wave-function formalism was used. The use of density-matrix methods has made it possible to substantially simplify the summation over unobservable projections of the angular momenta. This method has proved to be especially convenient for describing excitation of auto-ionized states of polarized atoms.[7–10] Preliminary polarization of the target atom opens up additional possibilities in the design of a complete quantum-mechanical photo-ionization experiment.[9]

In this work we consider the photoelectron angular distribution for resonant two-step ionization of auto-ionized states of Ba atoms with the configuration $6p8s(^3P_1)$ through the the 2ω-excited intermediate state $6p^2(^1S_0)$.

2. THEORY

For two-step ionization by polarized radiation the photoelectron angular distribution can be represented analytically in the form[11]

$$\frac{dW}{d\Omega} \propto \mathrm{Tr}(\varepsilon_i \varepsilon_e (R \rho_a \rho^{\gamma} R^{+})), \tag{1}$$

where dW is the probability of emitting a photoelectron within the solid angle $d\Omega$, ρ_a is the density matrix of the atoms in the intermediate state, ρ^{γ} is the density matrix of the ionizing radiation, ε_i and ε_e are the detection efficiency matrices of the ions and electrons, and R is the radiative interaction operator describing the transition in the presence of electromagnetic radiation between the intermediate and final state.

A multipole of the 2ω-excited intermediate state can be represented as a superposition of excitation channels through the intermediate states with total angular momentum J_a (Ref. 12):

$$\rho^{(2)}_{K_1 Q_1} = \sum_{J_a} \langle J_a \| R \| J_0 \rangle \langle J_1 \| R \| J_a \rangle \rho_{K_0 Q_0} \rho_{k_1 q_1} \langle J_a \| R \| J_0 \rangle^* \times \rho_{k_1' q_1'} \langle J_1 \| R \| J_a \rangle^* (E - E_{J_a} - \hbar\omega)^{-1}, \tag{2}$$

where $|J_0\rangle$ and E are the initial state and its energy, $|J_a\rangle$ and E_{J_a} are the intermediate state of two-photon excitation and its energy, $|J_1\rangle$ is the final state of the two-photon excitation process, $\langle J_a \| R \| J_0 \rangle$ is the reduced matrix element of the radiative interaction, $\rho_{K_0 Q_0}$ is the density matrix of the ground

state of the atom, $\rho_{k_1q_1}$ is the density matrix of the electromagnetic radiation of the first step, and $\hbar\omega$ is the energy of the electromagnetic radiation. The sum over J_a follows from the well-known formula of vector manifolds:

$$(\overline{e_s},\overline{e_r})=\sum_m (\overline{e_s},\overline{e_m})(\overline{e_m},\overline{e_r}). \tag{3}$$

In the description of the auto-ionization process we will employ the formalism of irreducible tensor operators. Following Refs. 11, 13, and 14, we write the state multipoles of the photons and atoms in the ground state, intermediate state, and final state.

a) The polarization density matrix of the electromagnetic radiation was obtained in Ref. 13. Employing the explicit expression for it, we write the state multipole of the dipole radiation

$$\rho^{\gamma}_{k_iq_i}(1,1)=\sum_{\lambda,\lambda'}(-1)^{1-\lambda'}(2k+1)^{1/2}\times\begin{pmatrix}1 & 1 & k_1\\ \lambda & \lambda' & -q_1\end{pmatrix}(\lambda|\rho|\lambda'), \tag{4}$$

where i is the number of the excitation step, $\left(\begin{smallmatrix}1 & 1 & k_1\\ \lambda & \lambda' & -q_1\end{smallmatrix}\right)$ is the $3j$-symbol, $(\lambda|\rho|\lambda')$ is the photon density matrix, expressed in its usual form in terms of the Stokes parameters ξ_j:

$$\langle\lambda|\rho|\lambda'\rangle=(1/2)\begin{pmatrix}1+\xi_2 & -\xi_3+i\xi_1\\ -\xi_3-i\xi_1 & 1-\xi_2\end{pmatrix}. \tag{5}$$

b) For atoms in the isotropically oriented ground state the state multipole has the following form:[11]

$$\rho_{K_0Q_0}(\gamma_0J_0,\gamma_0'J_0')=\hat{J}_0^{-1}\delta_{K0}\delta_{Q0}\delta_{J_0J_0'}, \tag{6}$$

where $\hat{J}\equiv(2J+1)^{1/2}$ holds for arbitrary angular momentum, J_0 is the total angular momentum of the initial state, and $\gamma_{(j)}$ are the remaining quantum numbers needed to describe the jth state of the atom.

c) Employing expression (2), we write an expression for the state multipole of the atom in the oriented intermediate state $|\gamma_1J_1\rangle$:

$$\begin{aligned}\rho^{(2)}_{K_1Q_1}(\gamma_1J_1,\gamma_1J_1)=&\sum_{J_a}(E_0-E_{J_a}+\hbar\omega)^{-1}\\ &+\langle\gamma_aJ_a\|R\|\gamma_0J_0,\gamma_aJ_a\rangle\\ &\times\langle\gamma_0J_0,\gamma_aJ_a\|R\|\gamma_aJ_a\rangle\\ &\times\langle\gamma_1J_1\|R\|\gamma_aJ_a\gamma_1J_1\rangle\\ &\times\langle\gamma_aJ_a,\gamma_1J_1\|R\|\gamma_1J_1\rangle\\ &\times\sum_{K_0,k_1}\hat{J}_a^2\hat{k}_1\hat{K}_0(K_aQ_a|K_0Q_0k_1q_1)\\ &\times\begin{Bmatrix}J_0 & J_0 & K_0\\ 1 & 1 & k_1\\ J_a & J_a & K_a\end{Bmatrix}\rho_{K_0Q_0}(\gamma_0J_0,\gamma_0J_0)\\ &\times D^{k_1*}_{Q_0q_1}(\hat{A})\rho^{\gamma}_{k_1q_1}(1,1)\\ &\times\sum_{K_a,k_1'}\hat{J}_1^2\hat{k}_1'\hat{K}_a(K_1Q_1|K_aQ_ak_1'q_1)\\ &\times\begin{Bmatrix}J_a & J_a & K_a\\ 1 & 1 & k_1'\\ J_1 & J_1 & K_1\end{Bmatrix}D^{k_1'*}_{Q_aq_1}(\hat{A})\rho^{\gamma}_{k_1'q_1}(1,1),\end{aligned} \tag{7}$$

where $(K_aQ_a|K_0Q_0k_1q_1)$ is the Clebsch–Gordan coefficient, $\{\ldots\}$ is the $9j$-symbol, and $D^{k_1'*}_{Q_aq_1}(\hat{A})$ is the matrix of finite rotations[15] carrying the treatment of the excitation process from the laboratory coordinate system to the atomic coordinate system defined by the polarization vector of the atom, in which the density matrix is diagonal.

d) We represent the multipole of the auto-ionized state of the atom in the form

$$\rho_{K_2Q_2}(\gamma_2J_2,\gamma_2J_2)=\langle\gamma_1J_1,\gamma_2J_2\|R\|\gamma_2J_2\rangle\times\langle\gamma_2J_2\|R\|\gamma_2J_2,\gamma_1J_1\rangle\rho'_{K_2Q_2}(J_2,J_2),$$

$$\begin{aligned}\rho'_{K_2Q_2}(J_2,J_2)=&\sum_{K_1k_2}\hat{J}_2^2\hat{k}_2\hat{K}_1(K_2Q_2|K_1Q_1k_2q_2)\\ &\times\begin{Bmatrix}J_1 & J_1 & K_1\\ 1 & 1 & k_2\\ J_2 & J_2 & K_2\end{Bmatrix}\rho^{(2)}_{K_1Q_1}(\gamma_1J_1,\gamma_1J_1)\\ &\times D^{k_2*}_{Q_1q_2}(\hat{A})\rho^{\gamma}_{k_2q_2}(1,1),\end{aligned} \tag{8}$$

where $|\gamma_2J_2\rangle$ is the auto-ionized state with total angular momentum J_2 and $\rho^{\gamma}_{k_2q_2}(1,1)$ is the state multipole of the electromagnetic radiation of the second step.

Since in the nonrelativistic approximation the ionization probability does not depend on the electron spin and the polarization characteristics of the states of the electron and ion are not fixed, the expression for the product of the state multipoles of the ion and photoelectron, arising as a result of decay of the auto-ionized state, can be represented in the form

$$\begin{aligned}&\rho_{K_iQ_i}(\gamma_iJ_i,\gamma_iJ_i)\rho_{k_eq_e}(lj,l'j')\\ &\quad=\frac{1}{4\pi}\langle lj,\gamma_iJ_i\|V\|\gamma_2J_2\rangle\\ &\qquad\times\langle\gamma_2J_2\|V\|l'j',\gamma_iJ_i\rangle\sum_{K_2}\rho_{K_2Q_2}(\gamma_2J_2,\gamma_2J_2)\\ &\qquad\times(-1)^{J_i+J_2+K_2+j}\hat{J}_2^2\hat{J}_i^{-1}\begin{Bmatrix}J_2 & j & J_i\\ j' & J_2 & K_2\end{Bmatrix},\end{aligned} \tag{9}$$

where $|\gamma_iJ_i\rangle$ is the state of the atom with one electron knocked out, V is the Coulomb interaction operator, and $\{...\}$ is the $6j$-symbol.

When the state multipole formalism is used, expression (1) transforms to[14]

$$\frac{dW}{d\Omega}=\sum \rho_{K_iQ_i}(\gamma_iJ_i,\gamma_i,J_i)\rho_{k_eq_e}(lj,l'j') \times\varepsilon^*_{K_iQ_i}(\gamma_iJ_i,\gamma_iJ_i)\varepsilon^*_{k_eq_e}(lj,l'j'). \quad (10)$$

where the sum is over all repeated indices, $\varepsilon^*_{K_iQ_i}(\gamma_iJ_i,\gamma_iJ_i)$ and $\varepsilon^*_{k_eq_e}(lj,l'j')$ are the detection efficiency tensors of the ions and electrons, respectively. In the case when the polarization characteristics of the state of the ion are not fixed, $\varepsilon_{K_iQ_i}(\gamma_iJ_i,\gamma_iJ_i)=\hat{J}_i\delta_{K_i0}\delta_{Q_i0}$ (Ref. 14) and the expression for the product of the detection efficiency tensors of the ions and the electrons has the form[14]

$$\varepsilon_{K_iQ_i}(\gamma_iJ_i,\gamma_iJ_i)\varepsilon_{k_eq_e}(lj,l'j') =\frac{1}{4\pi}\sum_{k_e}\hat{J}_i(-1)^{1/2-j}\overline{Z}\left(lj,l'j';\frac{1}{2}k_e\right)D^{k_e}_{q_e0}(R), \quad (11)$$

where $\overline{Z}(lj,l'j';\frac{1}{2}k_e)$ is the Huby function.[14]

Substituting Eqs. (8), (9), and (11) in Eq. (10), we obtain an expression for $dW/d\Omega$ (here we use the atomic system of units)

$$\frac{dW}{d\Omega}=2\pi(P2\pi\omega\alpha)^3\sum\langle\gamma_1J_1,\gamma_2J_2\|R\|\gamma_2J_2\rangle \times\langle\gamma_2J_2\|R\|\gamma_2J_2,\gamma_1J_1\rangle\frac{1}{4\pi}\langle lj,\gamma_iJ_i\|V\|\gamma_2J_2\rangle \times\langle\gamma_2J_2\|V\|l'j',\gamma_iJ_i\rangle\mathscr{K}(l,l',j,j',J_2,J_i,J_a,K_2,k_e),$$

where the sum is over l,l',j,j',J_2,J_i,J_a; P is the photon flux density, α is the fine-structure constant, and

$$\mathscr{K}(l,l',j,j',J_2,J_i,J_a,K_2,k_e) \equiv\hat{J}_2^2\sum_{K_2}\rho'_{K_20}(J_2,J_2) \times(-1)^{J_i+J_2+K_2+j'}\begin{Bmatrix}J_2 & j & J_i\\ j' & J_2 & K_2\end{Bmatrix} \times\sum_{k_e}(-1)^{1/2-j}\overline{Z}\left(lj,l'j';\frac{1}{2}k_e\right)D^{k_e}_{00}(\hat{R}^{-1}\hat{A}). \quad (12)$$

The necessity of summing over K_2 in the case of excitation of an auto-ionized state was first shown in Ref. 16. For excitation of such states from an isotropically oriented intermediate state, let us consider the case in which the coordinate system is related to the polarization plane so that $Q_2=q_e=0$. By virtue of the properties of the $6j$-symbol, expression (12) is nonzero for $K_2=0,2$ and transforms to[17]

$$W(\theta_e)=\frac{\sigma_s}{4\pi}[1+\beta\mathbb{P}_2(\cos\theta_e)], \quad (13)$$

where σ_s is the total cross section, β is the anisotropy parameter of the angular distribution, $\mathbb{P}_2(\cos\theta_e)$ is the Legendre polynomial of degree 2, and θ_e is the angle between the axis of the electron energy analyzer and the direction of polarization of the laser radiation.

3. EXPERIMENTAL SETUP

The time-of-flight electron energy analyzer was specially developed and prepared for recording the electron angular distribution and discriminating their energies in multiphoton ionization of atoms. Details of the experimental setup are given in Ref. 18; here we describe briefly its operation. The energy of the recorded electrons is less than 3 eV. The electron-optical circuit of the spectrometer employs a weak accelerating electrostatic field following a short interval of field-free electron drift during which ionization takes place. This design was chosen so that in the interaction region of the radiation and atoms electrostatic fields distorting the angular dependence of the electron distribution are absent. The Ba beam atomic density was varied in the range 10^{10}–10^{12} at/cm^3.

Ba atoms were excited and analyzed by pulsed dye-laser radiation oriented perpendicular to the direction of the Ba atomic beam. The dye-laser had the following parameters: spectral width of the lasing line 2 cm^{-1}, frequency scanning region 570–590 nm, pulse energy 1.5 mJ, pulse duration 20 ns. The degree of linear polarization was 98%. The interaction region of the atoms with the laser radiation was screened from the Earth's magnetic field with the help of a triple permalloy screen, where the thickness of each screen was 2 mm. The residual magnetic field in the electron drift region was 2 mG.

4. DISCUSSION

Three-photon ionization of the Ba atoms was realized according to the following scheme:[18]

$$\mathrm{Ba}[6s^2(^1S_0)]\xrightarrow{2\omega(581.9\ \mathrm{nm})}\mathrm{Ba}^*[6p^2(^1S_0)],$$
$$\mathrm{Ba}^*[6p^2(^1S_0)]\xrightarrow{\omega(581.9\ \mathrm{nm})}\mathrm{Ba}^*[6p8s(^3P_1)],$$
$$\mathrm{Ba}^*[6p8s(^3P_1)]\rightarrow\begin{bmatrix}\mathrm{Ba}^+6s(^2S_{1/2})\\ e^-[\varepsilon p_{1/2,3/2}]\end{bmatrix}. \quad (14)$$

The energy of the continuum corresponding to three-photon ionization coincided with the position of the auto-ionized state $6p8s(^3P_1)$.

By rotating the polarization plane of the laser radiation we measured the distribution of the photoelectrons corresponding to the $6s$ state of the ion (see Fig. 1). By virtue of the isotropic orientation of the $6p^2(^1S_0)$, whose choice was dictated by the experimental capabilities of the setup, the analytical expression for the electron angular distribution has the form (13).

For the electrons corresponding to the $6s$ state of Ba^+ we recorded the photoelectron angular distribution for different Ba atomic densities in the interaction region with the laser radiation. Expression (13) was used to fit the experimental dependence of the electron signal intensity on the rotation angle of the polarization plane of the radiation. The fit yielded the parameter value $\beta_2=0.74\pm0.1$, which remains within the limits of experimental error up to Ba atomic densities of $\sim5.0\times10^{-12}$ cm^{-3}.

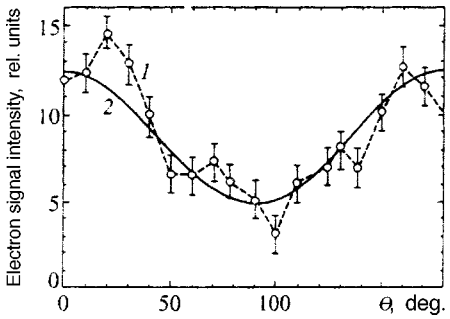

FIG. 1. Form of the dependence of the electron signal intensity on the rotation angle of the polarization plane of the laser radiation (curve *1*) 0° corresponds to collinear arrangement of the axis of the electron spectrometer relative to the polarization vector of the laser radiation. Curve *2* is a fit to the experimental curve using expression (13).

The main contribution to the probability of two-photon excitation comes from the excitation channel with participation of the state that is closest in energy to the virtual level. In Ref. 5 it was shown that for 2ω-excitation of the $6p^2(^1S_0)$ state the main contribution to the probability of two-photon excitation comes from the channel involving the participation of the resonant state $6s6p(^1P_1)$, which stands off from the virtual level of two-photon excitation by 808 cm^{-1}. The next selection-rule-allowed state stands off 13695 cm^{-1} from the virtual level. Therefore it is possible to limit the sum over J_a to the one term corresponding to the $6s6p(^1P_1)$ state.

For the case of the intermediate state $6p^2(^1S_0)$ ($J_2=0$), expression (12) simplifies substantially: $K_2=k_e$ and the numerical values for $\mathscr{K}(1,1,j,j',1,1,1/2,1,K_2,K_2) \equiv \mathscr{K}(j,j',K_2)$ are nonzero only for $K_2=0$ and 2 (see Table I). Comparison with experiment is rendered difficult because we have not calculated the Coulomb and dipole matrix elements.

5. CONCLUSION

We have considered the problem of the electron angular distribution in three-photon ionization of Ba atoms. As the intermediate states we used a discrete state and an autoionized state. Excitation of an even state with the help of a two-photon process ensures minimal depolarization of the atoms in the intermediate state. This fact is especially important for the accuracy of the results of the complete quantum-mechanical experiment on photoionization of polarized atoms,[9] which records the angular distribution of the electrons. From the point of view of experimental realization this variant of the setup of the complete experiment is the most convenient since it does not require determination of the spin state of the photoelectrons, which is a complicated experimental task. We hope that the method proposed here for decreasing the influence of depolarization of the atoms in the intermediate state on the recording of the photoelectron angular distribution will be of interest to experimentalists.

TABLE I.

j,j'	$\mathscr{K}(j,j',0)\times 1$	$\mathscr{K}(j,j',2)\times \mathrm{P}_2(\cos\theta)$
1/2, 1/2	−1.197	0
1/2, 3/2	0	$-0.219\mathrm{P}_2(\cos\theta_e)$
3/2, 3/2	−0.423	$0.155\mathrm{P}_2(\cos\theta_e)$

ACKNOWLEDGMENTS

This work was carried out with the partial support of the Russian Fund for Fundamental Research and the program ''Optics: Laser Physics.''

*)E-mail: elizarov@lpr.ioffe.rssi.ru

[1] M. Meyer, B. Müller, A. Nunnemann, Th. Prescher, E. von Raven, M. Richter, M. Schmidt, B. Sonntag, and P. Zimmermann, Phys. Rev. Lett. **59**, 2963 (1987).
[2] A. B. Fedotov, O. S. Ilyasov, N. I. Koroteev, and A. M. Zheltikov, Nuovo Cimento **14**, 1003 (1992).
[3] S. Baier, A. N. Grum-Grzhimailo, and N. M. Kabachnik, J. Phys. B: At. Mol. Opt. Phys. **27**, 3363 (1994).
[4] A. Yu. Elizarov and N. A. Cherepkov, Zh. Éksp. Teor. Fiz. **96** 1224 (1989) [Sov. Phys. JETP **69**, 695 (1989)].
[5] N. A. Cherepkov and A. Yu. Elizarov, J. Phys. B: At. Mol. Opt. Phys. **24**, 4169 (1991).
[6] B. Cleff and W. Mehlhorn, J. Phys. B: At. Mol. Opt. Phys. **7**, 593 (1974).
[7] N. M. Kabachnik and I. P. Sazhina, J. Phys. B: At. Mol. Opt. Phys. **9**, 1681 (1976).
[8] S. V. Bobashev, A. Yu. Elizarov, V. K. Prilipko, and N. A. Cherepkov, Laser Phys. **3**, 751 (1993).
[9] H. Klar, J. Phys. B: At. Mol. Opt. Phys. **13**, 4741 (1980).
[10] H. Klar and H. Kleinpoppen, J. Phys. B: At. Mol. Opt. Phys. **15**, 933 (1982).
[11] S. Devons and L. J. Goldfarb, *Handbuch der Physik* (Springer-Verlag, Berlin, 1957), p. 362.
[12] M. Goeppert-Mayer, Ann. Phys. (Leipzig) **9**, 273 (1931).
[13] M. Peshkin, Adv. Chem. Phys. **18**, 1 (1971).
[14] A. J. Ferguson, *Angular Momentum Methods in Gamma-Ray Spectroscopy* (North-Holland, Amsterdam, 1965), p. 246.
[15] D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988).
[16] E. G. Berezhko and N. M. Kabachnik, J. Phys. B: At. Mol. Opt. Phys. **10**, 2467 (1977).
[17] C. N. Yang, Phys. Rev. **74**, 764 (1948).
[18] A. Yu. Elizarov, JETP Lett. **62**, 23 (1996).

Translated by Paul F. Schippnick

Interaction of radiation and a relativistic electron in motion in a constant magnetic field

G. M. Filippov*)

I. N. Ul'yanov Chuvash State University, 428015 Cheboksary, Russia
(Submitted 27 May 1997)
Zh. Éksp. Teor. Fiz. **113**, 841–864 (March 1998)

This paper examines the effect of multiple photon emission on the quantum mechanical state of an electron emitting synchrotron radiation and on the intensity of that radiation. Calculations are done with a variant of perturbation theory based on the use of extended coherent states. A general formula is derived for the number of emitted photons, which allows taking into account their mutual interaction. A model problem is used to demonstrate the absence of the infrared catastrophe in the modified perturbation theory. Finally, the electron density matrix is calculated, and the analysis of this matrix makes it possible to conclude that the degree of the electron's spatial localization increases with the passage of time if the electron is being accelerated. © *1998 American Institute of Physics.* [S1063-7761(98)00703-3]

1. INTRODUCTION

The effect of radiation on the paths of charged particles in a synchrotron has already been analyzed (see, e.g., Ref. 1 and the literature cited therein). The analysis is based on the classical Lorentz–Dirac equation or on the solution of a kinetic equation whose coefficients are the probabilities of quantum transitions between various stationary states of an electron moving in the magnetic field of the synchrotron. Calculations have shown that in the absence of focusing in the magnetic field there is an increase in the radial fluctuations of the electron path and an increase in longitudinal fluctuations of the electron's momentum with the passage of time. In a focusing magnetic field, in the initial stages of electron acceleration, the presence of radiation leads to radiative damping, which damps radial and vertical oscillations (the so-called radiative damping effect). Lately, research has focused on the analysis of equations of the Lorentz–Dirac type in problems not necessarily related to synchrotron radiation (see, e.g., Refs. 2–4). New solutions of Lorentz–Dirac equations have been found for some special cases, and additional arguments from quantum electrodynamics are invoked to eliminate nonphysical solutions.

Despite the indisputable value of the results obtained by solving Lorentz–Dirac equations, it must be noted that some important properties of the states of a particle emitting radiation, properties that are not directly related to the path and do not directly influence the above effects—like an increase in radial fluctuations of the path—are excluded from these results. This is true, in particular, of the evolution of the particle's wave packet, which affects the radiation and hence the radiative friction and the path.

The present paper is an attempt to use a modified perturbation-theory approach to examine the effect of multiple photon emission on the evolution of the wave packet of a particle, in particular when the particle emits synchrotron radiation. Most papers devoted to the quantum mechanical theory of synchrotron radiation ignore this aspect. It is usually assumed that the particle emitting radiation has a wave function given by the solution of the Dirac equation. However, if the emitted radiation is taken into account, the particle is only a part of the quantum mechanical system and its state cannot be described with the completeness that is possible in principle in quantum theory.

A common approach to describing the states of particles that are members of a large system is to use the concept of the density matrix. This paper demonstrates that the evolution of the density matrix suggests that a particle goes with the passage of time into states that are more and more localized, with the particle motion described by the laws of classical mechanics with ever-increasing accuracy. Thus we have additional support for the validity of using the Lorentz–Dirac equation along with a clearer understanding of the incompleteness of the physical picture described by this equation.

The problem of the structure of the wave packets of emitting particles is related to the classical model of a distributed electron studied by Lorentz (see, e.g., Ref. 5). In quantum electrodynamics this model leads to the well-known problem of ultraviolet divergence, encountered in the calculations of the mass, charge, and energy of an elementary particle. Furthermore, the renormalization of charge in quantum electrodynamics reveals the internal inconsistency of the traditional Feynman formulation of perturbation theory (see, e.g., Ref. 6). It would be useful to follow the changes in the difficulties encountered by classical electrodynamics initiated by changes in the perturbation theory, to establish which of the above problems is invariant, so to say. It might turn out that in the modified theory some of these problems can be resolved without resorting to additional hypotheses. This might then lead to a new direction in the development of quantum electrodynamics and the theory of quantized fields in general. The present paper uses a model to show that at least in relation to the infrared catastrophe, the adopted modification in the theory does not lead to problems characteristic of the traditional form of the theory. Calculations are based on a general formula that describes the mutual interaction of the emitted photons as a manifestation of the non-

linearity inherent in quantum electrodynamics.

2. EMISSION OF PHOTONS BY A CLASSICAL CHARGED PARTICLE

We write the Hamiltonian describing the interaction of a free electromagnetic field and a particle carrying an electric charge Z (here we use atomic units: $\hbar=1$ and $|e|=1$):

$$\hat{H}_{\rm int}=-\frac{1}{c}\int \mathbf{j}\cdot\hat{\mathbf{A}}\,dV, \tag{1}$$

where the current density vector $\mathbf{j}$ is a function of coordinates and time. The vector potential operator is specified in a three-dimensional transverse gauge,

$$\hat{\mathbf{A}}=\sum_{\alpha,\mathbf{q}} g_q\{\hat{f}_{\alpha\mathbf{q}}\mathbf{e}_{\alpha\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{r}}+\hat{f}^{\dagger}_{\alpha\mathbf{q}}\mathbf{e}^{*}_{\alpha\mathbf{q}}e^{-i\mathbf{q}\cdot\mathbf{r}}\}, \tag{2}$$

as a standard linear form in the creation and annihilation operators ($\hat{f}^{\dagger}_{\alpha\mathbf{q}}$ and $\hat{f}_{\alpha\mathbf{q}}$) for photons in states with polarization α ($\alpha=1,2$), momentum $\mathbf{q}$, and energy $\omega=cq$. The polarization vectors $\mathbf{e}_{\alpha\mathbf{q}}$ have unit length and are orthogonal to $\mathbf{q}$. The coupling constants $g_q=(2\pi c^2/\omega\Omega)^{1/2}$ contain the normalization volume Ω, which does not enter into the final expressions and thus can be put equal to unity.

The total Hamiltonian is the sum of the free-photon Hamiltonian

$$\hat{H}_0=\sum_{\alpha,\mathbf{q}}\omega\,\hat{f}^{\dagger}_{\alpha\mathbf{q}}\hat{f}_{\alpha\mathbf{q}}+\text{const}$$

and the Hamiltonian (1). We pass to the interaction picture for field operators; for example,

$$\hat{A}(t)=e^{i\hat{H}_0t}\hat{A}e^{-i\hat{H}_0t}.$$

The equation describing the evolution of the state vector of the photon field, $|t)$ in the interaction picture,

$$i\frac{d}{dt}|t)=\hat{H}_{\rm int}(t)|t), \tag{3}$$

has in the given case an exact solution in the form of the direct product of vectors of photon coherent states,

$$|t)=\prod_{\alpha,\mathbf{q}}\exp[-i\chi_{\alpha\mathbf{q}}-\hat{f}_{\alpha\mathbf{q}}Q^{*}_{\alpha\mathbf{q}}+\hat{f}^{\dagger}_{\alpha\mathbf{q}}Q_{\alpha\mathbf{q}}]|t_0), \tag{4}$$

where the initial state vector coincides, to within an arbitrary phase factor, with the vacuum state of the photon field:

$$|t_0)=e^{i\phi_0}|\text{vac}),\quad \phi_0=\text{const},$$

$$Q_{\alpha\mathbf{q}}(t)=i\frac{g_q}{c}\int_{t_0}^{t}dt'\,\mathbf{e}^{*}_{\alpha\mathbf{q}}\cdot\mathbf{j}_{\mathbf{q}}(t')e^{i\omega t'},$$

$$\chi_{\alpha\mathbf{q}}(t)=\int_{t_0}^{t}\text{Im}[Q_{\alpha\mathbf{q}}(t')\dot{Q}^{*}_{\alpha\mathbf{q}}(t')]dt',$$

and $\mathbf{j}_{\mathbf{q}}(t)$ is the Fourier transform of the current density.

Using the exact solution (4), we can calculate all quantities of interest. For instance, the mean number of photons created by time t is given by

$$n^{(0)}_{\alpha\mathbf{q}}(t)=|Q_{\alpha\mathbf{q}}(t)|^2. \tag{5}$$

Note that this formula yields the mean number of emitted photons only as $t\to\infty$, since creation of a photon requires a time interval c/q long, which tends to infinity as $q\to0$. In what follows we use this interpretation of $n_{\alpha\mathbf{q}}$.

Suppose that photons are emitted by a point particle carrying electric charge Z and moving along a path $\mathbf{r}=\mathbf{r}_0(t)$. Then

$$\mathbf{j}_{\mathbf{q}}(t)=Z\mathbf{v}_0(t)\exp[-i\mathbf{q}\cdot\mathbf{r}_0(t)],$$

where $\mathbf{v}_0(t)=\dot{\mathbf{r}}_0(t)$. After summing over polarizations we can reduce the time derivative of the number of photons (as $t_0\to\infty$) to the form

$$\begin{aligned}\frac{d}{dt}\sum_{\alpha=1,2}n^{(0)}_{\alpha\mathbf{q}}(t)=\frac{Z^2}{c^2}g_q^2\int_{-\infty}^{\infty}d\tau\Bigg[&\mathbf{v}_0\left(t-\frac{|\tau|}{2}+\frac{\tau}{2}\right)\\&\times\mathbf{v}_0\left(t-\frac{|\tau|}{2}-\frac{\tau}{2}\right)-\frac{1}{q^2}\left(\mathbf{q}\cdot\mathbf{v}_0\left(t-\frac{|\tau|}{2}+\frac{\tau}{2}\right)\right)\left(\mathbf{q}\cdot\mathbf{v}_0\left(t-\frac{|\tau|}{2}-\frac{\tau}{2}\right)\right)\Bigg]\\&\times\exp\left\{i\omega\tau-i\mathbf{q}\cdot\left[\mathbf{r}_0\left(t-\frac{|\tau|}{2}+\frac{\tau}{2}\right)-\mathbf{r}_0\left(t-\frac{|\tau|}{2}-\frac{\tau}{2}\right)\right]\right\}.\end{aligned} \tag{6}$$

Applying this equation to the case of synchrotron radiation, we obtain

$$\begin{aligned}\frac{d}{dt}\sum_{\alpha}n^{(0)}_{\alpha\mathbf{q}}=Z^2\frac{v_0^2}{c^2}g_{\mathbf{q}}^2\int_{-\infty}^{\infty}d\tau\Bigg[&\cos\omega_0\tau-\frac{1}{2}\cos^2\Theta\\&\times(\cos\omega_0\tau+\cos((2t-|\tau|)\omega_0))\Bigg]\\&\times\exp\left\{i\omega\tau-2iqR\cos\Theta\sin\frac{\omega_0\tau}{2}\right.\\&\left.\times\cos\left(\omega_0\left(t-\frac{|\tau|}{2}\right)\right)\right\},\end{aligned} \tag{7}$$

where $\omega_0=eH_0/\gamma mc$, $v_0=R\omega_0$, H_0 is the magnetic field strength, R is the orbit's radius, m is the particle mass, and γ is the Lorentz factor. The angle Θ is the inclination of the vector $\mathbf{q}$ relative to the orbital plane.

The expression (7) is periodic in time, with period $T_0=2\pi/\omega_0$. Averaging over one period, we obtain

$$\begin{aligned}\frac{d}{dt}\sum_{\alpha}\overline{n^{(0)}_{\alpha\mathbf{q}}}=Z^2\frac{v_0^2}{c^2}g_q^2\int_{-\infty}^{\infty}d\tau\,e^{i\omega\tau}\Bigg[&\frac{1}{2}\cos^2\Theta J_2\\&\times\left(2qR\cos\Theta\sin\frac{\omega_0\tau}{2}\right)+\cos\omega_0\tau\\&\times\left(1-\frac{1}{2}\cos^2\Theta\right)J_0\left(2qR\cos\Theta\sin\frac{\omega_0\tau}{2}\right)\Bigg].\end{aligned} \tag{8}$$

Next, we allow for the fact that for any periodic function $F(\tau)$,

$$\int_{-\infty}^{\infty} e^{-i\omega\tau}F(\tau)d\tau=\sum_{n=-\infty}^{\infty} e^{in\omega T_0}\int_0^{T_0} e^{i\omega\tau}F(\tau)d\tau, \qquad (9)$$

where the sum of exponentials can be transformed into a sum of delta functions:

$$\sum_{n=-\infty}^{\infty} e^{in\omega T_0}=\frac{2\pi}{T_0}\sum_{n'=-\infty}^{\infty}\delta(\omega-n'\omega_0).$$

Combining this with (9), we can transform (8) to the following form:

$$\frac{d}{dt}\sum_{\alpha}\overline{n_{\alpha\mathbf{q}}^{(0)}}=Z^2\frac{v_0^2}{c^2}g_q^2\sum_{n=0}^{\infty}\frac{1}{\pi}\int_0^{\pi}dx\ e^{2inx}\Bigg\{\cos 2xJ_0 \times(2qR\cos\Theta\sin x)\left(1-\frac{1}{2}\cos^2\Theta\right) +\frac{1}{2}\cos^2\Theta J_2(2qR\cos\Theta\sin x)\Bigg\} \times 2\pi\delta(\omega-n\omega_0). \qquad (10)$$

Since $\omega=cq>0$, the sum in (10) goes from 0 to ∞.

Next we have[7]

$$\int_0^{\pi} e^{2i\mu x}J_{2\nu}(2a\sin x)dx=\pi e^{i\pi\mu}J_{\nu-\mu}(a)J_{\nu+\mu}(a),$$

and the recurrence formulas

$$J_{n+1}(z)+J_{n-1}(z)=\frac{2n}{z}J_n(z),$$

$$J_{n-1}(z)=\frac{n}{z}J_n(z)+J_n'(z),$$

$$J_{n+1}(z)=\frac{n}{z}J_n(z)-J_n'(z).$$

As a result, Eq. (10) becomes

$$\frac{d}{dt}\sum_{\alpha}\overline{n_{\alpha\mathbf{q}}^{(0)}}=Z^2g_q^2\sum_{n=0}^{\infty}2\pi\delta(\omega-n\omega_0) \times\left[\tan^2\Theta J_n\left(\frac{nv}{c}\cos\Theta\right) +\frac{v_0^2}{c^2}J_n'^2\left(\frac{nv}{c}\cos\Theta\right)\right]. \qquad (11)$$

Equation (11) can be used, in particular, to obtain the well-known Schott formula. Thus, for the mean intensity of synchrotron radiation the semiclassical theory yields results that coincide with classical results. The semiclassical theory provides additional information (in comparison to that provided by classical electrodynamics) only in the sense that it makes it possible to calculate the fluctuation in the number of the emitted photons, their mean energy, and total momentum. Similar results can be obtained for the case in which a charged particle moves along an arbitrary path.[8] In all cases we arrive at classical formulas for the mean intensity of the radiation emitted by the particle. Moreover, calculations of the mean electromagnetic field that accompanies a charged particle moving *in vacuo* lead to well-known expressions for the retarded potentials.[9] This agreement between the semiclassical and classical theories forms the basis for a more accurate quantum mechanical theory of interaction of radiation and an emitting particle.

3. QUANTUM MECHANICAL THEORY

We consider the interaction of an electron and the radiation emitted by that electron. We pass to the Furry representation and write the wave operator of the electron in the form of an expansion in the stationary states of type (A2) (see the Appendix):

$$\hat{\psi}=\sum_{\xi}\hat{d}_{\xi}\psi_{\xi},$$

where we have excluded the antiparticle operators, since allowing for the contribution of particle–antiparticle intermediate states leads only to small corrections to the phenomena considered. The creation and annihilation operators, $\hat{d}_{\xi}^{\dagger}$ and $\hat{d}_{\xi}$, must obey the standard Fermi commutation relations. The current density operator is approximately (without allowing for electron–positron pair contributions) given by

$$\hat{j}_a(t)=c\hat{\psi}^{\dagger}\alpha_a\hat{\psi}=c\sum_{\xi,\xi'}\hat{d}_{\xi}^{\dagger}\hat{d}_{\xi'}\psi_{\xi}^{*}(\mathbf{r})\alpha_a\psi_{\xi'}(\mathbf{r})e^{i(E_{\xi}-E_{\xi'})t} \qquad (12)$$

(from now on $a,b=1,2,3$ label the projections of vectors on the Cartesian coordinate axes).

We construct the operator

$$\hat{\mathbf{j}}_{\mathbf{q}}^{(0)}(t)=Z\dot{\mathbf{r}}_0(\mathbf{q},t)e^{-i\mathbf{q}\cdot\mathbf{r}_0(\mathbf{q},t)}\hat{\rho}_{\mathbf{q}}^{(0)}, \qquad (13)$$

where $\mathbf{r}_0(\mathbf{q},t)$ is a vector (which needs to be determined) that depends on the momentum transfer $\mathbf{q}$ and time t, and $\hat{\rho}_{\mathbf{q}}^{(0)}$ is the ''zeroth'' density operator at time $t=0$:

$$\hat{\rho}_{\mathbf{q}}^{(0)}=\sum_{\mathbf{k},\sigma}\hat{d}_{\mathbf{k}\sigma}^{\dagger}\hat{d}_{\mathbf{k}+\mathbf{q},\sigma}.$$

Here $\hat{d}_{\mathbf{k}\sigma}^{\dagger}$ and $\hat{d}_{\mathbf{k}\sigma}$ are the creation and annihilation operators for an electron in a state with momentum $\mathbf{k}$ and a projection of the electron spin on the z axis that takes the values $\sigma=\pm 1/2$. Note that the operator (13) is selected in a form that satisfies the energy conservation law.

We require that the running mean Fourier transform of the operator (12) coincide with the expectation value of (13):

$$(t|\hat{\mathbf{j}}_{\mathbf{q}}(t)|t)=\dot{\mathbf{r}}_0(\mathbf{q},t)e^{-i\mathbf{q}\cdot\mathbf{r}_0(\mathbf{q},t)}(t|\hat{\rho}_{\mathbf{q}}^{(0)}|t). \qquad (14)$$

Then $\mathbf{r}_0(\mathbf{q},t)$ must be approximately equal to the mean position of the particle at time t. We define the deviation of the current from the ''zeroth'' value to be

$$\Delta\hat{\mathbf{j}}_{\mathbf{q}}(t)=\hat{\mathbf{j}}_{\mathbf{q}}(t)-\hat{\mathbf{j}}_{\mathbf{q}}^{(0)}(t).$$

This deviation will be used to build the interaction operator in the new representation. The above transformation is convenient because the operators (13) commute at different times:

$$[\hat{j}^{(0)a}_{\mathbf{q}}(t),\hat{j}^{(0)b}_{\mathbf{q}'}(t')]=0. \tag{15}$$

Using (13), we write the electromagnetic interaction operator as a sum of two terms, $\hat{H}_{\rm int}(t)=\hat{H}^{(0)}_{\rm int}(t)+\hat{H}^{(1)}_{\rm int}(t)$, where

$$\hat{H}^{(0)}_{\rm int}(t)=-\frac{1}{c}\int \hat{\mathbf{j}}^{(0)}(t)\cdot\hat{\mathbf{A}}(t)dV, \tag{16}$$

$$\hat{H}^{(1)}_{\rm int}(t)=-\frac{1}{c}\int \Delta\hat{\mathbf{j}}(t)\cdot\hat{\mathbf{A}}(t)dV. \tag{17}$$

Then, by virtue of (15), the equation

$$i\frac{d}{dt}|t)=\hat{H}^{(0)}_{\rm int}(t)|t) \tag{18}$$

has an exact solution in the form of a direct product of the vectors of extended (or modified, in the terminology of Ref. 10) coherent states,

$$|t)=\prod_{\alpha,\mathbf{q}}\ \exp(-i\hat{\chi}_{\alpha\mathbf{q}}-\hat{f}_{\alpha\mathbf{q}}\hat{Q}^{\dagger}_{\alpha\mathbf{q}}+\hat{f}^{\dagger}_{\alpha\mathbf{q}}\hat{Q}_{\alpha\mathbf{q}})|0), \tag{19}$$

where (at $t_0=0$)

$$\hat{Q}_{\alpha\mathbf{q}}(t)=i\frac{g_q}{c}\int_0^t dt'\mathbf{e}^{*}_{\alpha\mathbf{q}}\cdot\hat{\mathbf{j}}^{(0)}_{\mathbf{q}}(t')e^{i\omega t'}, \tag{20}$$

$$\hat{\chi}_{\alpha\mathbf{q}}(t)=-\frac{i}{2}\int_0^t\{\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t')\hat{Q}_{\alpha\mathbf{q}}(t')-\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t')\hat{Q}_{\alpha\mathbf{q}}(t')\}dt'. \tag{21}$$

The initial state vector $|0)$ is the direct product of the vacuum state of the electromagnetic field, $|\mathrm{vac})$, and the vector of the initial state of the moving particle, $|\phi)$, described by the wave function $\phi(\mathbf{r})$, i.e., $|0)=|\phi,\mathrm{vac})$.

We have chosen t_0 to be zero rather than $-\infty$ due to the fact that, as further calculations show, the temporal sequence of changes in the state of a moving particle that interacts with the field of the radiation it emits is highly important. In this approach there are sure to be problems associated with the interaction turning on, which violates the charge conservation, and with the generation of virtual radiation, which is the consequence of such violation. To avoid the need to discard fictitious terms, one can resort to turning the interaction on slowly by replacing the constant Z with a slowly increasing charge $Z(1-e^{\epsilon t})$, where ϵ is small. The charge buildup time $\tau_{\rm in}=\epsilon^{-1}$ must be long compared to ω^{-1}, but short compared to the observation time (here t must be much longer than ω^{-1}). After we establish a method for evaluating the integrals for some definite value of q, we can extend it to any other value of q.

If we ignore the corrections generated by $\hat{H}^{(1)}_{\rm int}$, Eq. (19) fully solves the problem of calculating the physical quantities of interest. In particular, instead of (5) we have

$$n_{\alpha\mathbf{q}}(t)=(t|\hat{f}^{\dagger}_{\alpha\mathbf{q}}(t)\hat{f}_{\alpha\mathbf{q}}(t)|t),$$

which at $\mathbf{r}_0(\mathbf{q},t)=\mathbf{r}_0(t)$ leads to a result coinciding with (5). Thus, if we ignore $\hat{H}^{(1)}_{\rm int}$, the current variant of the quantum mechanical theory differs from the semiclassical one in calculations of the mean number of the emitted photons or the energy of these photons only when $\mathbf{r}_0(\mathbf{q},t)$ differs from $\mathbf{r}_0(t)$.

At the same time, corrections due to $\hat{H}^{(1)}_{\rm int}$ can be obtained for any convenient choice of the vectors $\mathbf{r}_0(\mathbf{q},t)$. Let us put $\mathbf{r}_0(\mathbf{q},t)=\mathbf{r}_0(t)$. Then $\hat{Q}_{\alpha\mathbf{q}}(t)=Q_{\alpha\mathbf{q}}(t)\hat{\rho}_{\mathbf{q}}$, where $Q_{\alpha\mathbf{q}}(t)$ is specified by its semiclassical expression

$$Q_{\alpha\mathbf{q}}(t)=i\frac{Z}{c}g_q\int_0^t dt'\mathbf{e}^{*}_{\alpha\mathbf{q}}\cdot\mathbf{v}_0(t')\exp\{i\omega t'-i\mathbf{q}\mathbf{r}_0(t')\}. \tag{22}$$

To construct a new ''modified'' perturbation theory in $\hat{H}^{(1)}_{\rm int}$, we introduce the zero-order evolution operator

$$\hat{U}_0(t)=\exp\left\{\sum_{\alpha,\mathbf{q}}\hat{Q}_{\alpha\mathbf{q}}(t)\hat{f}^{\dagger}_{\alpha\mathbf{q}}-\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t)\hat{f}_{\alpha\mathbf{q}}-i\hat{\chi}_{\alpha\mathbf{q}}(t)\right\}.$$

Then (19) can be written as $|t)=\hat{U}_0(t)|0)$. We also introduce a new representation of operators:

$$\tilde{A}(t)=\hat{U}^{\dagger}_0(t)\hat{A}(t)\hat{U}_0(t). \tag{23}$$

The state vector $|t\rangle$ in this representation obeys the equation

$$i\frac{d}{dt}|t\rangle=\tilde{H}^{(1)}_{\rm int}(t)|t\rangle. \tag{24}$$

Allowing for (24), we can reduce the expression for the mean number of photons to

$$n_{\alpha\mathbf{q}}(t)=n^{(0)}_{\alpha\mathbf{q}}(t)+\sum_{n=1}^{\infty}(-i)^n\int_0^t dt_1\int_0^{t_1}dt_2\cdots\int_0^{t_{n-1}}dt_n$$
$$\times\langle 0|[\cdots[[\tilde{f}^{\dagger}_{\alpha\mathbf{q}}(t)\tilde{f}_{\alpha\mathbf{q}}(t),\tilde{H}^{(1)}_{\rm int}(t_1)],$$
$$\tilde{H}^{(1)}_{\rm int}(t_2)],\ldots,\tilde{H}^{(1)}_{\rm int}(t_n)]|0\rangle. \tag{25}$$

4. CALCULATING CORRECTIONS IN THE MODIFIED THEORY

Writing the series in Eq. (25) explicitly, we find that the expansion contains terms proportional to even powers of Z. We collect the leading terms of this type, which contain Z^2 as a pre-exponential factor. Such terms exist in $n^{(0)}_{\alpha\mathbf{q}}(t)$ and in the first and second terms of the series (25). Note that here Z is also contained in the exponents entering into $\hat{U}_0$ and $\hat{U}^{\dagger}_0$. We calculate the first commutator in (25) via the following auxiliary formulas:

$$\tilde{f}_{\alpha\mathbf{q}}(t)=(\hat{f}_{\alpha\mathbf{q}}+\hat{Q}_{\alpha\mathbf{q}}(t))e^{-i\omega t},$$
$$\tilde{f}^{\dagger}_{\alpha\mathbf{q}}(t)=(\hat{f}^{\dagger}_{\alpha\mathbf{q}}+\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t))e^{-i\omega t},$$
$$[\hat{f}_{\alpha\mathbf{q}},\hat{U}_0(t)]=\hat{U}_0(t)\hat{Q}_{\alpha\mathbf{q}}(t),$$
$$[\hat{f}^{\dagger}_{\alpha\mathbf{q}},\hat{U}^{\dagger}_0(t)]=-\hat{U}^{\dagger}_0(t)\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t),$$
$$[\hat{f}^{\dagger}_{\alpha\mathbf{q}},\hat{U}_0(t)]=\hat{U}_0(t)\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t),$$
$$[\hat{f}_{\alpha\mathbf{q}},\hat{U}^{\dagger}_0(t)]=-\hat{U}^{\dagger}_0(t)\hat{Q}_{\alpha\mathbf{q}}(t). \tag{26}$$

We put

$$\hat{B}_{\alpha\mathbf{q}}(t)=\mathbf{e}_{\alpha\mathbf{q}}\cdot\Delta\hat{\mathbf{j}}_{-\mathbf{q}}(t),$$

so that

$$\hat{H}^{(1)}_{\rm int}(t)=-\frac{Z}{c}\sum_{\alpha,\mathbf{q}} g_q(\hat{f}_{\alpha\mathbf{q}}\hat{B}_{\alpha\mathbf{q}}(t)e^{-i\omega t}+\hat{f}^{\dagger}_{\alpha\mathbf{q}}\hat{B}^{\dagger}_{\alpha\mathbf{q}}(t)e^{i\omega t}). \tag{27}$$

Using (26), we can perform the following transformation:

$$\begin{aligned}&[\tilde{f}^{\,\dagger}_{\alpha\mathbf{q}}(t)\tilde{f}_{\alpha\mathbf{q}}(t),\tilde{H}^{(1)}_{\rm int}(t)]\\&\quad=\hat{U}^{\dagger}_0(t_1)[(\hat{f}^{\dagger}_{\alpha\mathbf{q}}+\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t,t_1))(\hat{f}_{\alpha\mathbf{q}}+\hat{Q}_{\alpha\mathbf{q}}(t,t_1)),\\&\quad\hat{H}^{(1)}_{\rm int}(t_1)]\hat{U}_0(t_1),\end{aligned} \tag{28}$$

where $\hat{Q}_{\alpha\mathbf{q}}(t,t_1)=\hat{Q}_{\alpha\mathbf{q}}(t)-\hat{Q}_{\alpha\mathbf{q}}(t_1)$. Since the operators $\hat{Q}_{\alpha\mathbf{q}}(t)$ already contain Z as a factor, the leading terms emerge as a result of the commutation of the photon operators and $\hat{H}^{(1)}_{\rm int}$:

$$\begin{aligned}&[\tilde{f}^{\,\dagger}_{\alpha\mathbf{q}}(t)\tilde{f}_{\alpha\mathbf{q}}(t),\tilde{H}^{(1)}_{\rm int}(t_1)]\\&\quad\approx\frac{Z}{c}g_q\hat{U}^{\dagger}_0(t_1)(e^{-i\omega t_1}\hat{B}_{\alpha\mathbf{q}}(t_1)(\hat{f}_{\alpha\mathbf{q}}+\hat{Q}_{\alpha\mathbf{q}}(t,t_1))\\&\quad-e^{i\omega t_1}(\hat{f}^{\dagger}_{\alpha\mathbf{q}}+\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t,t_1))\hat{B}^{\dagger}_{\alpha\mathbf{q}}(t_1))\hat{U}_0(t_1).\end{aligned} \tag{29}$$

If we average (29) over the initial state of the system by employing the equalities

$$\hat{f}_{\alpha\mathbf{q}}\hat{U}_0(t_1)|0)=\hat{Q}_{\alpha\mathbf{q}}(t_1)\hat{U}_0(t_1)|0),$$

$$(0|\hat{U}^{\dagger}_0(t_1)\hat{f}^{\dagger}_{\alpha\mathbf{q}}=(0|\hat{U}^{\dagger}_0(t_1)\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t_1),$$

we get

$$\begin{aligned}(0|[\tilde{f}^{\,\dagger}_{\alpha\mathbf{q}}(t)\tilde{f}_{\alpha\mathbf{q}}(t),\tilde{H}^{(1)}_{\rm int}(t_1)]|0)&=\frac{Z}{c}g_q(0|\tilde{U}^{\dagger}_0(t_1)\\&\times\{\hat{B}_{\alpha\mathbf{q}}(t_1)\hat{Q}_{\alpha\mathbf{q}}(t)e^{-i\omega t_1}\\&-e^{i\omega t_1}\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t)\hat{B}^{\dagger}_{\alpha\mathbf{q}}(t_1)\}\hat{U}_0(t_1)|0).\end{aligned} \tag{30}$$

In calculating the next correction in (25) we immediately discard terms that contain pre-exponential factors with Z raised to a power greater than two. This means that when we plug such terms into the second and subsequent terms of the sum in (25) into the expression for the first-order commutator, we can immediately discard terms containing the operators $\hat{Q}_{\alpha\mathbf{q}}$ and $\hat{Q}^{\dagger}_{\alpha\mathbf{q}}$. In the resulting expressions, the operators $\hat{f}_{\alpha\mathbf{q}}$ and $\hat{f}^{\dagger}_{\alpha\mathbf{q}}$ can be freely interchanged with the operators $\hat{U}_0$ and $\hat{U}^{\dagger}_0$, since their commutators contain higher-order corrections in Z, which we have just discarded.

In view of this, all terms in which the annihilation operators $\hat{f}_{\alpha\mathbf{q}}$ are to the right of other $\hat{f}$-operators, or in which the creation operators $\hat{f}^{\dagger}_{\alpha\mathbf{q}}$ are to the left of other $\hat{f}$-operators, must be dropped. In the remaining terms the operator products $\hat{f}_{\alpha\mathbf{q}}\hat{f}^{\dagger}_{\alpha/\mathbf{q}'}$ must be replaced by the commutators $\delta_{\alpha\alpha'}\Delta(\mathbf{q}-\mathbf{q}')$. By performing these transformations we reduce the leading terms that appear when we write the double commutator on the right-hand side of Eq. (25) explicitly to the form

$$\begin{aligned}&-\frac{Z^2}{c^2}g^2_q(\hat{U}^{\dagger}_0(t_1)e^{-i\omega(t_1-t_2)}\hat{B}_{\alpha\mathbf{q}}(t_1)\hat{U}_0(t_1)\\&\times\hat{U}^{\dagger}_0(t_2)\hat{B}^{\dagger}_{\alpha\mathbf{q}}(t_2)\hat{U}_0(t_2)+\hat{U}^{\dagger}_0(t_2)\hat{B}_{\alpha\mathbf{q}}(t_2)\hat{U}_0(t_2)\hat{U}^{\dagger}_0(t_1)\\&\times\hat{B}^{\dagger}_{\alpha\mathbf{q}}(t_1)\hat{U}_0(t_1)e^{i\omega(t_1-t_2)}).\end{aligned} \tag{31}$$

Collecting all terms of the specified order, we get

$$\begin{aligned}n_{\alpha\mathbf{q}}(t)&=|Q_{\alpha\mathbf{q}}(t)|^2-i\frac{Z}{c}g_q\int_0^t dt_1(0|\hat{U}^{\dagger}_0(t_1)\\&\times\{e^{-i\omega t_1}\hat{B}_{\alpha\mathbf{q}}(t_1)\hat{Q}_{\alpha\mathbf{q}}(t)\\&-e^{i\omega t_1}\hat{Q}^{\dagger}_{\alpha\mathbf{q}}(t)\hat{B}^{\dagger}_{\alpha\mathbf{q}}(t_1)\}\hat{U}_0(t_1)|0)\\&+\frac{Z^2}{c^2}g^2_q\int_0^t dt_1\int_0^{t_1}dt_2(0|\hat{U}^{\dagger}_0(t_1)\\&\times\{e^{-i\omega(t_1-t_2)}\hat{B}_{\alpha\mathbf{q}}(t_1)\hat{U}_0(t_1)\hat{U}^{\dagger}_0(t_2)\hat{B}^{\dagger}_{\alpha\mathbf{q}}(t_2)\\&\times\hat{U}_0(t_2)+\hat{U}^{\dagger}_0(t_2)\hat{B}^{\dagger}_{\alpha\mathbf{q}}(t_2)\hat{U}_0(t_2)\hat{U}^{\dagger}_0(t_1)\\&\times e^{i\omega(t_1-t_2)}\hat{B}_{\alpha\mathbf{q}}(t_1)\}\hat{U}_0(t_1)|0).\end{aligned} \tag{32}$$

If we now write all the terms in (32) that appear because of plugging the explicit expressions for $\hat{B}_{\alpha\mathbf{q}}(t)$ into (32), collect like terms, and do the necessary canceling, we arrive at the final result:

$$\begin{aligned}n_{\alpha\mathbf{q}}(t)&=\frac{Z^2}{c^2}g^2_q\int_0^t dt_1\int_0^t dt_2(0|\hat{U}^{\dagger}_0(t_1)\\&\times(\mathbf{e}_{\alpha\mathbf{q}}\cdot\hat{\mathbf{j}}^{\dagger}_{\mathbf{q}}(t_1))\hat{U}_0(t_1)\hat{U}^{\dagger}_0(t_2)\\&\times(\mathbf{e}^{*}_{\alpha\mathbf{q}}\cdot\hat{\mathbf{j}}_{\mathbf{q}}(t_2))\hat{U}_0(t_2)|0)e^{-i\omega(t_1-t_2)}.\end{aligned} \tag{33}$$

Note that in deriving (33), we did not take advantage of the fact that $\mathbf{r}_0(\mathbf{q},t)$ is independent of $\mathbf{q}$, with the result that the formula still holds in the general case, in which $\mathbf{r}_0(t)$ is replaced by $\mathbf{r}_0(\mathbf{q},t)$ in (22).

5. NUMBER OF PHOTONS

We assume that in the expansion of the initial state vector of the particle, $|0)$, the expansion coefficients $C_{\mathbf{k}_i}$ in states with definite momentum $|\mathbf{k}_i)$ have a peak at $\mathbf{k}_0$, and decrease as $\mathbf{k}_i$ deviates from $\mathbf{k}_0$, by the Gauss law

$$C_{\mathbf{k}_i}=(2\pi\delta^2_{\perp})^{1/2}(2\pi\delta^2_l)^{1/4}\exp\left[-\frac{p^2_i\delta^2_l}{4}-\frac{(\mathbf{k}_{i\perp}-\mathbf{k}_{0\perp})^2\delta^2_{\perp}}{4}\right],$$

where $\mathbf{k}=(\mathbf{k}_{\perp},p_i)$, $\mathbf{k}_{0\perp}$ is time-dependent, and $\delta_{\perp}$ and δ_l are the transverse and longitudinal packet widths (relative to the z axis). This representation follows from the study of electron states in a magnetic field in the Appendix. For relativistic electrons, the momentum uncertainty in the initial state is much less than the momentum proper. In real calculations of the numbers of emitted photons via (33), it is preferable to represent the current operators as expansions in states with

definite momentum at a given moment in time, with a time dependence characteristic of plane waves. In the present paper, this approximation is justified by the fact that due to the strong effect of the radiation on the particle's state in the comoving reference frame, an effect exceeding the one produced by the external field, we can ignore the quantization of levels in the time dependence of the operators. Indeed, even the classical theory of synchrotron radiation predicts that the mean energy of the photons emitted by a particle is much greater than ω_0. In view of this, the mean difference in particle energies before and after photon emission proves to be much greater than the separation between the levels of transverse motion. Under these conditions, allowance for level quantization in the time dependence of the operator can only lead to small corrections of order $1/\bar{n}$ (where $\bar{n}\sim\gamma^3$ is the mean ratio of the frequency of the emitted photon to ω_0).

As a result of the action of electron operators, the vectors $\mathbf{k}$ and $\mathbf{k}_1$ in the current operators in (33) are transformed into the vectors $\mathbf{k}_i-\Delta\mathbf{q}$, where $\Delta\mathbf{q}=\Sigma_s\mathbf{q}_s$, with $\mathbf{q}_1,\mathbf{q}_2,\ldots$ the momenta of the emitted photons. Replacing the given expression with $\mathbf{k}_i(t)=\mathbf{k}_i-\Delta\mathbf{k}(t)$, where $\Delta\mathbf{k}(t)$ is the mean momentum lost by the particle by time t, and plugging it into all the cofactors in (33) that are not in the exponential, we get

$$
\begin{aligned}
n_{\alpha\mathbf{q}}(t)=&\frac{Z^2}{c^2}g_q^2\int_0^t dt_1\int_0^t dt_2 e^{-i\omega(t_1-t_2)}\\
&\times\sum_{\mathbf{k}_i,\sigma'}|C_{\mathbf{k}_i}|^2(\mathbf{e}_{\alpha\mathbf{q}}\cdot\mathbf{v}^*_{i\sigma'}(\mathbf{q},t_1))\\
&\times(\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_{i\sigma'}(\mathbf{q},t_2))(\mathrm{vac};\mathbf{k}_i,\sigma|\hat{U}_0^\dagger(t_1)\hat{\rho}^\dagger_{\mathbf{q}}(t_1)\\
&\times\hat{U}_0(t_1)\hat{U}_0^\dagger(t_2)\hat{\rho}_{\mathbf{q}}(t_2)\hat{U}_0(t_2)|\mathbf{k}_i,\sigma;\mathrm{vac}),
\end{aligned}
\tag{34}
$$

where

$$
\hat{\rho}_{\mathbf{q}}(t)=\sum_{\mathbf{k},\sigma,\sigma'}\hat{d}^\dagger_{\mathbf{k}-\mathbf{q},\sigma'}\hat{d}_{\mathbf{k},\sigma}\exp\{i(\varepsilon_{\mathbf{k}-\mathbf{q}}-\varepsilon_{\mathbf{k}})t\},
$$

$$
\begin{aligned}
v^a_{i\sigma'}(\mathbf{q},t)=&\frac{c^2}{2\sqrt{\varepsilon_i\varepsilon_i'}}w^*_{\sigma'}\Bigg[\sqrt{\frac{\varepsilon_i'+mc^2}{\varepsilon_i+mc^2}}\,\sigma^a\sigma^b k_i^b(t)\\
&+\sqrt{\frac{\varepsilon_i+mc^2}{\varepsilon_i'+mc^2}}\,\sigma^b(k_i^b(t)-q^b)\sigma^a\Bigg]w_\sigma,
\end{aligned}
$$

with $\varepsilon_i=\varepsilon_{\mathbf{k}_i(t)}$, $\varepsilon_i'=\varepsilon_{\mathbf{k}_i(t)-\mathbf{q}}$, and summation over repeated indices is implied. The term corresponding to $\sigma'\neq\sigma$ describes emission processes accompanied by electron spin flip.

Further simplification is possible if the exponents in the density operators in (34) are transformed according to

$$
\varepsilon_{\mathbf{k}-\mathbf{q}}-\varepsilon_{\mathbf{k}}\approx\sum_s\mu(\mathbf{q},\mathbf{q}_s').
\tag{35}
$$

where $\mu(\mathbf{q},\mathbf{q}_s')$ are unspecified functions. In this approach, different photons are assumed to be almost independent, since otherwise we would have to speak of a strong correlation between the emission of two separate photons, which agrees neither with the semiclassical theory nor with the calculations below. In an approximation that is linear in $\Delta\mathbf{q}$, for $q\ll k_i$ we have

$$
\mu_i(\mathbf{q},\mathbf{q}_s')\approx(\nabla\varepsilon_{\mathbf{k}_i-\mathbf{q}}-\nabla\varepsilon_{\mathbf{k}_i})\mathbf{q}_s'\approx-\frac{\mathbf{q}\cdot\mathbf{q}_s'}{m\gamma_i},
\tag{36}
$$

where $\gamma_i=\varepsilon_{\mathbf{k}_i}/mc^2$. As $q_s'\to\infty$, the function $\mu(\mathbf{q},\mathbf{q}_s')$ ceases to depend on $\mathbf{q}_s'$.

Using the methods of calculating means employed in Ref. 9, we get

$$
\begin{aligned}
n_{\alpha\mathbf{q}}(t)=&\int_0^t dt_1\int_0^t dt_2\sum_{\mathbf{k}_i,\sigma'}|C_{\mathbf{k}_i}|^2\dot{Q}^*_{i\alpha\mathbf{q}}(t_1,\sigma')\dot{Q}_{i\alpha\mathbf{q}}(t_2,\sigma')\\
&\times\exp[-P_{i\mathbf{q}}(t_1,t_2)],
\end{aligned}
\tag{37}
$$

where

$$
\begin{aligned}
Q_{i\alpha\mathbf{q}}(t,\sigma')=&i\frac{Z}{c}g_q\int_0^t dt'\,\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_{i\sigma'}(\mathbf{q},t')\\
&\times\exp[i\omega t'-i\mathbf{q}\cdot\mathbf{r}_{i\sigma'}(\mathbf{q},t')],
\end{aligned}
$$

with $\dot{\mathbf{r}}_{i\sigma'}(\mathbf{q},t)=\mathbf{v}_{i\sigma'}(\mathbf{q},t)$. The exponent in (37) is given by

$$
\begin{aligned}
P_{i\mathbf{q}}(t_1,t_2)=&\sum_{\beta,\mathbf{q}',\sigma'}[|Q_{i\beta\mathbf{q}'}(t_1,\sigma')|^2\\
&\times(1-\exp[-i\mu_i(\mathbf{q},\mathbf{q}')t_1])\\
&+|Q_{i\beta\mathbf{q}'}(t_2,\sigma')|^2(1-\exp[i\mu_i(\mathbf{q},\mathbf{q}')t_2])\\
&-Q^*_{i\beta\mathbf{q}'}(t_1,\sigma')Q_{i\beta\mathbf{q}'}(t_2,\sigma')\\
&\times(1-\exp[-i\mu_i(\mathbf{q},\mathbf{q}')t_1])\\
&\times(1-\exp[i\mu_i(\mathbf{q},\mathbf{q}')t_2])].
\end{aligned}
\tag{38}
$$

Obviously,

$$
P^*_{i\mathbf{q}}(t_1,t_2)=P_{i\mathbf{q}}(t_2,t_1),
$$

$$
\lim_{q\to0}P_{i\mathbf{q}}(t_1,t_2)\to0,
$$

$$
\lim_{t_1\to t_2}P_{i\mathbf{q}}(t_1,t_2)\to0.
$$

Equation (37) contains the desired corrections to the semiclassical expression for the number of emitted photons. It assumes its semiclassical form for $|P_{i\mathbf{q}}(t_1,t_2)|\ll1$. From a physical standpoint, this difference between the formulas is due to the fact that in (37) we allow for interaction of the emitted photons, while in the semiclassical theory this factor is ignored. The probability distribution for the number of emitted photons in each state does not obey the Poisson law any longer, which a reflection of the nonlinearity of electromagnetic phenomena in the quantum theory.

Obviously, an equation like (37) can be used to study arbitrary motion of a particle, not just an electron in a synchrotron. To do so, we merely redefine the quantities $\mathbf{v}_{i\sigma'}(\mathbf{q},t)$, which in the simplest case can be approximately calculated for the mean of the vector $\mathbf{k}_i$ and averaged over spin (in this case, the velocities $\mathbf{v}(\mathbf{q},t)$ and the function (38) no longer depend on the indices i and σ').

Let us estimate $P_{\mathbf{q}}$ for the case in which the velocity $\mathbf{v}(\mathbf{q},t)$ is constant and equal to $\mathbf{v}_0$:

$$Q_{\alpha\mathbf{q}}(t)=\frac{Z}{c}g_{\mathbf{q}}\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_0}{\omega-\mathbf{q}\cdot\mathbf{v}_0}e^{i(\omega-\mathbf{q}\cdot\mathbf{v}_0)t}.$$

Plugging this into (38), we obtain an expression that is logarithmically divergent, due to the slow decrease in the integrands as $q'\to\infty$. This fact is the manifestation of ultraviolet divergence, often encountered in electrodynamics. In contrast to Feynman's perturbation theory, ultraviolet divergence does not lead to a catastrophe: it only means that (in contrast to the predictions of the semiclassical theory) a uniformly moving particle is not accompanied by transverse photons. This example is a clear demonstration of the dependence of the way in which the ultraviolet divergence depends on the perturbation theory employed. A detailed study of this problem lies outside the scope of the present paper, where we use the standard method of introducing a cutoff momentum $q_c\sim mc$ to remove the ultraviolet singularity. The resulting expression for the absolute value of the function (38) proves to be small and varies very slowly (logarithmically) with t_1 and t_2. An explicit estimate of the function (38) for $\mathbf{v}(\mathbf{q},t)$ constant will be made in Sec. 6.

6. INFRARED ASYMPTOTIC BEHAVIOR OF THE NUMBER OF PHOTONS

Let us discuss the asymptotic behavior of the function (37) as $\omega=qc\to 0$. In classical electrodynamics (see, e.g., Ref. 11) and in the semiclassical theory there is a characteristic frequency dependence of $n_{\alpha\mathbf{q}}$ as $\omega\to 0$, namely, $n_{\alpha\mathbf{q}}\propto 1/\omega^3$. Hence, upon integration with respect to momenta, the total number of emitted photons diverges logarithmically at the lower limit. Will allowing for the effect of emission on the state of the emitting particle (as in Eq. (37)) influence this pattern? To answer this question, we examine a model problem in which a charged particle moves at constant velocity $\mathbf{v}_1$ and, colliding at time $t_3>0$ with a point scatterer, suddenly changes its own velocity by a small quantity $\Delta\mathbf{v}=\mathbf{v}_2-\mathbf{v}_1$, $|\Delta\mathbf{v}|\ll v_1$, and then proceeds to move at constant velocity $\mathbf{v}_2$. The requirement that this jump in velocity be small simplifies all calculations considerably. Moreover, since a jump in velocity implies infinite acceleration, various nonphysical effects are to be expected. The requirement that the velocity jump be small makes the velocity almost a continuous function, so that such effects can be ignored. When Eq. (37) is employed in calculations, there is the problem of the interaction suddenly turning on at the initial moment in time, which violates energy conservation, and of generation of fictitious radiation, which is the consequence of such violation. To avoid the need to discard fictitious terms, one can use the procedure developed in Sec. 3 to turn the interaction on slowly.

Since the particle is assumed to have a definite velocity, we drop the subscript i in (38) and replace the vectors $\mathbf{v}_{i\sigma}$ with the current value of the velocity. We calculate the resulting functions $P_{\mathbf{q}}(t_1,t_2)$, assuming that $\tau_{\text{in}}\ll t_3\ll t$. To this end, we first estimate the quantities $f_{t_1,t_2}=\exp[\pm i\mathbf{q}\cdot\mathbf{q}'t_{1,2}/m\gamma]$ on the right-hand side of Eq. (38). Since a photon is emitted when the particle changes velocity, we consider the neighborhood of the point $t_1=t_3$, $t_2=t_3$, assuming that $t_3\sim\gamma_0 m/q^2$. The vector $\mathbf{q}'$ is the momentum transferred from the moving particle to the emitted quanta of electromagnetic field (photons). The mean value of this momentum is of order $m|\Delta\mathbf{v}|$, so that at $q\ll m|\Delta\mathbf{v}|$ the ratio q'/q can be large. Thus, the absolute value of the exponent in $f(t_1,t_2)$ in the range of parameters under investigation is large, and the exponentials are rapidly varying functions that make a negligible contribution to (38). Eliminating these contributions from the outset, we reduce (38) to the simpler form

$$P_{\mathbf{q}}(t_1,t_2)=\sum_{\beta,\mathbf{q}'}[|Q_{\beta\mathbf{q}'}(t_1)|^2+|Q_{\beta\mathbf{q}'}(t_2)|^2 - Q^*_{\beta\mathbf{q}'}(t_1)Q_{\beta\mathbf{q}'}(t_2)(1+e^{-i\mathbf{q}\cdot\mathbf{q}'(t_1-t_2)/m\gamma})]. \tag{39}$$

We now calculate the function (39) explicitly for $\mathbf{v}(\mathbf{q},t)=\mathbf{v}_0=\text{const}$. In this case, assuming that q_c is much less than the mean momentum of the emitting particle, we calculate the integral with respect to q' and obtain

$$P_{\mathbf{q}}(t_1,t_2)=\frac{Z^2}{4\pi^2c^3}\int do'\,\frac{(\mathbf{n}'\times\mathbf{v}_0)^2}{(1-\mathbf{n}'\cdot\mathbf{v}_0/c)^2} \times\{i\,\text{Si}(\omega_2(t_1-t_2))+i\,\text{Si}((\omega_2+\omega_1)(t_1-t_2)) +2\widetilde{C}-\text{Ci}(\omega_2|t_1-t_2|)-\text{Ci}(|\omega_2+\omega_1||t_1-t_2|) +\ln(\omega_2|\omega_2+\omega_1|(t_1-t_2)^2)\}, \tag{40}$$

where $\mathbf{n}'=\mathbf{q}'/q'$, $\omega_1=q_c\mathbf{n}'\cdot\mathbf{q}/m\gamma$, $\omega_2=(c-\mathbf{n}'\cdot\mathbf{v}_0)q_c$, $\text{Si}(\xi)$ and $\text{Ci}(\xi)$ are the sine and cosine integrals, and $\widetilde{C}=0.5772\ldots$ is Euler's constant. The function (40) vanishes at $t_1=t_2$ and slowly increases with the time difference $\Delta t=|t_1-t_2|$. In the nonrelativistic limit at large $\Delta t\gg 1/cq_c$, the function (40) can be approximated by the expression

$$P_{\mathbf{q}}(t_1,t_2)\approx\frac{2Z^2v_0^2}{3\pi c^3}[i\pi\,\text{sgn}(t_1-t_2) +2(\widetilde{C}+\ln(cq_c)+\ln|t_1-t_2|)]. \tag{41}$$

We remark on the smallness of the coefficient of the expression in square brackets. As Δt increases, the real part of (41) increases logarithmically, but the characteristic buildup time proves to be exponentially large, so that the function (41) can be considered small over the entire range of its arguments.

Now let us estimate the number of photons emitted by the electron in the entire course of its motion for the nonrelativistic case. Integrating by parts, we find, for instance, that

$$n_{\alpha\mathbf{q}}(t)=-i\frac{Z^2}{c^2}g_q^2\int_0^t dt_1\mathbf{e}_{\alpha\mathbf{q}}\cdot\mathbf{v}(t_1)e^{-i\omega t_1+i\mathbf{q}\cdot\mathbf{r}(t_1)} \times\left[\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}(t_2)}{\omega-\mathbf{q}\cdot\mathbf{v}(t_2)+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2}\right.$$

$$\times\exp\{i\omega t_2-i\mathbf{q}\cdot\mathbf{r}(t_2)-P_{\mathbf{q}}(t_1,t_2)\}\Big|_{t_2=0}^{t_2=t}$$

$$-\int_0^t dt_2\,\exp\{i\omega t_2-i\mathbf{q}\cdot\mathbf{r}(t_2)-P_{\mathbf{q}}(t_1,t_2)\}$$

$$\times\frac{\partial}{\partial t_2}\left(\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}(t_2)}{\omega-\mathbf{q}\cdot\mathbf{v}(t_2)+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2}\right)\Bigg].$$

Allowance for the value of the first term inside the square brackets at the lower limit is unjustified because of the violation of charge conservation at $t=0$. If we turn the interaction on slowly, then this contribution is zero. First we integrate by parts with respect to t_1, using the same ideas that we used in integrating with respect to t_2. We obtain

$$n_{\alpha\mathbf{q}}(t)\approx\frac{Z^2}{c^2}g_q^2\left|\frac{\mathbf{e}_{\alpha\mathbf{q}}\cdot\mathbf{v}_2}{\omega-\mathbf{q}\cdot\mathbf{v}_2+i\partial P_{\mathbf{q}}(t,t_2)/\partial t_2}\right|_{t_2=t}^2$$

$$+\frac{Z^2}{c^2}g_q^2\int_0^t dt_2\int_0^t dt_1\,\exp\{i\omega(t_2-t_1)$$

$$+i\mathbf{q}\cdot(\mathbf{r}(t_1)-\mathbf{r}(t_2))-P_{\mathbf{q}}(t_1,t_2)\}\,\frac{\partial}{\partial t_1}$$

$$\times\Bigg[\frac{\mathbf{e}_{\alpha\mathbf{q}}\cdot\mathbf{v}(t_1)}{\omega-\mathbf{q}\cdot\mathbf{v}(t_1)-i\partial P_{\mathbf{q}}/\partial t_1}$$

$$\times\frac{\partial}{\partial t_2}\left(\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}(t_2)}{\omega-\mathbf{q}\cdot\mathbf{v}(t_2)+i\partial P_{\mathbf{q}}/\partial t_2}\right)\Bigg],\qquad(42)$$

where we have discarded the rapidly oscillating terms, which contribute nothing to the overall expression for the number of emitted photons. The first term on the right-hand side of Eq. (42) corresponds to the part of the transverse field that follows the moving particle, and is related neither to changes in the particle's velocity nor to the radiation. Hence in all calculations of the characteristics of the radiation that follow, we allow only for the second (integral) term.

In calculating the time derivatives in (42) we encounter continuous and delta-functions terms, with the latter being a reflection of the discontinuity in velocity, the derivatives $\dot{Q}_{\alpha\mathbf{q}}$ and $\dot{P}_{\mathbf{q}}$. For instance,

$$\frac{\partial}{\partial t_2}\left(\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}(t_2)}{\omega-\mathbf{q}\cdot\mathbf{v}(t_2)+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2}\right)$$

$$=i\theta(t_3-t_2)\,\frac{\partial^2P_{\mathbf{q}}(t_1,t_2)}{\partial t_2^2}\,\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_1}{(\omega-\mathbf{q}\cdot\mathbf{v}_1+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2)^2}$$

$$+i\theta(t_2-t_3)\,\frac{\partial^2P_{\mathbf{q}}(t_1,t_2)}{\partial t_2^2}$$

$$\times\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_2}{(\omega-\mathbf{q}\cdot\mathbf{v}_2+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2)^2}$$

$$+\delta(t_2-t_3)\left[\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_2}{\omega-\mathbf{q}\cdot\mathbf{v}_2+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2}\right|_{t_2=t_3+0}$$

$$-\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_1}{\omega-\mathbf{q}\cdot\mathbf{v}_1+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2}\Bigg|_{t_2=t_3-0}\Bigg].\qquad(43)$$

Here $\theta(\xi)$ is the Heaviside step function.

The relationship between the continuous and delta-function terms in (43) can be evaluated as follows. The total contribution of the θ-functions can again be calculated by parts, which again results in a delta-function contribution multiplied by the magnitude of the discontinuity of the integrand at $t_2=t_3$. This jump includes the second derivative of $P_{\mathbf{q}}$ as a factor whose order of magnitude can be estimated to be the product of the first derivative and the mean value of the frequency of the emitted photon. The latter cannot exceed the energy lost by the moving particle, and it is therefore proportional to the small parameter $\lambda=\mathbf{v}_1\cdot\Delta\mathbf{v}/v_1^2$. Clearly, allowing for the continuous terms in (43) would mean allowing for the next terms in the series expansion of the integrals in λ. The leading term is still the contribution of the delta function, the only contribution we consider.

Using the condition that the interaction is turned on slowly, we find that

$$Q_{\alpha\mathbf{q}}(t)=\frac{Z}{c}g_q\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_1}{\omega-\mathbf{q}\cdot\mathbf{v}_1}e^{i(\omega-\mathbf{q}\cdot\mathbf{v}_1)t},\quad \tau_{\text{in}}\ll t\leqslant t_3.$$

For $t_3<t$ the result is different:

$$Q_{\alpha\mathbf{q}}(t)=\frac{Z}{c}g_q\Bigg[\left(\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_1}{\omega-\mathbf{q}\cdot\mathbf{v}_1}-\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_2}{\omega-\mathbf{q}\cdot\mathbf{v}_2}\right)e^{i(\omega-\mathbf{q}\cdot\mathbf{v}_1)t_3}$$

$$+\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_2}{\omega-\mathbf{q}\cdot\mathbf{v}_2}e^{i(\omega-\mathbf{q}\cdot\mathbf{v}_2)t}\Bigg].$$

For $\tau_{\text{in}}\ll t_2\leqslant t_3$ we have

$$\frac{\partial P_{\mathbf{q}}(t_1,t_2)}{\partial t_2}=-i\sum_{\beta,\mathbf{q}'}Q^*_{\beta\mathbf{q}'}(t_1)$$

$$\times\Bigg[\frac{Z}{c}g_{q'}(\mathbf{e}^*_{\beta\mathbf{q}'}\cdot\mathbf{v}_1)e^{i(\omega-\mathbf{q}\cdot\mathbf{v}_1)t_2}$$

$$\times\left(1+\exp\left\{-\frac{i\mathbf{q}\cdot\mathbf{q}'(t_1-t_2)}{m\gamma}\right\}\right)$$

$$+\frac{\mathbf{q}\cdot\mathbf{q}'}{m\gamma}Q_{\beta\mathbf{q}'}(t_2)\exp\left\{-\frac{i\mathbf{q}\cdot\mathbf{q}'(t_1-t_2)}{m\gamma}\right\}\Bigg].$$

Finally, for $t_2>t_3$ we have

$$\frac{\partial P_{\mathbf{q}}(t_1,t_2)}{\partial t_2}=\sum_{\beta,\mathbf{q}'}\Bigg[-2\,\frac{Z^2}{c^2}\,g_{q'}^2\,\sin\left\{(\omega'-\mathbf{q}'\cdot\mathbf{v}_2)\,\frac{t_2-t_3}{2}\right\}$$

$$\times(\mathbf{e}^*_{\beta\mathbf{q}'}\cdot\mathbf{v}_2)\left(\frac{\mathbf{e}_{\beta\mathbf{q}'}\cdot\mathbf{v}_1}{\omega'-\mathbf{q}'\cdot\mathbf{v}_1}-\frac{\mathbf{e}_{\beta\mathbf{q}'}\cdot\mathbf{v}_2}{\omega'-\mathbf{q}'\cdot\mathbf{v}_2}\right)$$

$$-i\,\frac{\mathbf{q}\cdot\mathbf{q}'}{m\gamma}Q^*_{\beta\mathbf{q}'}(t_1)Q_{\beta\mathbf{q}'}(t_2)$$

$$\times\exp\left\{-\frac{i\mathbf{q}\cdot\mathbf{q}'(t_1-t_2)}{m\gamma}\right\}-iQ^*_{\beta\mathbf{q}'}(t_1)\,\frac{Z}{c}g_{q'}$$

$$\times(\mathbf{e}^*_{\beta\mathbf{q}'}\cdot\mathbf{v}_2)e^{i(\omega'-\mathbf{q}'\cdot\mathbf{v}_2)t_2}$$

$$\times\left(1+\exp\left\{-\frac{i\mathbf{q}\cdot\mathbf{q}'(t_1-t_2)}{m\gamma}\right\}\right)\Bigg].$$

Note that $\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2$ is continuous in t_1.

Let us calculate the delta-function contribution to the integrals with respect to t_2 in (42), letting $t\to\infty$:

$$n_{\alpha\mathbf{q}}(\infty)=\frac{Z^2}{c^2}g_q^2\int_0^\infty dt_1\ \exp[-i\omega(t_1-t_3)+i\mathbf{q}\cdot(\mathbf{r}_0(t_1)-\mathbf{r}_0(t_3))-P_{\mathbf{q}}(t_1,t_3)]$$
$$\times\frac{\partial}{\partial t_1}\Bigg[\frac{\mathbf{e}_{\alpha\mathbf{q}}\cdot\mathbf{v}(t_1)}{\omega-\mathbf{q}\cdot\mathbf{v}(t_1)-i\partial P_{\mathbf{q}}(t_1,t_3)/\partial t_1}$$
$$\times\Bigg(\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_2}{\omega-\mathbf{q}\cdot\mathbf{v}_2+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2}\Bigg|_{t_2=t_3+0}$$
$$-\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_1}{\omega-\mathbf{q}\cdot\mathbf{v}_1+i\partial P_{\mathbf{q}}(t_1,t_2)/\partial t_2}\Bigg|_{t_2=t_3-0}\Bigg].$$

Now we integrate with respect to t_1, again limiting ourselves to delta-function contributions. Allowing for the fact that $P_{\mathbf{q}}(t_2,t_1)=P^*_{\mathbf{q}}(t_1,t_2)$, we obtain

$$n_{\alpha\mathbf{q}}(\infty)=\frac{Z^2}{c^2}g_q^2\Bigg|\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_2}{\omega-\mathbf{q}\cdot\mathbf{v}_2+i\partial P_{\mathbf{q}}(t_3,t_2)/\partial t_2}\Bigg|_{t_2=t_3+0}$$
$$-\frac{\mathbf{e}^*_{\alpha\mathbf{q}}\cdot\mathbf{v}_1}{\omega-\mathbf{q}\cdot\mathbf{v}_1+i\partial P_{\mathbf{q}}(t_3,t_2)/\partial t_2}\Bigg|_{t_2=t_3-0}\Bigg|^2. \quad (44)$$

This equation solves the problem. If we neglect the derivatives of $P_{\mathbf{q}}$ in the denominators, (44) coincides with the standard expression for the number of low-frequency photons emitted in a collision, an expression that can derived in classical electrodynamics[11] and in quantum electrodynamics if we use standard perturbation theory.[6] Let us estimate the derivatives of $P_{\mathbf{q}}$ in the denominators in (44). We have

$$\frac{\partial P_{\mathbf{q}}(t_3,t_2)}{\partial t_2}\Bigg|_{t_2=t_3+0}=-2i\frac{Z^2}{c^2}$$
$$\times\sum_{\mathbf{q}'}g^2_{q'}\frac{(\mathbf{q}'\times\mathbf{v}_1)(\mathbf{q}'\times\mathbf{v}_2)}{q'^2(\omega'-\mathbf{q}'\cdot\mathbf{v}_1)}+O_1(q),$$

where $O_1(q)$ is of first order in q. For q small, noting that

$$\frac{\partial P_{\mathbf{q}}(t_3,t_2)}{\partial t_2}\Bigg|_{t_2=t_3-0}=-2i\frac{Z^2}{c^2}$$
$$\times\sum_{\mathbf{q}'}g^2_{q'}\frac{(\mathbf{q}'\times\mathbf{v}_1)^2}{q'^2(\omega'-\mathbf{q}'\cdot\mathbf{v}_1)}+O_1(q),$$

and that $\mathbf{v}_2\approx\mathbf{v}_1$, we obtain

$$n_{\alpha\mathbf{q}}(\infty)=\frac{Z^2}{c^2}g_q^2\left|\frac{\mathbf{e}_{\alpha\mathbf{q}}\cdot\mathbf{v}_2}{\omega-\mathbf{q}\cdot\mathbf{v}_2+\Delta}-\frac{\mathbf{e}_{\alpha\mathbf{q}}\cdot\mathbf{v}_1}{\omega-\mathbf{q}\cdot\mathbf{v}_1+\Delta}\right|^2, \quad (45)$$

where

$$\Delta=2\frac{Z^2}{c^2}\sum_{\mathbf{q}'}g^2_{q'}\frac{(\mathbf{q}'\times\mathbf{v}_1)^2}{q'^2(\omega'-\mathbf{q}'\cdot\mathbf{v}_1)}. \quad (46)$$

In the nonrelativistic limit $v_0\ll c$, from (46) we obtain $\Delta\approx 4Z^2v_1^2q_c/3\pi c^2$. Equation (45) does not contain the infrared singularity. A deviation from the $n_{\alpha\mathbf{q}}\propto 1/\omega^3$ law with decreasing ω begins at an energy of order Δ. The lower the energy of relative motion of the charged particle and point scatterer, the lower the aforementioned energy. This estimate also holds if the velocity of the particle changes not suddenly but over a time interval that is short compared to the time of production of a low-energy photon.

7. CALCULATING THE DENSITY MATRIX FOR THE CASE OF SYNCHROTRON RADIATION

We now use the above approach to calculate the density matrix of an emitting particle. The exact expression for the density matrix in the representation realized by the transformation (23) has the form

$$\gamma(\mathbf{x},\mathbf{x}',t)=\langle t|\tilde{\psi}^\dagger(\mathbf{x},t)\tilde{\psi}(\mathbf{x}',t)|t\rangle. \quad (47)$$

We calculate (47) in the first approximation, replacing the vector $|t\rangle$ by the initial state vector $|0\rangle$. Using the Baker–Hausdorff rule with proper transformation of the evolution operators $\hat{U}_0(t)$ and $\hat{U}_0^\dagger(t)$, and the operators $\hat{Q}_{\alpha\mathbf{q}}(t)$ in the form (20), we easily find that

$$\gamma(\mathbf{x},\mathbf{x}',t)=\gamma_0(\mathbf{x},\mathbf{x}',t)\exp[-S(\mathbf{x}-\mathbf{x}',t)], \quad (48)$$

where $\gamma_0(\mathbf{x},\mathbf{x}',t)=\phi^*(\mathbf{x},t)\phi(\mathbf{x}',t)$ is the value of the density matrix that does not account for emission and is determined by the wave function $\phi(\mathbf{x},t)$ of the exactly described state of an electron in an external magnetic field. The function $S(\mathbf{x}-\mathbf{x}',t)$ in the exponent is given by

$$S(\mathbf{x}-\mathbf{x}',t)=\sum_{\mathbf{q},\alpha}|Q_{\alpha\mathbf{q}}(t)|^2[1-e^{i\mathbf{q}\cdot(\mathbf{x}-\mathbf{x}')}], \quad (49)$$

which vanishes at $\mathbf{x}=\mathbf{x}'$. As $|\mathbf{x}-\mathbf{x}|\to\infty$, the function (46) acquires its maximum value, equal to the total number of photons emitted by the given moment in time.

The mean momentum of the particle can be calculated as follows:

$$\overline{\mathbf{p}}(t)=\mathbf{p}_0+i\int|\phi(\mathbf{x},t)|^2\nabla' S(\mathbf{x},\mathbf{x}',t)|_{\mathbf{x}=\mathbf{x}'}\,d^3x. \quad (50)$$

This means that the gradient $\nabla' S(\mathbf{x},\mathbf{x}',t)$ determines the rate of decrease of the mean particle momentum due to emission of photons. If the initial state was stationary, $|\phi(\mathbf{x},t)|^2$ does not depend on time. In this case, the mean force acting on the particle is

$$\mathbf{F}_b=i\int|\phi_0(\mathbf{x},t)|^2\nabla' S(\mathbf{x},\mathbf{x}',t)|_{\mathbf{x}'=\mathbf{x}}d^3x. \quad (51)$$

The calculation of the function $S(\mathbf{x},\mathbf{x}',t)$ for the case of synchrotron radiation is similar to the calculation of the photon production rate in Sec. 4. Noting that S actually depends on the difference $\mathbf{r}=|\mathbf{x}-\mathbf{x}'|$, we obtain the value of S averaged over one period:

$$\overline{S(\mathbf{r},t)}=t\frac{Z^2}{c}\int_0^\pi d\theta\,\sin\,\theta\sum_{n=1}^{\infty}n\omega_0\left[\cot^2\theta J_n^2\left(\frac{nv_0}{c}\sin\,\theta\right)+\frac{v_0^2}{c^2}J_n'^2\left(\frac{nv_0}{c}\sin\,\theta\right)\right]\left(1-J_0\left(r\frac{n\omega_0}{c}\sin\,\theta_0\,\sin\,\theta\right)\times\exp\left(ir\frac{n\omega_0}{c}\cos\,\theta_0\,\cos\,\theta\right)\right),\tag{52}$$

where θ_0 is the polar angle of the vector $\mathbf{r}$ with respect to the axis perpendicular to the orbital plane, $r=|\mathbf{r}|$. In the ultrarelativistic case the following approximate formula is more convenient:

$$\overline{S(\mathbf{r},t)}=t\frac{2^{2/3}Z^2\omega_0}{\pi c}\int_0^\infty d\varsigma\varsigma^{1/3}\int_0^\pi d\theta\,\sin\,\theta\left\{\cot^2\,\theta\times\mathrm{Ai}^2\left[\left(\frac{\varsigma}{2}\right)^{2/3}\left(1-\frac{v_0^2}{c^2}\sin^2\,\theta\right)\right]+\frac{v_0^4}{c^4}\frac{2^{2/3}\sin^2\,\theta}{\varsigma^{2/3}}\times\mathrm{Ai}'^2\left[\left(\frac{\varsigma}{2}\right)^{2/3}\left(1-\frac{v_0^2}{c^2}\sin^2\,\theta\right)\right]\right\}\left(1-J_0\left(\sin\,\theta_0\times\sin\,\theta r\frac{\varsigma\omega_0}{c}\right)\exp\left(i\,\cos\,\theta_0\,\cos\,\theta\,r\frac{\varsigma\omega_0}{c}\right)\right),\tag{53}$$

where $\mathrm{Ai}(z)$ is the Airy function and $\mathrm{Ai}'(z)$ is its derivative. Obviously, the imaginary part of the averaged expression for S given by (52) and (53) is zero.

According to Ref. 10, the density matrix (48) describes an ensemble of pure states (in the sense of von Neumann), whose properties are determined by the behavior of $e^{-S(\mathbf{r},t)}$. The expansion of the matrix (48) in the density matrices of these pure states can be written

$$\gamma(\mathbf{x},\mathbf{x}',t)=\int d^3a\Phi_{\mathbf{a}}^*(\mathbf{x},t)\Phi_{\mathbf{a}}(\mathbf{x}',t)N_{\mathbf{a}}(t),\tag{54}$$

where $\Phi_{\mathbf{a}}$ are the wave functions of the pure states, and $N_{\mathbf{a}}(t)d^3a$ are the probabilities that these states are realized at the given moment in time. The functions $\Phi_{\mathbf{a}}(\mathbf{x},t)$ are proportional to the products of the wave function $\phi(\mathbf{x},t)$ and the functions $\chi(\mathbf{x}-\mathbf{a},t)$, where $\chi(\mathbf{x},t)$ is the solution of the integral equation

$$G(\mathbf{r},t)=e^{-S(\mathbf{x}-\mathbf{x}',t)}=\int\chi^*(\mathbf{x}-\mathbf{a},t)\chi(\mathbf{x}'-\mathbf{a},t)d^3a.\tag{55}$$

But what about the existence and uniqueness of the solution of this equation? If we write (55) in the Fourier representation

$$G_{-\mathbf{q}}(t)=\chi_{\mathbf{q}}^*(t)\chi_{\mathbf{q}}(t),$$

the absolute value of the desired function is uniquely defined, but not the phase. However, this uncertainty is a direct consequence of the translation invariance of Eq. (55), whose general solution, therefore, has the form

$$\chi_{\mathbf{q}}=e^{i\alpha}\sqrt{G_{-\mathbf{q}}},$$

with arbitrary real α. A solution exists if the Fourier transform $G_{\mathbf{q}}$ is a real nonnegative quantity. That it is real follows directly from the fact that $\mathrm{Im}\,S(\mathbf{r},t)$ is odd and $\mathrm{Re}\,S(\mathbf{r},t)$ even under inversion; the nonnegativity follows from the fact that $\mathrm{Re}\,S(\mathbf{r},t)$ increases monotonically with distance r.

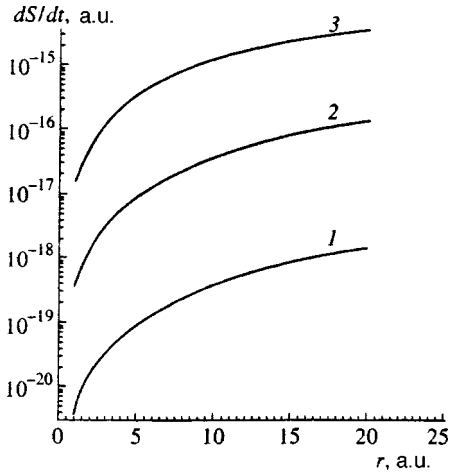

FIG. 1. Rate of buildup of the function $S(\mathbf{r},t)$ determining the spatial width of the electron's wave packet in the FIAN-60 synchrotron as a function of distance r in observations at various angles relative to the direction of the magnetic field: $\theta_0=\pi/200$ (curve *1*), $\pi/20$ (curve *2*), and $\pi/2$ (curve *3*).

The effective size of the localization region for the initial state in the orbital plane is $\delta\rho\sim\sqrt{R/\gamma v_0}$ (see Appendix). The quantity $\delta\rho$ is usually much larger than atomic dimensions. The localization region for the initial state along the magnetic field is infinitely large, which is due to the initial uncertainty in the z-component of the momentum. The latter is obviously determined by the macroscopic parameters of the actual experimental layout.

Radiation can substantially alter the picture, and lead to spatial localization of the emitting particle in a region whose size is of the order of atomic dimensions. To estimate the rate of variation of the widths of the states $\Phi_{\mathbf{a}}(t)$ with the passage of time, the function $S(\mathbf{r},t)$ was calculated numerically for a set of parameters characteristic of the FIAN-60 synchrotron ($E=0.68$ GeV and $R=2$ m). Figure 1 depicts $\bar{\bar{S}}(\mathbf{r})$ as a function of r for three orientations of $\mathbf{r}$. The upper curve illustrates the behavior of this factor in the plane of the electron's orbit.

8. MAIN CONCLUSIONS

The perturbation theory developed in this paper has made it possible to establish that certain fundamental problems of quantum electrodynamics are not invariant when the type of perturbation theory is altered. With respect to the infrared divergence, this is shown by employing the simple example of an emitting particle that undergoes a sudden change in velocity. The results have been obtained for the nonrelativistic case, since the study of emitting relativistic particles requires a detailed analysis of the ultraviolet asymptotic behavior of the integrand in (38).

The method of removing ultraviolet divergences by introducing a cutoff momentum, which was adopted in the present paper, is not covariant under Lorentz transformations, and therefore cannot be used in a consistent relativistic theory. But even preliminary studies show that in the new

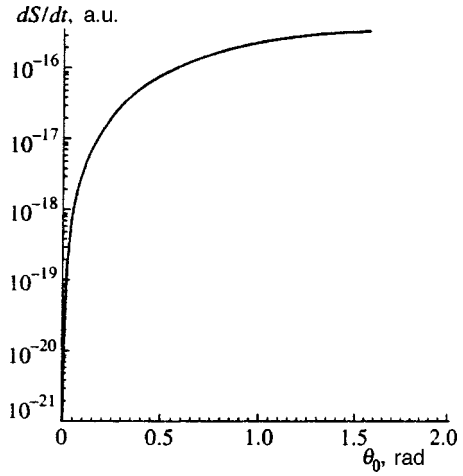

FIG. 2. The same as in Fig. 1 but as a function of the observation angle θ_0 for a fixed distance $r=5$ a.u. The angle θ_0 is measured in radians.

approach the problem of ultraviolet divergence is not catastrophic, in contrast to the case in ordinary quantum electrodynamics. It is to be hoped that further research will lead to progress in understanding this problem.

Density matrix calculations have shown that reduction of the spatial dimensions of the localization region for the emitting electrons to atomic dimensions can be achieved over a macroscopically long time interval τ_c of some tenths of a second. Can the present theory be applied to such long times? The situation is complicated by the fact that in the course of one orbital revolution, the particle is subject to a solenoidal electric field that balances the loss of energy to photon emission. If we assume that this field acts during a time interval so short that it only accelerates the particle's wave packet as a whole and is unable to change the particle's internal parameters substantially, then there is no reason why to do estimates we cannot extend the theory to the entire duration of the particle's motion in the synchrotron. The time τ_c is much shorter than it takes the packet to spread due to the nonequidistant nature of the spectrum of the transverse-motion levels. What is observed is an anisotropy in the packet's width: the packet is most strongly squeezed perpendicular to the magnetic field, and least strongly parallel to the field. Figure 2 depicts the angular dependence of the same quantity as in Fig. 1 at a fixed value of r. The considerable elongation of the packet in the direction of the magnetic field is obvious.

The possibility of strong spatial localization of the emitting particles means that if the acceleration cycle in the synchrotron is long enough, the motion of the particles can be described to high accuracy by the equations of classical mechanics. Nevertheless, this does not mean that the intensity of the radiation must agree with the predictions of classical electrodynamics. Indeed, a localized state in quantum mechanics is completely different in its properties from a localized state in Newton's classical theory. The justification for using Newton's equations of motion to calculate the paths followed by wave packets is provided by Ehrenfest's theorem, but the decisive factor in calculating the intensity of the radiation is the momentum of the particle, rather than the position.

In quantum mechanics, a state with a definite momentum is completely delocalized, and in this way differs substantially from states of type $\Phi_{\mathbf{a}}$. There is thus no way in which we can intuitively interpret calculations of the characteristics of radiation using classical ideas. The characteristic common feature of the formulas derived in this paper is the fact that allowing for the mutual interaction of the emitted photons reduces the radiative intensity. A similar result was obtained by Landau and Pomeranchuk,[12] who studied the radiation emitted by charged particles moving in continuous media (the Landau–Pomeranchuk effect).

The physics of this phenomenon amount to the fact that random collisions of an emitting particle with particles of the medium can reduce the path length over which the radiative intensity builds up coherently. Something similar is observed when photons are emitted into vacuum: multiple emission of photons can mimic the multiple collisions in a continuous medium that lead to a reduction in radiative intensity.

The numerical calculations were done together with V. A. Aleksandrov. This work was supported by a grant from the Russian Fund for Fundamental Research (Grant No. 97-02-160-58).

APPENDIX: COHERENT STATES OF A RELATIVISTIC ELECTRON IN A UNIFORM MAGNETIC FIELD

Let

$$\mathbf{A}=(-\tfrac{1}{2}yH_0,\ \tfrac{1}{2}xH_0,\ 0),$$

where H_0 is the strength of the magnetic field directed along the z axis. The motion of an electron in such a field, which obeys the Dirac equation, has been the topic of numerous studies (see, e.g., Refs. 13–15). The solution given below differs from the well-known one only in some details.

We introduce the lowering operators for two independent oscillators:

$$\hat{a}_1=\frac{1}{2}\sqrt{m\omega_L}\left(x+\frac{i\hat{p}_x}{m\omega_L}-iy+\frac{\hat{p}_y}{m\omega_L}\right),$$

$$\hat{a}_2=\frac{1}{2}\sqrt{m\omega_L}\left(x+\frac{i\hat{p}_x}{m\omega_L}+iy-\frac{\hat{p}_y}{m\omega_L}\right),$$

where $\omega_L=|e|H_0/2mc$ is the Larmor frequency. The frequency of the first independent oscillator is twice the Larmor frequency, $\omega_1=2\omega_L$, while the frequency of the second oscillator is zero. The set of lowering and raising operators (which are conjugates of lowering operators) satisfies the standard Bose commutation relations. The operators $\hat{a}_1$ and $\hat{a}_1^\dagger$ describe the orbital motion of an electron in a magnetic field, while the operators $\hat{a}_2$ and $\hat{a}_2^\dagger$ describe the position, fluctuations, and other characteristics of the center of the osculating circular orbit, whose mean radius is R.

We next introduce the matrix operator

$$\hat{D}=\begin{pmatrix}\hat{p}_z & -2i\sqrt{m\omega_L}\hat{a}_1\\ 2i\sqrt{m\omega_L}\hat{a}_1^\dagger & -\hat{p}_z\end{pmatrix}.$$

The energies of the electron's quantum states are

$$E_\xi = \sqrt{m^2c^4 + p^2c^2 + 4\omega_L mc^2\left(n_1 + \sigma + \frac{1}{2}\right)}, \qquad \text{(A1)}$$

where the label $\xi=(n_1,n_2,\sigma,p)$ simply indicates the set of quantum numbers in parentheses. The $n_1,n_2=0,1,2,\ldots$ label the quantum states of the independent oscillators, with n_1 being the principal quantum number. We denote the projection of momentum on the z axis by p. The discrete variable σ takes two values, $\pm 1/2$, corresponding to two possible projections of spin on the direction of the magnetic field. The bispinor describing a stationary state of an electron in a magnetic field is given by

$$\psi_\xi(\mathbf{r},t) = \frac{1}{\sqrt{2E_\xi}} \begin{pmatrix} \sqrt{E_\xi + mc^2}\,\varphi_\xi(\mathbf{r}) \\ \dfrac{c}{\sqrt{E_\xi + mc^2}}\,\hat{D}\varphi_\xi(\mathbf{r}) \end{pmatrix} e^{-iE_\xi t}, \qquad \text{(A2)}$$

where $\varphi_\xi(\mathbf{r})$ is a spinor of the form

$$\varphi_\xi(\mathbf{r}) = e^{ipz}\frac{1}{\sqrt{n_1!n_2!}}(\hat{a}_1^\dagger)^{n_1}(\hat{a}_2^\dagger)^{n_2}\varphi_{0\sigma}(\rho). \qquad \text{(A3)}$$

Here

$$\varphi_{0\sigma}(\rho) = \sqrt{\frac{m\omega_L}{\pi}}\exp\left(-\frac{m\omega_L\rho^2}{2}\right)\chi_\sigma,$$

with

$$\rho^2 = x^2 + y^2, \quad \chi^*_{1/2} = (1,0), \quad \chi^*_{-1/2} = (0,1),$$

and the normalization length along the z axis is taken equal to unity.

An arbitrary solution of the Dirac equation is a linear combination of bispinors of type (A2). Just what linear combination corresponds to the initial state of an electron in the synchrotron? For standard values of synchrotron parameters (for example, for the FIAN-60 synchrotron), the mean value of n_1 is very large (of order 10^{13}), and if the assumption that photons are emitted largely independently of one another is true, so is Poisson's law. In this case, the expected value of the relative fluctuation of the number n_1 is extremely small, $\lambda = \overline{\Delta n_1}/\overline{n_1} \sim 10^{-6}$. Hence, essentially all terms of the desired linear combination can be expanded in powers of λ, with the result that

$$E_\xi \approx E_{\sigma p} + \frac{2\omega_L}{\gamma_{\sigma p}}\Delta n_1,$$

where $E_{\sigma p}$ is the value of E_ξ at $\xi=(\overline{n_1},\overline{n_2},\sigma,p)$, and $\gamma_{\sigma p} = E_{\sigma p}/mc^2$.

We see that the spectrum is essentially uniformly spaced, with the levels being separated by the mean orbital frequency $\omega_{\sigma p} = 2\omega_L/\gamma_{\sigma p}$ of the electron about the magnetic field. When the relative fluctuation of n_1 is small, we can put $E_\xi \approx E_{\sigma p}$ in all nonexponential factors.

The linear combination corresponding to the above requirements has the form

$$\psi_{\sigma p}(\mathbf{r},t) = \frac{1}{\sqrt{2E_{\sigma p}}} \begin{pmatrix} \sqrt{E_{\sigma p} + mc^2}\,\varphi_{\sigma p}(\mathbf{r},t) \\ \dfrac{c}{\sqrt{E_{\sigma p} + mc^2}}\,\hat{D}\varphi_{\sigma p}(\mathbf{r},t) \end{pmatrix} \times \exp(-i\Delta E_{\sigma p}t), \qquad \text{(A4)}$$

where $\Delta E_{\sigma p} = E_{\sigma p} - \omega_{\sigma p}\overline{n_1}$, and

$$\varphi_{\sigma p}(\mathbf{r},t) = e^{ipz}\exp[\sqrt{\overline{n_1}}(e^{i\alpha_1}\hat{a}_1^\dagger(t) - e^{-i\alpha_1}\hat{a}_1(t)) + \sqrt{\overline{n_2}}(e^{i\alpha_2}\hat{a}_2^\dagger(t) - e^{-i\alpha_2}\hat{a}_2(t))]\varphi_{0\sigma}(\rho),$$

with $\hat{a}_1^\dagger(t) = \hat{a}_1^\dagger \exp(i\omega_{\sigma p}t)$ and $\hat{a}_2^\dagger(t) = \hat{a}_2^\dagger$, where α_1 and α_2 are constant phases; the momentum along the z axis is assumed equal to p. The components of the current density vector in the state (A4) are

$$j^x_{\sigma p} = \frac{2c^2}{E_{\sigma p}}\sqrt{m\omega_L\overline{n_1}}\,|\varphi_{\sigma p}(\mathbf{r},t)|^2\sin(\omega_{\sigma p}t + \alpha_1),$$

$$j^y_{\sigma p} = -\frac{2c^2}{E_{\sigma p}}\sqrt{m\omega_L\overline{n_1}}\,|\varphi_{\sigma p}(\mathbf{r},t)|^2\cos(\omega_{\sigma p}t + \alpha_1),$$

$$j^z_{\sigma p} = -\frac{c^2 p}{E_{\sigma p}}|\varphi_{\sigma p}(\mathbf{r},t)|^2. \qquad \text{(A5)}$$

The packet's rms width in the radial direction in the state (A4) is determined by the radial behavior of the function $\varphi_{0\sigma}(\rho)$, and can be estimated to be $\Delta\rho = \sqrt{2Rc/E_{\sigma p}}$. In the azimuthal direction, the stationary states of type (A2) are completely delocalized. Indeed, in these states the angular momentum is well-defined, and by virtue of the uncertainty relation for action–angle variables, they cannot be localized in angle.

In contrast, the state (A4) has no definite angular momentum, but its angular width is limited, and is of order $\Delta\phi \sim 1/\sqrt{\overline{n_1}}$ in the azimuthal direction (we assume that the uncertainty in the position of the orbit's center is much smaller than the orbit's radius, so that $n_2 \ll n_1$), which after being multiplied by the orbit's radius yields a distance roughly equal to $\Delta\rho$ (for the FIAN-60 synchrotron this distance is about one micrometer).

The packet width along the z axis is governed by such macroscopic parameters of the device as the diaphragm widths, and for this reason it can exceed the radial or azimuthal width many times over. In this case the packet can be represented by a linear combination of states of type (A4):

$$\psi_\sigma(\mathbf{r},t) = \sum_p C_p\psi_{\sigma p}(\mathbf{r},t), \qquad \text{(A6)}$$

where the constants C_p satisfy the normalization condition and guarantee, e.g., a Gaussian dependence on the z projection of the momentum with midpoint at $p=0$:

$$C_p = (2\pi\delta_0^2)^{1/4}e^{-p^2\delta_0^2/4}.$$

If we assume that the spatial width of the packet along the z axis is much greater than the radial width, then $\delta_0 \gg \Delta\rho$, and in this case the state (A6) is associated with a small symmetric ellipsoid elongated in the direction of the magnetic field and revolving in this orientation in a circular orbit about an

axis parallel to z. To estimate the time of packet spread in the radial or azimuthal direction, we must keep the next term in the expansion of the energy E_ξ in powers of Δn_1. This yields the value of the time of packet spreading due to the nonequidistant levels of transverse motion, $\tau_1 \sim \gamma_0 R^2 \sqrt{\Delta n_1}$. Here γ_0 is the Lorentz factor for the electron beam in a synchrotron. For the FIAN-60 synchrotron the time τ_1 was estimated to be about ten seconds.

*)E-mail: Phys@chgu.chuvashia.su

[1] I. M. Ternov and V. V. Mikhaĭlin, *Synchrotron Radiation* [in Russian], Énergoatomizdat, Moscow (1986).
[2] A. I. Nikishov, Zh. Éksp. Teor. Fiz. **110**, 510 (1996) [JETP **83**, 274 (1996)].
[3] V. I. Ritus, Zh. Éksp. Teor. Fiz. **110**, 526 (1996) [JETP **83**, 282 (1996)].
[4] G. F. Efremov, Zh. Éksp. Teor. Fiz. **110**, 1629 (1996) [JETP **83**, 896 (1996)].
[5] D. D. Iwanenko and A. A. Sokolov, *Klassische Feldtheorie*, Akademie-Verlag, Berlin (1953).
[6] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 2nd ed., Pergamon Press, Oxford (1982).
[7] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Sums, Series and Products*, Academic Press, New York (1980).
[8] G. M. Filippov, Poverkhnost' No. 3, 52 (1997).
[9] G. M. Filippov, Poverkhnost' No. 12, 73 (1995).
[10] G. M. Filippov, Zh. Éksp. Teor. Fiz. **101**, 1634 (1992) [Sov. Phys. JETP **74**, 871 (1992)].
[11] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed., Pergamon Press, Oxford (1975).
[12] L. D. Landau and I. Ya. Pomeranchuk, Dokl. Akad. Nauk SSSR **92**, 735 (1953).
[13] V. V. Dodonov, E. V. Kurmyshev, and V. I. Man'ko, in *Proc. of the P. N. Lebedev Physics Institute* [in Russian], Vol. **176**, Moscow (1986), p. 128.
[14] V. G. Bagrov, D. M. Gitman, and V. D. Skarzhinskiĭ, in *Proc. of the P. N. Lebedev Physics Institute* [in Russian], Vol. **176**, Moscow (1986), p. 151.
[15] V. G. Bagrov, D. M. Gitman, and V. D. Skarzhinskiĭ, Preprint P. N. Lebedev Physics Institute No. 101, Moscow (1986).

Translated by Eugene Yankovsky

Recoil correction to hydrogen energy levels: A revision

A. S. Yelkhovsky

Budker Institute of Nuclear Physics and Physics Department, Novosibirsk University, 630090 Novosibirsk, Russia

(Submitted 26 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 865–879 (March 1998)

Recent calculations of the order of $(Z\alpha)^4(m/M)$ Ry pure recoil correction to the hydrogen energy levels are critically revised. The origins of errors made in the previous studies are elucidated. In the framework of a systematic approach, a new result is obtained for the S levels. It amounts to -16.4 kHz in the ground state and -1.9 kHz in the $2S$ state. © *1998 American Institute of Physics.* [S1063-7761(98)00803-8]

1. INTRODUCTION

The correction to the S levels of hydrogen atom, which is first order in m/M and fourth order in $Z\alpha$, has become recently a point of controversy. Initially, this correction was calculated in Ref. 1. A different result for the same correction was subsequently obtained in Ref. 2. While the same (exact in $Z\alpha$) starting expression for the pure recoil correction was employed in both papers, the methods of calculation and, in particular, the regularization schemes used were rather different. To resolve the discrepancy between the two results, an attempt was undertaken in Ref. 3 to prove the correctness of the earlier result of Ref. 1 by applying the method of calculation used by the present author in Ref. 2. An extra contribution due to the peculiarities of the regularization procedure was found by the authors of Ref. 3, which exactly compensates for the difference between the result of Ref. 2 and the result of Ref. 1. This finding has led the authors of Ref. 3 to the conclusion that ''discrepancies between the different results for the correction of the order of $(Z\alpha)^6(m/M)$ to the energy levels of the hydrogen-like ions are resolved and the correction of this order is now firmly established.''

Assuming that the criticism of Ref. 3 is completely valid, we nevertheless cannot agree with the conclusion cited above. The point is that in emphasizing the importance of an explicit regularization of divergent expressions, the authors of Ref. 3 pay no attention to an accurate matching of regularized contributions.

In fact, one usually starts from an exact expression, which can be easily checked to have a finite value. Different approximations must then be used to handle this expression at different scales. In this way some auxiliary parameter(s), which enable us to separate the applicability domains for different approximations, are introduced. Finally, a necessary condition for the sum of the calculated contributions to be correct is its independence of any scale-separating parameter.

In the present paper we systematically pursue this line of reasoning for a recalculation of the order of $(Z\alpha)^6 m^2/M$ correction to the hydrogen energy levels. We discuss only the S levels since for higher angular-momentum levels the result has actually been firmly established.[2,4] Since the controversy mentioned above concerns details of a regularization at the subatomic scale, the dependence of the results on the principal quantum number n is also known. We therefore perform all the calculations for the ground state and then restore the n dependence in the final result.

To make the presentation self-contained we rederive some known results, using sometimes new approaches. The general outline of the problem is given in Sec. 2. Sections 3, 4, and 5 deal with the Coulomb, magnetic, and seagull contributions, respectively. The correspondence between various results is discussed in Sec. 6. A couple of minor computational issues are addressed in the Appendices.

The Coulomb gauge of the electromagnetic potentials and relativistic units $\hbar = c = 1$ are used throughout the paper. Leaving aside the radiative corrections, we set $Z = 1$ in what follows.

2. GENERAL OUTLINE

The first recoil correction to a bound state energy of the relativistic electron in the Coulomb field is an average value of the nonlocal operator,[5–7,1,8]

$$\Delta E_{\rm rec} = -\frac{1}{M}\int \frac{d\omega}{2\pi i}\,\langle (\mathbf{p}-\mathbf{D}(\omega,\mathbf{r}'))G(\mathbf{r}',\mathbf{r}|E+\omega) \times (\mathbf{p}-\mathbf{D}(\omega,\mathbf{r}))\rangle, \tag{1}$$

which is taken over an eigenstate of the Dirac equation in the Coulomb field,

$$H\psi(\mathbf{r}) = E\psi(\mathbf{r}), \quad H = \boldsymbol{\alpha}\mathbf{p} + \beta m - \frac{\alpha}{r}. \tag{2}$$

In (1), $\mathbf{p}$ is the electron momentum operator, $\mathbf{D}(\omega,\mathbf{r})$ describes an exchange by the transverse (magnetic) quantum,

$$\mathbf{D}(\omega,\mathbf{r}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\, e^{i\mathbf{k}\mathbf{r}}\,\frac{4\pi\alpha\boldsymbol{\alpha}_k}{k^2-\omega^2}, \quad \boldsymbol{\alpha}_k \equiv \boldsymbol{\alpha} - \frac{\mathbf{k}(\boldsymbol{\alpha}\mathbf{k})}{k^2}, \tag{3}$$

and

$$G(\mathbf{r}',\mathbf{r}|E+\omega) = \left(E+\omega-\boldsymbol{\alpha}\mathbf{p}'-\beta m+\frac{\alpha}{r'}\right)^{-1}\delta(\mathbf{r}'-\mathbf{r}) \tag{4}$$

is the Green's function for the Dirac equation in the Coulomb field. The integration contour in (1) goes from minus infinity to zero below the real axis, rounds zero from above, and then proceeds to plus infinity above the real axis.

Since we will calculate the correction (1) perturbatively, i.e., as a power series in α, we can decompose (1) into three parts:

$$\Delta E_{\rm rec} = \mathscr{C} + \mathscr{M} + \mathscr{S}. \tag{5}$$

Specifically, the Coulomb, the magnetic, and the seagull contributions, which correspond to the **pp**, **pD**+**Dp**, and **DD** terms from (1), respectively.

3. COULOMB CONTRIBUTION

It is natural to continuously transform the integration contour into the sum of two subcontours, thus splitting the Coulomb contribution into two terms,

$$\mathscr{C} = \left\langle \frac{p^2}{2M} \right\rangle - \frac{1}{M} \langle \mathbf{p} \Lambda_- \mathbf{p} \rangle, \tag{6}$$

where Λ_- is the projector to the set of negative-energy Dirac–Coulomb eigenstates. The former term in (6) results from the integration along the upper half of the infinite circumference and its value is determined by the atomic scale $p \sim m\alpha$. Being the average of the local operator, this term can be easily calculated exactly. The latter term in (6) arises as an integral along the contour C_-, wrapping the half-axis $(-\infty,0)$ in the counterclockwise direction. This term is completely saturated by momenta from the relativistic scale $p \sim m$. It can therefore be calculated without any regularization:[1,2]

$$-\frac{1}{M} \langle \mathbf{p} \Lambda_- \mathbf{p} \rangle_{\alpha^6} = \frac{m^2 \alpha^6}{M}. \tag{7}$$

4. MAGNETIC CONTRIBUTION

Using the identity

$$\langle \mathbf{p} G \mathbf{D} + \mathbf{D} G \mathbf{p} \rangle = \frac{1}{\omega} \langle [\mathbf{p}, H] G \mathbf{D} + \mathbf{D} G [H, \mathbf{p}] + \{\mathbf{p}, \mathbf{D}\} \rangle, \tag{8}$$

which follows directly from the equation for the Green's function, we can extract from the general expression for the magnetic contribution,

$$\mathscr{M} = \frac{1}{M} \int \frac{d\omega}{2\pi i} \langle \mathbf{p} G \mathbf{D} + \mathbf{D} G \mathbf{p} \rangle, \tag{9}$$

its local part,

$$\frac{1}{M} \int_{C_-} \frac{d\omega}{2\pi i} \frac{1}{\omega} \langle \{\mathbf{p}, \mathbf{D}(\omega, \mathbf{r})\} \rangle = -\frac{1}{2M} \langle \{\mathbf{p}, \mathbf{D}(0, \mathbf{r})\} \rangle. \tag{10}$$

Due to the rapid convergence of the integral in (9) at infinity, we can reduce the integration contour to C_-. By virtue of the virial relations (see Ref. 9 and the references cited there), the sum of the local parts of the Coulomb and magnetic contributions takes a simple form:[6]

$$\left\langle \frac{p^2}{2M} - \frac{1}{2M} \{\mathbf{p}, \mathbf{D}(0, \mathbf{r})\} \right\rangle = \frac{m^2 - E^2}{2M}. \tag{11}$$

Physically, this contribution to the recoil correction is induced by an instantaneous part of the electron-nucleus interaction.

4.1. Long distances

Immediate integration with respect to ω in (9) gives:[2]

$$\mathscr{M} = -\frac{\alpha}{M} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \left\langle \mathbf{p} \left(\sum_+ \frac{|m\rangle\langle m|}{k + E_m - E} - \sum_- \frac{|m\rangle\langle m|}{E - E_m + k} \right) \frac{4\pi \boldsymbol{\alpha}_k}{k} e^{i\mathbf{k}\mathbf{r}} \right\rangle, \tag{12}$$

where $\Sigma_{+(-)}$ represents the sum over discrete levels supplied by the integral over the positive-(negative-) energy part of the continuous spectrum.

4.1.1. Positive Energies. In the leading nonrelativistic approximation, the first term in Eq. (12) reads

$$\mathscr{M}_+ = \frac{\alpha}{Mm} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \left\langle \mathbf{p} \mathscr{G}(\mathbf{r}', \mathbf{r} | E - k) \frac{4\pi e^{i\mathbf{k}\mathbf{r}}}{k} \mathbf{p}_k \right\rangle, \tag{13}$$

where $\mathscr{G}(\mathbf{r}', \mathbf{r} | E - k)$ is the Green's function for the Schrödinger equation in the Coulomb field, and the average is taken over the nonrelativistic wave function. For the ground state we use

$$\psi(\mathbf{r}) = \sqrt{\frac{(m\alpha)^3}{\pi}} e^{-m\alpha r}, \quad E = -\frac{m\alpha^2}{2}. \tag{14}$$

Only the p-wave term from the partial expansion

$$\mathscr{G}(\mathbf{r}', \mathbf{r} | \omega) = \sum_l (-)^l (2l+1) P_l(\mathbf{n}'\mathbf{n}) \mathscr{G}_l(r', r | \omega) \tag{15}$$

survives the integration over the angles:

$$\mathscr{M}_+ = -\frac{m\alpha^3}{M\pi} \int_0^\infty dk\, k \int_{-1}^1 dx (1 - x^2) \times \langle \mathscr{G}_1(r', r | E - k) e^{ikrx} \rangle. \tag{16}$$

For the nonrelativistic Green's function in the Coulomb field we use the integral representation from Ref. 10,

$$\mathscr{G}_1\left(r', r \middle| -\frac{\kappa^2}{2m}\right) = \frac{im}{2\pi\sqrt{r'r}} \times \int_0^\pi \frac{ds}{\sin s} \frac{\exp\{i[2(m\alpha/\kappa)s + \kappa(r'+r)/\tan s]\}}{1 - \exp\{i2(m\alpha/\kappa)\pi\}} \times J_3\left(\frac{2\kappa\sqrt{r'r}}{\sin s}\right). \tag{17}$$

The integrals over r and r' in (16) can be easily calculated after expanding the Bessel function in a power series. The result can be expressed in the form

$$\mathscr{M}_+ = \frac{2^7 3 m^5 \alpha^6}{M\pi} \int_0^\infty \frac{dk\, k}{\kappa^5} \times \int_{-1}^1 dx (1-x^2) \int_C dt\, \frac{t^{1-m\alpha/\kappa}}{(a-bt)^4} \times \frac{1}{1-\exp\{i2\pi m\alpha/\kappa\}}, \tag{18}$$

where $\kappa = \sqrt{2m(k-E)}$, the contour C is the unit circumference $|t|=1$ directed clockwise, and

$$a = \left(1+\frac{m\alpha}{\kappa}\right)\left(1+\frac{m\alpha}{\kappa}-\frac{ikx}{\kappa}\right),$$

$$b = \left(1-\frac{m\alpha}{\kappa}\right)\left(1-\frac{m\alpha}{\kappa}+\frac{ikx}{\kappa}\right).$$

Integration by parts conveniently extracts from the last integral in (18) the terms which are nonvanishing at large momenta:

$$\mathscr{M}_+ = -\frac{2^5 m^2 \alpha^5}{M\pi} \int_0^1 dy (1-y^2) \int_{-1}^1 dx (1-x^2) F(x,y), \tag{19}$$

where

$$F(x,y) = \frac{2}{b(a-b)^3} - \frac{1-y}{b^2(a-b)^2} - \frac{y(1-y)}{ab^2(a-b)} + \frac{1-y^2}{a^2b^2} - \frac{y(1-y^2)}{a^3 b} \int_0^1 \frac{dt\, t^{-y}}{1-(b/a)t}. \tag{20}$$

Here we introduced a new integration variable $y \equiv m\alpha/\kappa$. Since

$$\frac{k}{\kappa} = \frac{\alpha}{2}\frac{1-y^2}{y}, \tag{21}$$

we can do a power series expansion of (19) with respect to α up to the first order by expanding the integrand with respect to y up to the first order (note that $a-b = 4y - 2ikx/\kappa$):

$$(1-y^2)F(x,y) \approx \frac{2}{(a-b)^3} - \frac{1}{2(a-b)} + \frac{1}{2} - \frac{y^2}{2(a-b)} - \frac{y}{2} + y\,\ln(a-b). \tag{22}$$

Here the last term emerges as a result of expansion of the integral in (20),

$$\int_0^1 \frac{dt\, t^{-y}}{1-(b/a)t} = \frac{1}{1-y} F\left(1,1-y;2-y;\frac{b}{a}\right), \tag{23}$$

where $F(1,1-y;2-y;b/a)$ is the Gauss hypergeometric function. Integrating now (22) first with respect to x, and then with respect to y, from 0 to some y_0 ($\alpha^{1/2} \ll y_0 \ll 1$), we obtain

$$\int_0^{y_0} dy (1-y^2) \int_{-1}^1 dx (1-x^2) F(x,y) \approx \frac{\pi}{32\alpha} - \frac{1}{48y_0^2} - \frac{1}{12}\ln\frac{4y_0^2}{\alpha} + \frac{1}{48} - \frac{1}{9} + \frac{2y_0}{3} + \frac{2y_0^2}{3} \times \ln 4y_0 - \frac{3y_0^2}{4} - \frac{\pi\alpha}{32}. \tag{24}$$

On the other hand, we can ignore α in $F(x,y)$ in the interval $[y_0,1]$. In the sum of two integrals, the dependence on the auxiliary parameter y_0 disappears, and we obtain the result

$$\mathscr{M}_+ = \frac{m^2\alpha^5}{M\pi}\left\{-\frac{\pi}{\alpha} + \frac{8}{3}\ln\frac{1}{\alpha} + \frac{8}{3}\ln\frac{\mathrm{Ry}}{\langle E\rangle_{1S}} + \frac{16}{3}\ln 2 + \frac{32}{9} - \pi\alpha\right\}. \tag{25}$$

Here we introduce the Bethe logarithm[11]

$$16\int_0^1 dy\, y\, \frac{F(1,1-y;2-y;((1-y)/(1+y))^2)-1}{(1+y)^4(1-y)} = \ln\frac{\mathrm{Ry}}{\langle E\rangle_{1S}} + 2\ln 2 + \frac{11}{6}. \tag{26}$$

In (25), the order α^4 term is the lowest-order contribution to (10), the order α^5 terms are in accord with the result of Salpeter,[12] and the order α^6 term coincides with the retardation correction, which is found in Ref. 2, Eq. (14), by a different method.

It can be easily seen that the order α^6 contribution to the positive-energy part of (12) is exhausted by the sum of the contributions to (10) and (25). Actually, relativistic corrections are at least of the α^2 relative order. The retardation reveals itself beginning with the α^5 order (25). Hence the relativistic corrections for the retardation are at least of the order of α^7.

4.1.2. Negative Energies. Virtual transitions to the negative-energy states give rise to the second term in (12). In the leading nonrelativistic approximation, it is[2]

$$\mathscr{M}_- = \frac{\alpha^2}{4m^2 M} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \left\langle \frac{4\pi}{k'^2} \frac{4\pi \mathbf{k}_{k'}^2}{k^2} \right\rangle, \tag{27}$$

where $\mathbf{k}_{k'} = \mathbf{k} - \mathbf{k}'(\mathbf{k}\mathbf{k}')/k'^2$, and $\mathbf{k}' = \mathbf{p}' - \mathbf{p} - \mathbf{k}$; here $\mathbf{p}$ and $\mathbf{p}'$ are the arguments of the wave function and its conjugate, respectively. The integral over k diverges logarithmically (the leading linear divergence vanishes due to the numerator, which at $k\to\infty$ becomes transverse to itself, and hence rises only like k, not k^2). To treat this divergence we use the following formal trick:[2] subtract from (27) the same expression with $k'^2+\lambda^2$ substituted in place of k'^2. For $\lambda \gg m\alpha$, the subtracted term is completely determined by a scale much less than the atomic scale, so that we find the term below by using a relativistic approach.

The regularized version of (27) can be written in the form

$$\mathscr{M}_- - \mathscr{M}_-^r = -\frac{\alpha^2}{4m^2 M} \left\langle (p_i' - p_i) \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{4\pi k_j}{k^2} \times \left(\delta_{ij} - \frac{k_i' k_j'}{k'^2}\right)\left(\frac{4\pi}{k'^2} - \frac{4\pi}{k'^2+\lambda^2}\right)\right\rangle. \tag{28}$$

In the coordinate representation, the integral above is

$$\frac{in_j}{r^2}\left(\delta_{ij}-\frac{\partial_i\partial_j}{\lambda^2}\right)\frac{1-e^{-\lambda r}}{r}=\frac{in_i}{r^2}\int_0^{-\lambda}d\sigma\left(1-\frac{\sigma^2}{\lambda^2}\right)e^{\sigma r}. \tag{29}$$

After substitution into (28) it gives

$$\mathscr{M}_-^r-\mathscr{M}_-^r=-\frac{\alpha^2}{4m^2M}\left\langle 4\pi\delta(\mathbf{r})\int_0^{-\lambda}d\sigma\left(1-\frac{\sigma^2}{\lambda^2}\right)+\frac{1}{r^2}\int_0^{-\lambda}d\sigma\sigma\left(1-\frac{\sigma^2}{\lambda^2}\right)e^{\sigma r}\right\rangle. \tag{30}$$

Finally, the result of a trivial calculation of the average over the ground state is

$$\mathscr{M}_-^r-\mathscr{M}_-^r=\frac{m^2\alpha^6}{M}\left(2\ln\frac{\varepsilon}{\alpha}-1\right), \tag{31}$$

where $\varepsilon\equiv\lambda/2m$.

4.2. Short distances

Since in the nonrelativistic approximation the subtracted term, $\mathscr{M}_-^r$, is ultraviolet divergent, we must calculate it beyond this approximation, i.e., using a relativistic approach. It is more convenient in this approach to postpone the integration over ω to the last stage of calculation. As we will see below, the reversed order of integration (first over space variables, then over frequency) makes the calculations quite simple. The price for the technical advantage is that a regulator contribution is calculated not only for the negative-energy part, but also for the positive-energy part of **M**. Surely, the instantaneous contribution can be put aside, so that only two first terms from the right-hand side of (8) are considered below.

For the subtracted term, we have the new expansion parameter, $m\alpha/\lambda$, and hence the Coulomb interaction during a single magnetic exchange can be treated perturbatively. The order $m\alpha^6/M$ contributions arise due to only two first terms of the Green's function expansion in the Coulomb interaction, $G^{(0)}$ and $G^{(1)}$. Let us begin with the second contribution:

$$\mathscr{M}_G^r=\frac{2}{M}\int_{C_-}\frac{d\omega}{2\pi i}\frac{1}{\omega}\langle[\mathbf{p},H]G^{(1)}\mathbf{D}^r\rangle, \tag{32}$$

Here

$$\mathbf{D}^r=\int\frac{d^3\mathbf{k}}{(2\pi)^3}e^{i\mathbf{kr}}\frac{4\pi\alpha\boldsymbol{\alpha}_k}{k^2+\lambda^2-\omega^2}.$$

and we can disregard the atomic momenta in comparison with λ and m:

$$\mathscr{M}_G^r=-\frac{\alpha^3\psi^2}{\pi M}\int_{C_-}\frac{d\omega}{i\omega}\times\left\langle\frac{4\pi\mathbf{p}'}{p'^2}\frac{2m+\omega+\boldsymbol{\alpha}\mathbf{p}'}{p'^2-\Omega^2}\frac{4\pi}{q^2}\frac{\omega+\boldsymbol{\alpha}\mathbf{p}}{p^2-\Omega^2}\frac{4\pi\boldsymbol{\alpha}_p}{p^2-\mathscr{K}^2}\right\rangle. \tag{33}$$

The notations of Ref. 2 are used: $\psi^2\equiv|\psi(0)|^2$, the angle brackets denote here integrations over $\mathbf{p}$ and $\mathbf{p}'$ together with the average over the spinor $u_\alpha=\delta_{\alpha1}$, $\mathbf{q}=\mathbf{p}'-\mathbf{p}$, and

$$\mathscr{K}\equiv\sqrt{\omega^2-\lambda^2},\quad \Omega\equiv\sqrt{2m\omega+\omega^2}.$$

The average over the spin degrees of freedom gives

$$\langle(2m+\omega+\boldsymbol{\alpha}\mathbf{p}')(\omega+\boldsymbol{\alpha}\mathbf{p})\boldsymbol{\alpha}_p\mathbf{p}'\rangle=\omega\mathbf{p}'^2_p=\omega\mathbf{p}'_p\mathbf{q}. \tag{34}$$

Then, after transition to the coordinate representation we obtain

$$\mathscr{M}_G^r=\frac{2\alpha^3\psi^2}{mM}\int_{C_-}\frac{d\omega}{i\omega}\int_0^\infty dr\left(\partial_i\frac{e^{i\Omega r}-1}{\Omega^2 r}\right)\times n_j\left[\left(\delta_{ij}+\frac{\partial_i\partial_j}{\Omega^2}\right)\frac{e^{i\Omega r}-1}{r}-(\Omega\to\mathscr{K})\right]. \tag{35}$$

The integration over r is simple but lengthy. The result is

$$\mathscr{M}_G^r=-\frac{\alpha^3\psi^2}{2mM}\int_{C_-}\frac{d\omega}{i\omega}\left\{\frac{\mathscr{K}}{\Omega}-\frac{\mathscr{K}^2}{\Omega^2}\ln\left(1+\frac{\Omega}{\mathscr{K}}\right)+(\Omega\leftarrow\mathscr{K})+2\ln\left(1+\frac{\mathscr{K}}{\Omega}\right)\right\}. \tag{36}$$

Here the contour of integration goes counterclockwise around the cut that connects points $-2m$ and $-\lambda$. According to the Feynman rules, $\Omega=i|\Omega|$, while $\mathscr{K}=+(-)|\mathscr{K}|$ on the lower (upper) edge of this cut. Since the integrand is regular at small ω, we can set $\lambda=0$ (recall that $\lambda\ll m$) and obtain

$$\mathscr{M}_G^r=\frac{\alpha^3\psi^2}{mM}\int_0^1 dx\left(\frac{\sqrt{1-x}}{x^{3/2}}-\frac{1}{x^2}\arctan\sqrt{\frac{x}{1-x}}-\frac{1}{\sqrt{x(1-x)}}\right)=-\frac{3}{2}\frac{\pi\alpha^3\psi^2}{mM}. \tag{37}$$

To calculate the contribution due to $G^{(0)}$ we must account properly for the wave-function's short-distance behavior:

$$\mathscr{M}_\psi^r=-\frac{\alpha^3\psi^2}{\pi M}\int_{C_-}\frac{d\omega}{i\omega}\left\langle\left(\frac{4\pi\mathbf{p}'}{p'^2}\frac{2m+\omega+\boldsymbol{\alpha}\mathbf{p}'}{p'^2-\Omega^2}\frac{4\pi\boldsymbol{\alpha}_q}{q^2-\mathscr{K}^2}+\frac{4\pi\boldsymbol{\alpha}_{p'}}{p'^2-\mathscr{K}^2}\frac{\omega+\boldsymbol{\alpha}\mathbf{p}'}{p'^2-\Omega^2}\frac{4\pi\mathbf{q}}{q^2}\right)\frac{2m+\boldsymbol{\alpha}\mathbf{p}}{p^2}\frac{4\pi}{p^2}\right\rangle. \tag{38}$$

Averaging over the spin part of the wave function, we obtain

$$\mathscr{M}_\psi^r=-\frac{\alpha^3\psi^2}{\pi M}\int\frac{d\omega}{i}\left\langle\left(\frac{4m}{\omega}+1\right)\times\frac{4\pi}{p'^2(p'^2-\Omega^2)}\frac{4\pi}{q^2-\mathscr{K}^2}\frac{4\pi\mathbf{p}_q^2}{p^4}-\frac{4\pi}{(p'^2-\mathscr{K}^2)(p'^2-\Omega^2)}\frac{4\pi}{q^2}\frac{4\pi\mathbf{p}_{p'}^2}{p^4}\right\rangle. \tag{39}$$

Again, the six-dimensional integral over $\mathbf{p}$ and $\mathbf{p}'$ turns into a simple integral over r in the coordinate representation. This integral is

$$\mathscr{M}^r_\psi=\frac{2\alpha^3\psi^2}{M}\int\frac{d\omega}{i}\left\{\left(\frac{4m}{\omega}+1\right)\left[\frac{1}{\Omega^2}\ln\left(1+\frac{\Omega}{\mathscr{K}}\right)+\frac{1}{\mathscr{K}^2}\ln\left(1+\frac{\mathscr{K}}{\Omega}\right)-\frac{1}{\Omega\mathscr{K}}\right]+\frac{1}{2m\omega}\ln\frac{\mathscr{K}}{\Omega}\right\}. \quad (40)$$

Finally, the integration along the same contour as above gives the following expression for the nonvanishing terms in the limit $\varepsilon\to 0$:

$$\mathscr{M}^r_\psi=\frac{m^2\alpha^6}{M}\left(\frac{2}{\varepsilon}-\frac{32}{9\pi\sqrt{\varepsilon}}\int_0^\infty\frac{d\theta}{\sqrt{\cosh\theta}}+2\ln\frac{1}{\varepsilon}\right). \quad (41)$$

We see that, as expected, the logarithmic in ε term cancels the corresponding term in (31). The more singular in ε terms can only be the result of the regularization procedure applied to the positive-energy contribution (25). Since the latter is nonsingular at short distances, this procedure is actually unnecessary, i.e., it can produce only positive powers of $m\alpha/\lambda$. An explicit calculation can be found in Appendix A.

4.3. Total magnetic contribution

In the sum of all contributions due to a single magnetic exchange any dependence on the scale that separates the parameter ε cancels out, and we obtain

$$\mathscr{M}_{\alpha^6}+\left\langle\frac{p^2}{2M}\right\rangle_{\alpha^6}=\frac{m^2\alpha^6}{M}\left(-1+2\ln\frac{1}{\alpha}-1-\frac{3}{2}\right). \quad (42)$$

Here -1 on the right-hand side is due to the (long-range) effect of retardation [see Eq. (25) and Eq. (14) of Ref. 2], $2\ln(1/\alpha)-1$ comes from the whole range of scales from $m\alpha$ to m, and $-3/2$ is the short-range contribution.

5. SEAGULL CONTRIBUTION

5.1. Long distances

The best way to analyze the atomic scale contribution is to begin by taking the integral with respect to ω. It appears that in the order of interest only the positive-energy intermediate states should be considered:[2]

$$\mathscr{S}_+=\frac{\alpha^2}{2M}\int\frac{d^3\mathbf{k}}{(2\pi)^3}\times\left\langle\frac{4\pi}{k'^2}\frac{2\mathbf{p}'_{k'}+i[\sigma\mathbf{k}']}{2m}\frac{4\pi}{k^2}\frac{2\mathbf{p}_k+i[\sigma\mathbf{k}]}{2m}\right\rangle. \quad (43)$$

A simple power counting shows that only bilinear in $\mathbf{k}$ and $\mathbf{k}'$ term gives rise to the ultraviolet divergence. To regularize this divergence, we subtract from the divergent term the regulator contribution, which at large distances is

$$-\frac{\alpha^2}{4m^2M}\left\langle\frac{4\pi\mathbf{k}'}{k'^2+\lambda'^2}\frac{4\pi\mathbf{k}}{k^2+\lambda^2}\right\rangle, \quad (44)$$

while $m\alpha\ll\lambda$, $\lambda'\ll m$. In the coordinate representation the regularized version of (43) is

$$\mathscr{S}_+-\mathscr{S}^r_+=\frac{\alpha^2}{4m^2M}\left\langle 2\mathbf{p}\frac{1}{r^2}\mathbf{p}+\frac{1}{r^4}-\left(\nabla\frac{e^{-\lambda' r}}{r}\right)\left(\nabla\frac{e^{-\lambda r}}{r}\right)\right\rangle. \quad (45)$$

The average over the ground state is ($\varepsilon'=\lambda'/2m$):

$$\mathscr{S}_+-\mathscr{S}^r_-=\frac{m^2\alpha^6}{M}\left\{2\frac{\varepsilon'^2+\varepsilon'\varepsilon+\varepsilon^2}{\alpha(\varepsilon'+\varepsilon)}+1-2\ln\frac{\varepsilon'+\varepsilon}{\alpha}+\frac{2\varepsilon'\varepsilon}{(\varepsilon'+\varepsilon)^2}\right\}. \quad (46)$$

Here 1 appears due to the nonsingular operator $\mathbf{p}r^{-2}\mathbf{p}$. The first term in the curly brackets represents the regulator contribution to the previous order. In Appendix B, an appearance of this term as a short-range contribution to the $m\alpha^5/M$ order is shown explicitly. In what follows we calculate the subtracted term, whose nonrelativistic version (44) is ultraviolet divergent, in the framework of a relativistic approach.

5.2. Short distances

As in the case of the single magnetic exchange, only two first terms of the Green's function expansion in the Coulomb interaction contribute to the $m^2\alpha^6/M$ order. For the $G^{(1)}$'s contribution we have

$$\mathscr{S}^r_G=\frac{\alpha^3\psi^2}{2\pi M}\int_{C_-}\frac{d\omega}{i}\times\left\langle\frac{4\pi\boldsymbol{\alpha}_{p'}}{p'^2-\mathscr{K}'^2}\frac{\omega+\boldsymbol{\alpha}\mathbf{p}'}{p'^2-\Omega^2}\frac{4\pi}{q^2}\frac{\omega+\boldsymbol{\alpha}\mathbf{p}}{p^2-\Omega^2}\frac{4\pi\boldsymbol{\alpha}_p}{p^2-\mathscr{K}^2}\right\rangle. \quad (47)$$

Calculation along the same lines as in the case of $\mathscr{M}^r_G$ gives the result

$$\mathscr{S}^r_G=\frac{\pi\alpha^3\psi^2}{Mm}(4\ln 2-2), \quad (48)$$

which is nonsingular in the limit $\lambda,\lambda'\to 0$.

As for the contribution due to $G^{(0)}$, it can be extracted from

$$\frac{\alpha^3\psi^2}{2\pi M}\int_{C_-}\frac{d\omega}{i}\left\langle\frac{4\pi\boldsymbol{\alpha}_{p'}}{p'^2-\mathscr{K}'^2}\frac{\omega+\boldsymbol{\alpha}\mathbf{p}'}{p'^2-\Omega^2}\times\frac{4\pi\boldsymbol{\alpha}_q}{q^2-\mathscr{K}^2}\frac{2m+\boldsymbol{\alpha}\mathbf{p}}{(p^2+\gamma^2)^2}4\pi\right\rangle+(\lambda\leftrightarrow\lambda') \quad (49)$$

as a zeroth-order term of the Laurent series in $\gamma\equiv m\alpha$ (this series begins with an order $1/\gamma$ term, which describes the seagull contribution to the $m^2\alpha^5/M$ order at short distances, discussed in Appendix B). The average over the spin part of the wave function is

$$(2m\omega\boldsymbol{\alpha}_{p'}\boldsymbol{\alpha}_q+\boldsymbol{\alpha}_{p'}(\boldsymbol{\alpha}\mathbf{p}')\boldsymbol{\alpha}_q(\boldsymbol{\alpha}\mathbf{p}))\rangle=-(\omega^2+[p'^2-\Omega^2])\times\left(1+\frac{(\mathbf{p}'\mathbf{q})^2}{p'^2q^2}\right)+2\mathbf{p}'\mathbf{q}. \quad (50)$$

The term in the square brackets can be omitted. In fact, the corresponding part of (49) does not depend on m and hence (merely on dimensional grounds) contributes to the $m^2\alpha^5/M$ order only. The first term then gives the nonsingular contribution in the limit $\lambda,\lambda'\to 0$:

$$-\omega^2\left(1+\frac{(\mathbf{p}'\mathbf{q})^2}{p'^2q^2}\right)\to\frac{\pi\alpha^3\psi^2}{Mm}(1-4\ln 2). \tag{51}$$

Finally, analysis of the last term in (50) deserves more attention since here we have the infrared singularity. Integrated over the space variables, this term gives

$$2\mathbf{p}'\mathbf{q}\to\frac{2\alpha^3\psi^2}{Mm}\int_{C_-}\frac{d\omega}{i\omega}\{f(\Omega,\mathscr{K})-f(\mathscr{K}',\mathscr{K})\}, \tag{52}$$

where

$$f(x,y)=\ln\left(1+\frac{x}{y}\right)-\frac{xy}{(x+y)^2} \tag{53}$$

(recall that $\mathscr{K}'=\sqrt{\omega^2-\lambda'^2}$). For $\varepsilon\ll 1$ we obtain

$$\frac{2\alpha^3\psi^2}{Mm}\int_{C_-}\frac{d\omega}{i\omega}f(\Omega,\mathscr{K})=\frac{\pi\alpha^3\psi^2}{Mm}\left(-2\ln\frac{1}{\varepsilon}+4\ln 2-1\right). \tag{54}$$

Calculation of the integral with $f(\mathscr{K}',\mathscr{K})$ is slightly more cumbersome since it does not contain a small parameter. The contour C_- for this integral encompasses in the counterclockwise direction the cut connecting the points $-\lambda$ and $-\lambda'$. Continuous deformation of C_- leads to the equation

$$\int_{C_-}d\omega\ldots=\int_{C_+}d\omega\ldots-2\pi i\operatorname*{Res}_{\omega=0}\ldots-2\pi i\operatorname*{Res}_{\omega=\infty}\ldots, \tag{55}$$

where... represents $f(\mathscr{K}',\mathscr{K})/\omega$, and the contour C_+ goes in the clockwise direction around the cut that connects the points λ and λ'. Using the evident relations,

$$\int_{C_+}d\omega\ldots=-\int_{C_-}d\omega\ldots, \tag{56}$$

$$\operatorname*{Res}_{\omega=0}\frac{1}{\omega}f(\mathscr{K}',\mathscr{K})=f(\lambda',\lambda), \tag{57}$$

$$\operatorname*{Res}_{\omega=\infty}\frac{1}{\omega}f(\mathscr{K}',\mathscr{K})=-\ln 2+\frac{1}{4}, \tag{58}$$

we obtain

$$\frac{2\alpha^3\psi^2}{Mm}\int_{C_-}\frac{d\omega}{i\omega}f(\mathscr{K}',\mathscr{K})=\frac{\pi\alpha^3\psi^2}{Mm}\left(2\ln\frac{2\varepsilon}{\varepsilon+\varepsilon'}+\frac{2\varepsilon\varepsilon'}{(\varepsilon+\varepsilon')^2}-\frac{1}{2}\right). \tag{59}$$

5.3. Total seagull contribution

As can be seen from (46), (48), (51), (54), and (59), the total seagull contribution to the $m^2\alpha^6/M$ order does not depend on the scale that separates the parameters λ and λ'. The contribution is

$$\mathscr{S}_{\alpha^6}=\frac{m^2\alpha^6}{M}\left(1-2\ln\frac{2}{\alpha}+\frac{1}{2}+4\ln 2-2\right), \tag{60}$$

where 1 is the long-range contribution, $4\ln 2-2$ is the short-range contribution, and the remaining terms obtain their values on the whole range of scales from $m\alpha$ to m.

6. CONCLUSIONS

In complete agreement with the result of Ref. 13, the total correction of the $m^2\alpha^6/M$ order does not contain $\ln\alpha$. It consists of two terms,

$$\Delta E_{\rm rec}=\left.\frac{m^2-E^2}{2M}\right|_{\alpha^6}+\frac{m^2\alpha^6}{Mn^3}(2\ln 2-3). \tag{61}$$

The former term is completely determined by the atomic scale and depends nontrivially on the principal quantum number n,

$$\left.\frac{m^2-E^2}{2M}\right|_{\alpha^6}=\frac{m^2\alpha^6}{2Mn^3}\left(\frac{1}{4}+\frac{3}{4n}-\frac{2}{n^2}+\frac{1}{n^3}\right). \tag{62}$$

As for the latter term, our calculations show that it has its origin at the scale of the order of m.

The correction (61) shifts the hydrogen ground state by -16.4 kHz and the $2S$ state by -1.9 kHz. These figures are comparable[14] or even exceed the uncertainties of the recent Lamb shift measurements.[15]

The result (61) differs from those obtained in Refs. 1 and 3 and in Ref. 2. Let us first discuss the origin of the difference in the latter case. In Ref. 2, it was erroneously assumed that the cancellation of singular operators at the atomic scale does not leave a nonvanishing remainder. The present calculation shows that because of the difference in the particular features of a cutoff procedure used to regularize the average values of different singular operators, some finite contributions survive the cancellation process.

Unfortunately, the same error was repeated in Ref. 3. The long-range contribution was found there in the framework of some particular regularization scheme. It was then added to the short-range contribution calculated in Refs. 1 and 2 by completely different regularization procedures. The regularization dependence of the results obtained in Ref. 3 can be seen, for example, in Eq. (29) of Ref. 3, where the integration over k', which is limited above by the parameter σ', gives rise to a finite (depending on σ'/σ) contribution to the result. This contribution was erroneously omitted from Eq. (29) of Ref. 3.

The error made in Ref. 1 is a computational error, which is caused by inaccurate treatment of the frequency dependence in the integral (42) of Ref. 1 [ironically, because of a typographical error, only the important factors, $(\omega^2-k_1^2)^{-1}$ and $(\omega^2-k_2^2)^{-1}$, are skipped in Eq. (42) of Ref. 1]. In what follows we rederive our result employing the regularization scheme used in Ref. 1.

First of all, the result for the long-range contribution (46) of Ref. 1 (''the third term'') is consistent with the result of our work [1 in Eq. (46)].

As for the remaining contributions, let us begin with one general comment. In their analysis of the integral (42), the authors of Ref. 1 use the symmetrization in ω, since, as they wrote, ''generally there are three regions of photon energy, $\omega\sim\alpha^2$, $\omega\sim\alpha$, and $\omega\sim 1$, that give a contribution and the

middle region is almost eliminated by the symmetrization''. In order to avoid the discussion whether the middle region is eliminated or not, we will recalculate the contributions of the first and the second terms in Eq. (43) of Ref. 1 without the symmetrization in ω. Since the symmetrization procedure is no more than a technical trick, the result of a calculation should not depend on whether this procedure is applied or not.

To determine the high-energy part of the first and second term contribution, we set $\varepsilon'=\varepsilon=0$ in (49) and cut off the low-energy end $|\omega|<m\epsilon$ from the contour C_-. The result for the short-range (high-energy) contribution to the integral (42) of Ref. 1 can then be obtained:

$$\Delta E=\frac{m^2\alpha^6}{M}\,2\ln\frac{\epsilon}{2}. \tag{63}$$

The sum of the order $m^2\alpha^6/M$ contributions to Eqs. (51), (54), and (57) of Ref. 1 is smaller by a factor of 2. An extra factor 1/2 emerges there due to the symmetrization in ω, since the contribution of the contour C_+, which wraps the half-axis $(m\epsilon,\infty)$, vanishes.

Let us now consider the low energies. Only the second term in Eq. (43) of Ref. 1 contributes there. According to Eqs. (42) and (43) of Ref. 1, this contribution (with the typos corrected) is

$$\Delta E=\frac{\alpha^2}{Mm}\int_{C_L}\frac{d\omega}{2\pi i}\int\frac{d^3\mathbf{k}_1}{(2\pi)^3}\int\frac{d^3\mathbf{k}_2}{(2\pi)^3}\int\frac{d^3\mathbf{p}}{(2\pi)^3}\times\psi(\mathbf{p}+\mathbf{k}_1)\,\frac{4\pi\mathbf{k}_1}{k_1^2-\omega^2}\,\frac{1}{2m\omega-p^2}\,\frac{4\pi\mathbf{k}_2}{k_2^2-\omega^2}\,\psi(\mathbf{p}+\mathbf{k}_2). \tag{64}$$

Here the contour C_L goes from $-m\epsilon$ to 0 below the real axis and then from 0 to $m\epsilon$ above it. Recall now that the high-energy contribution (63) is calculated on the assumption that $\epsilon\gg\alpha$. This means that in (64) we can ignore p^2, which is of the order of $(m\alpha)^2$, in comparison with $2m\omega$, which is shown below to be of the order of $m^2\alpha$. We can then easily come to the coordinate representation and obtain

$$\Delta E=\frac{\alpha^2}{2Mm^2}\int_{C_L}\frac{d\omega}{2\pi i}\,\frac{1}{\omega-0}\left\langle\left(\nabla\,\frac{e^{i|\omega|r}}{r}\right)^2-\frac{1}{r^4}\right\rangle. \tag{65}$$

Since the integration contour does not wrap the zero point, we can safely add the operator $-1/r^4$, which is annihilated by the ω integration. The result of taking the average over the ground state is

$$\Delta E=-\frac{2m^2\alpha^6}{M}\int_{C_L}\frac{d\omega}{2\pi i}\,\frac{1}{\omega-0}\left(2\ln\left(1-\frac{i|\omega|}{m\alpha}\right)+\frac{2i|\omega|}{m\alpha}\,\frac{1}{1-i|\omega|/m\alpha}+\frac{3}{2}\left(\frac{\omega}{m\alpha}\right)^2\frac{1}{1-i|\omega|/m\alpha}\right). \tag{66}$$

Here we see that the natural scale for ω is in fact $m\alpha$. Since $|\omega|$ is positive on the lower half of C_L, the integral given above in dimensionless units reads

$$\Delta E=-\frac{m^2\alpha^6}{\pi M}\int_0^{\epsilon/\alpha}dx\left(\frac{4}{x}\arctan x-\frac{4}{1+x^2}-3\,\frac{x^2}{1+x^2}\right). \tag{67}$$

The result of integration,

$$\Delta E=\frac{m^2\alpha^6}{M}\left(3\,\frac{\epsilon}{\pi\alpha}-2\ln\frac{\epsilon}{\alpha}+\frac{1}{2}\right), \tag{68}$$

which is added to all the other seagull contributions, gives for the order $m^2\alpha^6/M$ seagull correction:

$$\mathscr{S}_{\alpha^6}=\frac{m^2\alpha^6}{M}\left(-2\ln\frac{1}{\alpha}+2\ln 2-\frac{1}{2}\right), \tag{69}$$

in complete agreement with our result (60).

I wish to thank M. Eides, H. Grotch, and A. Milstein for stimulating discussions. I also thank S. Karshenboim, who informed me about recent experimental work. This research was financed in part by the Russian Fund for Fundamental Research (Grant No. 97-02-18450) and by the Universities of Russia Program (Grant No. 95-0-5.5-130).

APPENDIX A

The extra terms in (41) should be canceled by the regulator counterpart of (13), which differs from (13) by $\sqrt{k^2+\lambda^2}$ used instead of k. As in the main text, we approximate the sum over the positive-energy intermediate states by the nonrelativistic Green's function and the matrix element of $\boldsymbol{\alpha}$ by $\mathbf{p}/m$. In this approximation, the regulator contribution is

$$\mathscr{M}^r_+=\frac{\alpha}{Mm}\int\frac{d^3\mathbf{k}}{(2\pi)^3}\left\langle\mathbf{p}\mathscr{G}(\mathbf{r}',\mathbf{r}|E-\sqrt{k^2+\lambda^2})\,\frac{4\pi e^{i\mathbf{kr}}}{\sqrt{k^2+\lambda^2}}\,\mathbf{p}_k\right\rangle. \tag{A1}$$

After the transformations the regulator version of the expression (18) is

$$\mathscr{M}^r_+=\frac{2^7 3m^5\alpha^6}{M\pi}\int_0^\infty\frac{dk\;k^2}{\kappa^5\omega}\int_{-1}^1dx(1-x^2)\times\int_C dt\,\frac{t^{1-m\alpha/\kappa}}{(a-bt)^4}\,\frac{1}{1-\exp(2\pi i m\alpha/\kappa)}, \tag{A2}$$

where $\kappa=\sqrt{2m(\omega-E)}$, $\omega=\sqrt{k^2+\lambda^2}$, the contour C, and the functions a and b are defined in the text. Only singular terms in the expansion (22) operate at ranges of the order of λ^{-1}. For those terms the integrals over k and x become elementary and give

$$\mathscr{M}^r_+=\frac{m^2\alpha^6}{M}\left\{-\frac{2}{\varepsilon^2}\left(\ln\frac{\varepsilon}{\alpha}-1\right)+\frac{2}{\varepsilon}-\frac{32}{9\pi\sqrt{\varepsilon}}\times\int_0^\infty\frac{d\theta}{\sqrt{\cosh\,\theta}}\right\}. \tag{A3}$$

The second and the third terms therefore coincide with the corresponding terms in (41). The new singularity $\propto\varepsilon^{-2}$ is

the regulator contribution to the instantaneous part of the magnetic exchange (10):

$$-\frac{1}{2M}\langle\{\mathbf{p},\mathbf{D}^r(0,\mathbf{r})\}\rangle$$
$$\approx -4\pi\alpha\left\langle\frac{p_i'+p_i}{2m}\frac{p_j'+p_j}{2M}\frac{\delta_{ij}-q_iq_j/q^2}{q^2+\lambda^2}\right\rangle$$
$$=-2\frac{m^2\alpha^6}{M\varepsilon^2}\left(\ln\frac{\varepsilon}{\alpha}-1\right). \tag{A4}$$

The leading contribution to (49) is

$$\mathscr{S}^r=\frac{8m^3\alpha^5}{M}\int_{C_-}\frac{d\omega}{2\pi i}\int\frac{d^3\mathbf{p}}{(2\pi)^3}\frac{4\pi}{p^2-\mathscr{K}'^2}$$
$$\times\frac{\omega}{p^2-\Omega^2}\frac{4\pi}{p^2-\mathscr{K}^2}. \tag{B1}$$

After the integration with respect to **p** it becomes

$$\mathscr{S}^r=\frac{m\alpha^5}{M\pi}\frac{1}{\varepsilon^2-\varepsilon'^2}\int_{C_-}d\omega\;\omega\left(\frac{1}{\Omega+\mathscr{K}'}-\frac{1}{\Omega+\mathscr{K}}\right). \tag{B2}$$

Up to terms of the first order in ε, ε' we obtain

$$\mathscr{S}^r=\frac{m^2\alpha^5}{M\pi}\left(3-2\frac{\varepsilon^2\ln(2/\varepsilon)-\varepsilon'^2\ln(2/\varepsilon')}{\varepsilon^2-\varepsilon'^2}\right.$$
$$\left.-2\pi\frac{\varepsilon'^2+\varepsilon'\varepsilon+\varepsilon^2}{\varepsilon'+\varepsilon}\right). \tag{B3}$$

The last term compensates for the leading contribution to (46)

[1] K. Pachucki and H. Grotch, Phys. Rev. A **51**, 1854 (1995).
[2] A. S. Yelkhovsky, Zh. Éksp. Teor. Fiz. **110**, 431 (1996) [JETP **83**, 230 (1996)].
[3] M. Eides and H. Grotch, Phys. Rev. A **55**, 3351 (1997).
[4] E. A. Golosov, I. B. Khriplovich, A. I. Milstein, and A. S. Yelkhovsky, Zh. Éksp. Teor. Fiz. **107**, 393 (1995) [JETP **80**, 208 (1995)].
[5] M. A. Braun, Zh. Éksp. Teor. Fiz. **64**, 413 (1973) [Sov. Phys. JETP **37**, 211 (1973)].
[6] V. M. Shabaev, Theor. Math. Phys. **63**, 588 (1985); in *The First Soviet-British Symposium on Spectroscopy of Multicharged Ions*, Troitsk, 1986, Spectroscopic Council of the USSR Academy of Sciences (1986), p. 238.
[7] A. S. Yelkhovsky, Preprint Budker INP 94-27, E-prints archive hep-th/9403095.
[8] V. M. Shabaev, Phys. Rev. A **57**, 59 (1998).
[9] J. H. Epstein and S. T. Epstein, Am. J. Phys. **30**, 266 (1962).
[10] A. I. Mil'shtein and V. M. Strakhovenko, Phys. Lett. A **90**, 447 (1982).
[11] Ya. I. Granovsky, Zh. Éksp. Teor. Fiz. **56**, 605 (1969) [Sov. Phys. JETP **29**, 333 (1969)].
[12] E. E. Salpeter, Phys. Rev. **87**, 328 (1952).
[13] R. N. Fell, I. B. Khriplovich, A. I. Milstein, and A. S. Yelkhovsky, Phys. Lett. A **181**, 172 (1993).
[14] M. Weitz, A. Huber, F. Schmidt-Kaler, D. Leibfried, and T. W. Hänsch, Phys. Rev. Lett. **72**, 328 (1994); D. J. Berkeland, E. A. Hinds, and M. G. Boshier, Phys. Rev. Lett. **75**, 2470 (1995); S. Bourzeix, B. de Beauvoir, F. Nez *et al.*, Phys. Rev. Lett. **76**, 384 (1996); B. de Beauvoir, F. Nez, L. Julien *et al.*, Phys. Rev. Lett. **78**, 440 (1997).
[15] T. Udem, A. Huber, B. Gross, J. Reichert, M. Prevedelli, M. Weitz, and T. W. Hansch, Phys. Rev. Lett. **79**, 2646 (1997).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Plasma waves in a nonideal plasma

A. A. Valuev, A. S. Kaklyugin, and G. É. Norman*)

Moscow Physical Society, 121019 Moscow, Russia
(Submitted 10 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 880–896 (March 1998)

This paper shows how the concepts commonly used for a Debye plasma—Landau damping, collisional damping, short-range and long-range collisions, and plasma waves—must be revised to describe a nonideal electron–ion plasma. The degrees of freedom of a nonideal plasma are divided into collective and individual. The increase and saturation of the fraction of collective degrees of freedom as the coupling constant increases is discussed. The Tatarskiĭ approach for a system of coupled oscillators makes it possible to model the collective degrees of freedom of a nonideal plasma by a set of Langevin oscillators in a thermostat. The correlation energy and the energy of the plasma waves are found. The concepts developed here made it possible to determine the dispersion of the plasma waves and their damping. The effect of damping on the discrepancy between the position of the maximum of the dynamic structure factor and the real part of the solution of the dispersion equation is considered. The effective collision frequency of the individual degrees of freedom (the electrons) is estimated, taking into account both short-range pairwise scattering and scattering at plasma waves. [S1063-7761(98)00903-2]

1. INTRODUCTION

Plasma (Langmuir) waves are a specific characteristic of a gaseous (Debye) plasma (see, for example, Refs. 1–5). The formulas obtained for a gaseous plasma restrict plasma waves to wave numbers $k/\kappa \ll 1$ and coupling constant $\gamma=\beta e^2 n^{1/3} \ll 1$,[4] where $\kappa=a_e^{-1}$ is the inverse electron Debye radius, $a_e=[T/(4\pi e^2 n_e)]^{1/2}$, n_e and n are the densities of the electrons and the positive charges, and $T=\beta^{-1}$ is the temperature.

However, plasma waves in a nonideal plasma have been observed in early experiments[6–9] and with simulation by the molecular-dynamics method.[10–13] In this connection, the existence of plasma waves in a nonideal plasma has been considered proven in a number of papers, and has been used to solve various problems of the theory of such a plasma.[14–20] At the same time, no one has succeeded in making a satisfactory transition from a Debye plasma[1–5] to a nonideal plasma.[6–20] Thus, an entire region of the theory of a nonideal plasma relating to plasma waves and corresponding to the concepts developed in Refs. 1–5 has been lacking.

This paper attempts to develop a theory of a nonideal plasma that follows the concepts of Refs. 1–5 and the method of collective variables.[21,22] In doing so, we start from the restriction of the possible wave vectors of the longitudinal plasma modes, $kn^{-1/3}<\text{const}$, which was covered in Refs. 23–25, and from the associated limitation of the saturation of the collective degrees of freedom.[24–27]

The expressions for a Debye plasma are used in this paper with values of the coefficients from Ref. 1. However, because of the approximate character of the existing models of a Debye plasma,[1–5] the choice of numerical values is ambiguous. This is even more true of a nonideal plasma. We have directed most of our attention to qualitative dependences.

Section 2 briefly presents the necessary information from the theory of a Debye plasma. Section 3 reviews the analogous concepts—Landau damping, collisional damping, the region of existence of plasma waves, and the number of collective degrees of freedom—for a nonideal plasma by means of qualitative estimates.

Section 4 separates the degrees of freedom of a nonideal plasma into collective and individual. It is shown that the equations of motion for the collective variables reduce to equations of the Langevin type, so that it is possible to introduce dispersion characteristics. Section 5 discusses the energy of the fluctuations of the longitudinal electric field. It is shown how the theory developed for an ideal plasma can be extended to a nonideal plasma. Both the correlation energy and the energy of the plasma waves are found. The distribution of these energies over wave numbers and frequencies is discussed.

The concepts that have been developed are used in Sec. 6 to determine the dispersion characteristics. The results are compared with experimental (actual and computer-simulated) observations of plasma waves in a nonideal plasma. The frequencies of the interactions of the electron (the individual degrees of freedom) with the charges and the plasma waves are considered in Sec. 7. The results are compared with the data of a molecular-dynamics experiment.

A nonideal plasma is a disordered system with strong interaction. It is assumed that, for this reason, a nonideal plasma can be studied theoretically either by extrapolating and interpolating rigorous expansions for a weakly nonideal plasma and by approaches developed in the theory of metals or by computer simulation.[28,29] In the Conclusion, it is dis-

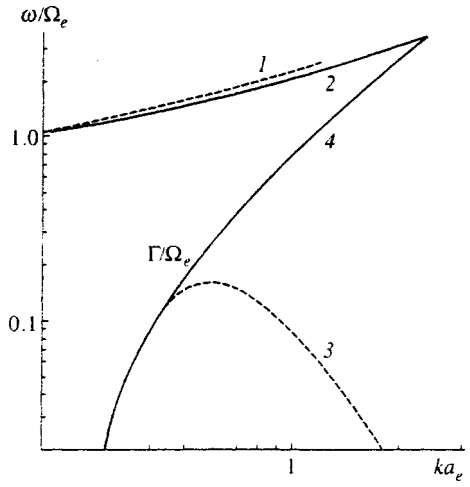

FIG. 1. Dispersion curves for an ideal plasma. The real part (*1, 2*) and the imaginary part (*3, 4*) of the frequency of the plasma waves, computed from Eqs. (1) and (2) and according to Refs. 30 and 31, respectively.

cussed why it is nevertheless promising to construct such a theory *ab initio*.

2. THE DEBYE PLASMA

In a low density plasma, longitudinal plasma waves can propagate with a dispersion law

$$\omega^2(k)=\Omega_e^2(1+3a_e^2k^2) \tag{1}$$

and with a collisionless Landau damping rate of[1]

$$\frac{\Gamma_L(k)}{\Omega_e}=0.14\left(\frac{\kappa}{k}\right)^3\exp\left[-\frac{1}{2}\left(\frac{\kappa}{k}\right)^2\right], \tag{2}$$

which is independent of γ [here $\Omega_e=(4\pi e^2n_e/m)^{1/2}$ is the plasma frequency]. The damping rate Γ_L is exponentially small for $k/\kappa\ll 1$ and increases as k/κ increases (Fig. 1). However, Eq. (2) is valid only in the limit $k/\kappa\ll 1$ and becomes inapplicable in the region $k/\kappa\approx 1$, i.e., even before Γ_L becomes a quantity of the same order of magnitude as the frequency, and the concept of propagating plasma waves becomes meaningless.[1] The $\omega(k)$ and $\Gamma_L(k)$ dependences can be extended into the region of large k by solving a complex functional equation, as described in Bekefi's monograph[30] with a citation to Ref. 31. The result shown in Fig. 1 makes it possible to find the value of $k_0=\alpha\kappa$ ($\alpha=2.34$) for which $\Gamma_L(k_0)=\omega(k_0)$ [in this case, $\omega(k)/\Omega_e=3.5$]. The quantitative results of Ref. 31 cannot be considered very reliable in the region $\Gamma_L/\omega\approx 1$, since the expressions used in this case formally go beyond the limits within which they are applicable. However, this value of α agrees with Ref. 21, where it is assumed that the motion of the particles in a plasma is moderately collective for $k\approx 2\kappa$, while the degrees of freedom with $k\approx\kappa$ can already be considered completely collective. These estimates refine the quantitative concepts used in Ref. 1.

The collisional damping Γ_c of the plasma waves (plasmons) is expressed in Ref. 1 in terms of the frequency ν_{ei} of the electron–ion collisions in the Lorentzian model:

$$\Gamma_c=\nu_{ei}/3\sqrt{2\pi}. \tag{3}$$

In a Debye plasma, ν_{ei} is mainly determined by long-range collective interactions with scattering at small angles and is described by the Landau collision integral

$$\nu_{ei}=\langle n_e v\,\sigma_t(v)\rangle, \tag{4}$$

where v is the electron velocity, the brackets denote averaging over velocities, σ_t is the electron momentum-transfer cross section

$$\sigma_t\propto\int_0^{\pi/2}(1-\cos\chi)\,\frac{\cos(\chi/2)}{\sin^3(\chi/2)}\,d\chi. \tag{5}$$

and χ is the scattering angle. Since $\chi\ll 1$, the expression under the integral in Eq. (5) approximately equals $d\chi/\chi$. Then

$$\Gamma_c/\Omega_e=\ln\Lambda\,\gamma^{3/2}/3, \tag{6}$$

where $\Lambda=\rho_{\max}/\rho_{\min}=(1/4\pi)^{1/2}\gamma^{3/2}$ (the total Debye radius $a=a_e/2^{1/2}$ is usually taken as the maximum impact parameter $\rho_{\max}$, while $\rho_{\min}=r_L=e^2\beta$ is taken as the minimum).

Equation (6) is valid for $\gamma\ll 1$ ($\ln\Lambda\gg 1$) and becomes inapplicable as γ increases before the value $\Gamma_c/\Omega_e\approx 1$ is reached. To estimate Γ_c for large γ, the value of $\ln\Lambda$ was usually fixed when it attained a value of ≈ 3 (Ref. 32) (experimental data on electrical conductivity for $\gamma>0.1$ indicates that this is possible). Then we have $\Gamma_c/\Omega_e=1$ when $\gamma_0\approx 1$.

According to the above estimates, the region in which plasma waves exist is thus limited by the inequalities

$$k<k_0\approx\alpha\kappa, \tag{7a}$$

$$\gamma<\gamma_0\approx 1. \tag{7b}$$

The region of existence of plasma waves is sometimes limited even more strongly by the inequality

$$\Omega_e/\nu_{ei}>1.$$

It is this inequality that Iakubov and Khrapak,[33] for example, refer to when they assert that plasma waves cease to exist in a nonideal plasma. This explains why plasma waves are not mentioned in the monographs of Ebeling *et al.*[28] and Fortov and Yakubov.[29]

Another important characteristic for our treatment is the number s of collective degrees of freedom. The value of s can be estimated as the ratio of the volume of a sphere, $4\pi k_0^3/3$, to $(2\pi)^3/V$, which gives the density of possible wave vectors in volume V. In a Debye plasma,

$$\frac{s}{3N_e}=\frac{1}{9\pi^{1/2}}\,\alpha^3\gamma^{3/2}\ll 1, \tag{8}$$

where N_e is the number of electrons in volume V. It can be seen that almost all the degrees of freedom remain individual.

3. THE DAMPING RATE

Estimates of Γ_L and Γ_c valid for $\gamma\ll 1$ require substantial revision for the region $\gamma\gtrsim 1$. The key here is to naturally limit the wavelengths of the plasma waves to about the mean

distance between the charged particles;[23] in terms of wave numbers, it corresponds to a bound of the Debye wave number q_0:[34]

$$k<q_0=(6\pi^2)^{1/3}n_e^{1/3}. \tag{9}$$

In a Debye plasma, $q_0\gg\kappa$, and the limitation of inequality (7a) is stronger than that of inequality (9). As γ increases, the inequality reverses: $q_0<\kappa$.

Landau damping is determined by electrons moving with a velocity equal to the phase velocity v_f of the wave:[1]

$$v_f=\frac{\omega(\mathbf{k})}{k}>\begin{cases}\Omega_e/k_0\simeq v_T, & \gamma\ll 1,\\ \Omega_e/q_0\simeq v_T\kappa n^{-1/3}>v_T, & \gamma\geqslant 1.\end{cases} \tag{10}$$

In a Debye plasma, such electrons are thermal. In a non-Debye plasma, as γ increases, a smaller and smaller number of electrons, corresponding to the far tail of the Maxwell distribution, contribute to the Landau damping, and the maximum value of $\Gamma_L(q_0)$ decreases with increasing γ. Thus, as γ increases, the region of k values in which strong Landau damping could occur shifts toward larger k and is displaced beyond the limit $k=q_0$. In other words, the role of Landau damping decreases as γ increases in the nonideal region.

The estimates made in Sec. 2 for the collisional damping rate become meaningless in a nonideal plasma, since, as γ increases, a larger and larger role begins to be played by strong interactions at distances of the order of and less than $n^{-1/3}$. Three regions of γ can be distinguished, in each of which interactions of a different type predominate:

(a) The region $\gamma\ll 1$ ($a\gg n^{-1/3}$). Weak collective interactions at distances $\approx a$ are the determining factors.

(b) The region $\gamma\sim 0.1$ ($a\sim n^{-1/3}$). The Debye character of the screening is maintained, but strong interactions at distances of $\approx n^{-1/3}$ become the main interactions, while involving multiple particles.

(c) The region $\gamma>1$ ($a<n^{-1/3}$). The Debye radius loses the meaning of a screening radius, and screening now takes place at distances of $\approx n^{-1/3}$ (this is also suggested by the results of numerical simulation[35, 1]). The character of the interaction substantially changes in this case:[36] Because the Coulomb wells of the close-lying ions overlap, the electron–ion interaction potential effectively becomes short-range. The main contribution to ν_{ei} now comes from strong pairwise interactions at distances of $<n^{-1/3}$ with scattering at large angles.

To estimate the collisional damping rate for any γ, it is convenient to introduce damping rate Γ_1 and Γ_2, caused by weak and pairwise strong collisions, respectively. The first of them is described, as before, by the Landau collision integral and the second by the Boltzmann collision integral, both being expressed in terms of the Coulomb scattering integral in Eq. (5). Going in Eq. (5) from angles to impact parameters ρ, after integrating in the first case from $n_e^{-1/3}/2$ to a, and in the second case from zero to $n_e^{-1/3}/2$, we get for the region $\gamma\ll 1$

$$\frac{\Gamma_1}{\Omega_e}=\frac{\gamma^{3/2}}{6}\ln\frac{\Lambda_1^2+1}{\Lambda_2^2+1} \tag{11}$$

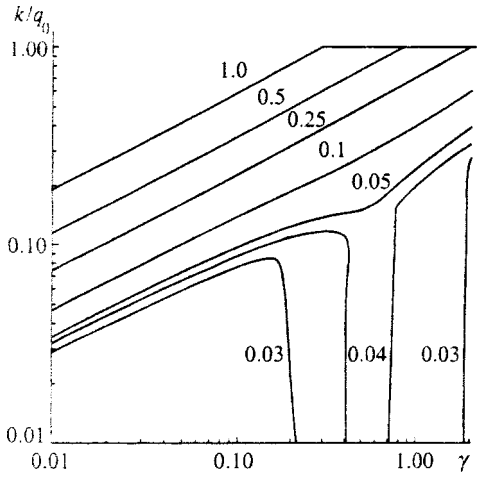

FIG. 2. Region of existence of plasma waves in a nonlinear plasma. The horizontal line is $k=q_0$, and the curves are the contours $(\Gamma_c+\Gamma_L)/\omega(k)=\text{const}$. The numbers near the curves indicate the values of the constant.

and for the region $\gamma>1$

$$\frac{\Gamma_2}{\Omega_e}=\frac{\gamma^{3/2}}{6}\ln(\Lambda_2^2+1), \tag{12}$$

where $\Lambda_1=a/r_L$, $\Lambda_2=n_e^{-1/3}/2r_L$. There is a certain indeterminacy (from 0.5 to 1) in the choice of the coefficient in Λ_2. After averaging in Eq. (4), the value of mv^2 under the logarithm sign is replaced by T (this estimate also introduces indeterminacy). In the limit $\gamma\ll 1$ we have $\Gamma_1\gg\Gamma_2$ and Γ_1 increases as $\gamma^{3/2}$. For $\gamma\gg 1$ the damping rate Γ_2 decreases as $\gamma^{-1/2}$. Consequently, a maximum must exist in the intermediate region $\gamma\sim 0.1$. Its value can be estimated by summing Γ_1 and Γ_2:

$$\Gamma_c=\Gamma_1+\Gamma_2. \tag{13}$$

Here $\Gamma_1=0$ holds for $\Lambda_1<\Lambda_2$, since the concept of weak collisions becomes meaningless when $a<n^{-1/3}$. Although neither Eq. (11) nor (12) can generally be used in the intermediate region, such an estimate can be supported by the fact that all such collisions result in deviations by small angles (the extreme case is the flight of an electron precisely between two ions). Small deviations make a small contribution to Γ_c. This is valid, of course, if Eq. (13) actually gives small values of Γ_c. According to the estimate of Eq. (13), taking into account the noted indeterminacies, the maximum of Γ_c/Ω_e is $0.08-0.2\ll 1$ and is located at $\gamma=0.8-1.6$.

It follows from these estimates that plasma waves remain weakly damped even when $\gamma\approx 1$ and $k\approx q_0$. The region of their existence is substantially wider than that given by inequalities (7); it is shown in Fig. 2, where the contours of the total damping rate are also plotted. It can be seen that either Landau damping or collisional damping dominates, depending on the values of k and γ.

In considering the number of collective degrees of freedom in a nonideal plasma, it should be borne in mind that we have $\alpha\kappa=q_0$ when $\gamma=\gamma_L$. In Fig. 2, $\gamma_L=0.25$. For $\gamma>\gamma_L$ inequality (8) becomes invalid. It follows from the definition of the Debye wave number q_0 that

$$s/3n_e = 1/3 \quad \text{when} \quad \gamma > \gamma_L . \tag{14}$$

The value 1/3 corresponds to the fact that only longitudinal oscillations are considered.

Thus, in a nonideal plasma, the number of collective degrees of freedom has increased by comparison with a Debye plasma, while Landau damping has ceased to limit them.

4. COLLECTIVE DESCRIPTION OF A NONIDEAL PLASMA

Since there are numerous well-defined collective modes in a nonideal plasma, it is natural to attempt to use for such a plasma the method of collective variables developed by Bohm[21] to describe a gas plasma. This method was used to obtain the dispersion behavior of Eqs. (1) and (2), which follows from the kinetic equations of Refs. 1 and 30. The collective variables were defined in the form

$$\rho_k(t) = \sum_j^N \exp[-ikx_j(t)].$$

The case of a nonideal plasma is far more complex. We are dealing with a system of nonlinearly interacting oscillators[21] in which, unlike the case of an ideal plasma, this interaction cannot be neglected.

In order to obtain equations of motion for collective excitations in a nonideal plasma that can be used for subsequent qualitative analysis, it is necessary to modify the method of collective variables. In order to do this in the present paper, we immediately use the heuristic estimates from Sec. 3; i.e., we assume that the damping of the plasma waves is small. Then, by analogy with Ref. 21, the Hamiltonian H of the system can be written as

$$H = H_{\rm os}(\rho_k) + H_{\rm in}(q,p) + \Delta H_{kk'} + \Delta H_{\rm os-in}, \tag{15}$$

where $H_{\rm os}(\rho_k)$ is the Hamiltonian corresponding to the collective degrees of freedom, $H_{\rm in}(q,p)$ is the Hamiltonian of the individual degrees of freedom, q and p are the coordinates and momenta of the separate particles, and $\Delta H_{kk'}$ and $\Delta H_{\rm os-in}$ are corrections that allow for the interactions of the collective modes with each other and with the individual degrees of freedom. Strictly speaking, in a nonlinear plasma, it is impossible to separate the collective and individual variables in the last two terms. Writing them thus in Eq. (15) is justified only because of their assumed smallness.

We shall not give the explicit form for the Hamiltonian in Eq. (15); this is rather complex, and there is no need for it in the subsequent estimates in this paper. Therefore, we immediately write the equations of motion for the collective variables, following from Eq. (15):

$$\ddot{\rho}_k = -\omega_k^2 \rho_k + F_{kk'} + F_{\rm os-in}, \tag{16}$$

where the last two terms are forces corresponding to the corrections in the Hamiltonian of Eq. (15). The system of Eqs. (16) recalls the system for weakly coupled oscillators considered in Ref. 37. They differ only in the presence of the term $F_{\rm os-in}$, describing the interaction of an oscillator with the mean of the individual particles, which can be regarded as interaction with a thermostat. Reference 37 investigated the relaxation of a Hamiltonian system of coupled oscillators at times small by comparison with the Poincaré recurrence time. In this case, equations like Eq. (16) were represented in the form

$$\ddot{\rho}_k = -\omega_k^2 \rho_k + \Gamma_{\rm eff}\, \dot{\rho}_k + y, \tag{17}$$

where $\Gamma_{\rm eff}$ is the effective coefficient of friction, while $y(t)$ is a δ-correlated random force, $\langle y(t)y(t')\rangle = 2D_\rho \delta(t-t')$, where D_ρ is the diffusion coefficient in ρ space. In our case, only the presence of the thermostat makes it easy to go from Eq. (16) to Eq. (17).

Thus, the equation of motion of the collective variables can be reduced, at least in principle, to simpler equations of Langevin type. The simplification, in particular, consists of the fact that, unlike Eq. (16), Eq. (17) depends only on its wave vector. The damping due to the interaction of different collective modes, following from $F_{kk'}$, and the damping caused by interaction with the individual degrees of freedom and following from $F_{\rm os-in}$ are combined into an effective coefficient of friction.

Note that Eq. (17) differs from the equation for damped oscillations by the presence of random excitation $y(t)$. Its appearance could have been foreseen initially, since Eq. (17) models a subsystem of the collective degrees of freedom of the entire system of Eqs. (15) at constant temperature.

As in the ordinary method of collective variables, isolating the collective degrees of freedom limits the motion of the individual quasiparticles to a neighborhood ΔE of a certain hypersurface in phase space with dimension $3N_e - s$. However, unlike the case of a Debye plasma, this neighborhood is apparently rather large (because of the fluctuation–dissociation theorem of Ref. 37, $\Delta E \propto D_\rho = nk^2 \Gamma_{\rm eff} v_T$), which expresses the far stronger coupling between the collective and individual degrees of freedom.

Equation (17) can be analyzed similarly to the usual equation for a stochastic oscillator.[38] In this case, the concept of the normal frequency of an oscillator becomes meaningless, and its role is taken over by the frequency ω_1 at which the maximum response to the random force $y(t)$ of Eq. (17) occurs. For this frequency, the expression $\omega_1^2 = \omega_k^2 - \Gamma_{\rm eff}^2$, which is well known in mechanics, is valid. The quantity $\Gamma_{\rm eff}$, unlike the case of Landau damping, has a dissipative character. The quantities ω_k and $\Gamma_{\rm eff}$ must of course be determined by successively carrying out the transformations given in Eqs. (15)–(17). However, for the estimates in this paper, we use $\omega_k = \omega(k)$ from Eq. (1) or Ref. 31 and $\Gamma_{\rm eff} = \Gamma_L + \Gamma_c$ from the estimates in Sec. 3.

5. THE POTENTIAL ENERGY OF THE PLASMA AND THE ELECTRIC-FIELD FLUCTUATIONS

5.1. The correlation energy

The mean potential energy of a system with an electromagnetic field is expressed in terms of the mean square electric field (the contribution of the magnetic field is small):

$$U = \langle E^2 \rangle / (8\pi). \tag{18}$$

In thermal equilibrium at temperature T, the space–time spectrum of the electric-field fluctuations and consequently the mean square electric field, according to the fluctuation–

dissipation theorem, are determined by the spectrum of a particular dissipative quantity—the imaginary part of the inverse longitudinal permittivity:[2,30]

$$U=-\frac{T}{(2\pi)^4}\int d^3k\int\frac{d\omega}{\omega}\,\mathrm{Im}\left[\frac{1}{\varepsilon(\mathbf{k},\omega)}\right]. \tag{19}$$

This integral can be computed by various means. If the integration over ω is done first, it is obvious that, because of the analytical properties of the expression under the integral, the integral in Eq. (19) over ω can be expressed in terms of values of $\varepsilon(\mathbf{k},0)$. We represent the factors of the integrand in the form

$$\frac{1}{\omega}=\frac{1}{2}\lim_{\Delta\to 0}\left[\frac{1}{\omega+i\Delta}+\frac{1}{\omega-i\Delta}\right].$$

$$-\mathrm{Im}\left[\frac{1}{\varepsilon(\mathbf{k},\omega)}\right]=\frac{1}{2i}\left\{\left[1-\frac{1}{\varepsilon(\mathbf{k},\omega)}\right]-\left[1-\frac{1}{\varepsilon^*(\mathbf{k},\omega)}\right]\right\},$$

after which, the integral in Eq. (19) can be computed:[2]

$$U=\frac{T}{2(2\pi)^3}\int d^3k\left[1-\frac{1}{\varepsilon(\mathbf{k},0)}\right]=\frac{T}{(2\pi)^2}\int F(k)dk. \tag{20}$$

The integral in Eq. (20) diverges [in particular, in a Debye plasma, the integrand is proportional to $a_e^2k^2/(1+a_e^2k^2)$ and goes to infinity when it is integrated over d^3k]. This is the ordinary Coulomb divergence, associated with the infinite self-energy of the electromagnetic field. There is physical interest in a finite quantity, namely, the energy gain $U_{\mathrm{corr}}=U-U_0$ from the mutual correlations, where U_0 is the energy of an uncorrelated Coulomb system. The quantity U_{corr} is the potential part of the thermodynamic function, i.e., the internal energy of the system. To compute U_0, we use the fact that the correlations weaken as the temperature is increased; i.e.,

$$U_0=\lim_{T\to\infty}U.$$

This limit can be represented in the form of an integral over the wave vectors:[23]

$$U_0=\frac{1}{2(2\pi)^3}\int d^3k\lim_{T\to\infty}\left[T\left(1-\frac{1}{\varepsilon(\mathbf{k},0)}\right)\right]. \tag{21}$$

As $T\to\infty$, the plasma becomes a Debye plasma, and therefore the limit of the quantity in brackets equals $T(\kappa/k)^2=4\pi n_e e^2/k^2$. From this we get that

$$U_{\mathrm{corr}}=-\frac{T}{2a_e^3(2\pi)^2}\int d(ka_e)\times\left\{1-\left[1-\frac{1}{\varepsilon(\mathbf{k},\omega)}\right](ka_e)^2\right\}. \tag{22}$$

The quantity U_0 can also be written in more general form. Since $\varepsilon(\mathbf{k},0)\simeq 1+o(1/T)$ at high temperatures, the integrand in Eq. (21) can be written with the same accuracy as $T(\varepsilon(\mathbf{k},0)-1)$, and U_{corr} can be obtained[39] in the form

$$U_{\mathrm{corr}}=-\frac{T}{2}\int\frac{d^3k}{(2\pi)^3}\frac{[\varepsilon(\mathbf{k},0)-1]^2}{\varepsilon(\mathbf{k},0)}, \tag{23}$$

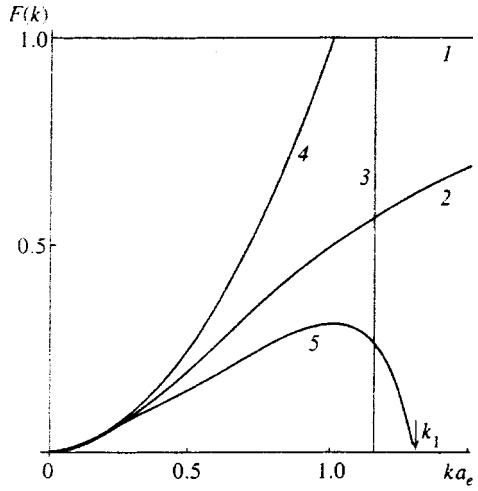

FIG. 3. Energy density $F(k)$ of the longitudinal modes: *1*—uncorrelated charges, *2*—correlated charges, *3*—$k=q_0$, *4*—collective modes in the absence of damping,[30] *5*—collective modes with damping taken into account. The definition of k_1 is given in the text.

which agrees with the general expression for the correlation energy [Eq. (85.15) in Ref. 40]. The integral in Eq. (22) corresponds to the region between straight line *1* [the first term in braces in Eq. (22)] and curve *2* (the second term) in Fig. 3 and is finite. In the case of a Debye plasma, at room temperatures, we get

$$U_{\mathrm{corr}}=-\frac{T}{a_e^3(2\pi)^2}\int_0^\infty\frac{d(ka_e)}{1+(ka_e)^2}, \tag{24}$$

which gives the well-known expression for the Coulomb energy[38]

$$U_{\mathrm{corr}}/NT=-\sqrt{\pi}\gamma^{3/2}.$$

To make the transition to a nonideal plasma, the integration in Eqs. (18)–(24) must be carried not to infinity, but to $k\approx q_0$. In a Debye plasma, $q_0\gg\kappa$, and this restriction is unimportant. In a nonideal plasma, the indicated strong inequality is not satisfied, and the integration must be carried only to q_0 (the region between curves *1*, *2* and *3* in Fig. 3). Then, by extrapolating the Debye expression for $\varepsilon(k,0)$ into the strong-coupling region (this is valid at least in the long-wavelength limit[1)]), for U_{corr} we get

$$U_{\mathrm{corr}}/NT=-(2/\pi^{1/2})\gamma^{3/2}\arctan(q_0a_e). \tag{25}$$

This expression was obtained in Ref. 23 by somewhat different means.

5.2. Electric-field energy of longitudinal plasma waves

There is particular interest in the energy of the longitudinal plasma waves. This energy can be estimated in the equilibrium case by assigning $T/2$ to each collective degree of freedom and assuming that the number of collective degrees of freedom equals the volume of the phase space of wave numbers.[4,5] This volume is about κ^3 for a Debye plasma and about q_0^3 for a nonideal plasma. It is assumed in this estimate that the plasma waves are undamped for $k<\kappa$ (or q_0), whereas these waves do not exist for $k>\kappa$ (or q_0).

To refine the numerical coefficients, compute the energy distribution over k taking damping into account, and obtain a smooth transition from the Debye case to the nonideal case, Eq. (19) should be integrated first over ω and then over k in the region of where these modes exist. In Ref. 30, such an estimate is made for a low-density plasma on the assumption of infinitely small damping, $\mathrm{Im}[\varepsilon(k,\omega)]\to 0$. In this case, the integrand in Eq. (19)

$$A(\omega,k)=\omega^{-1}\ \mathrm{Im}[\varepsilon^{-1}(k,\omega)]$$

is equal to

$$A(\omega,k)=(\pi/2)\Omega_e\delta(\omega-\Omega_e)/\omega, \tag{26}$$

and, after integrating over ω, the ratio of the electric-field energy of the collective modes to the thermal energy of the electrons equals

$$\xi=\frac{\langle E^2\rangle}{8\pi\cdot(3/2)n_eT}=\frac{1}{6\pi^2n_e}\int_0^{k_{\max}}k^2dk. \tag{27}$$

It is assumed in Ref. 30 that $k_{\max}=\kappa$. Then

$$\xi=0.17\gamma^{3/2}. \tag{28}$$

However, if the choice $k_{\max}=2.34\kappa$ is correlated with Fig. 1, we get

$$\xi=2.17\gamma^{3/2}. \tag{29}$$

In such an estimate, the damping of the waves is neglected in the approximation to $k_{\max}=2.34\kappa$; i.e., the estimate given by Eq. (29) overestimates ξ.

It might be thought that the truth lies between Eqs. (28) and (29). However, the estimates of Eqs. (28) and (29) can be refined without assuming that the damping is infinitely small by numerically integrating $A(\omega,k)$ in the region where the collective modes exist [the maximum of $A(\omega_k)$ for a given k] by using the following expression for the permittivity:[30]

$$\varepsilon(k,\omega)=1-(\Omega_e/\omega)^2\Phi(x),$$

$$\Phi(x)=-2x^2\Big[1-2x\ \exp(-x^2)\times\int_0^x\exp(t^2)dt-i\sqrt{\pi}x\ \exp(-x^2)\Big],$$

$$x=\omega/(\sqrt{2}kv_T).$$

It is easy to integrate $A(\omega,k)$ over ω in the region $ka_e\leqslant 0.7$, where the collective modes are well defined (Fig. 4, curves *1* and *2*). The result of integration over ω is given by curve *5* in Fig. 3. It can be seen that this curve already begins to deviate from curves *2* and *4* when $ka_e\approx 0.3$. For larger ka_e values, the estimate acquires an extremely approximate character, since it is now difficult to separate the contributions to the integral of $A(\omega,k)$ from the individual and the poorly defined collective modes (curve *3* in Fig. 4).

The integral over k from zero to infinity, or actually to k_1 ($\kappa<k_1<2.34\kappa$), of the function corresponding to curve *5* in Fig. 3 determines the quantity

$$\xi=0.11\gamma^{3/2}, \tag{30}$$

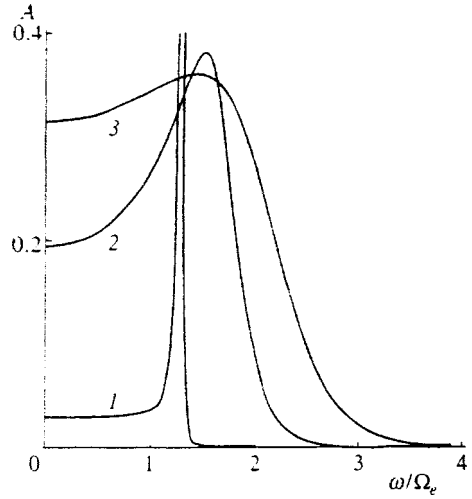

FIG. 4. Power spectrum $A(\omega,k)$ of the longitudinal modes: *1*—$ka_e=0.3$, *2*—0.7, *3*—1.

which is less than the estimates from Eqs. (28) and (29).

In the case of Eq. (29), the limiting value of wave number q_0 becomes comparable with $\alpha\kappa$ at γ_L, whereas, in the case of Eq. (28), this occurs for some value $\gamma>\gamma_L$. The value of ξ reaches its limiting value of 1/3 for these values of γ and then does not change as γ increases. However, according to the estimate given by Eq. (30), only the coefficient begins to decrease at γ values for which q_0 is less than k_1, whereas the value of ξ itself becomes equal to 1/3 when $\gamma\approx 15$.

Recalling that these estimates are approximate, we should expect that the value of γ at which $\xi=1/3$ is attained lies between 0.25 and 15 and can be refined (at this stage of the investigations) by comparing with the results of a numerical experiment.

6. DISPERSIVE BEHAVIOR

We take Eq. (1) as the starting point for estimating the dispersion of the plasma waves. The dispersion term proper in it, $3\Omega_e^2a_e^2k^2$, has a gas-dynamic character. The coefficient for k^2, which equals $3T$, is $r=(\partial P/\partial n)_\eta$ for an ideal gas ($P=nT$) in a process with polytrope index $\eta=3$.[4] For a nonideal plasma, r can be computed by using Eq. (25) for the Coulomb energy. We get

$$r=3T[1-(2/3\sqrt{\pi})\gamma^{3/2}\ \arctan(q_0a_e)]; \tag{31}$$

i.e., the value is smaller than for an ideal gas. The results of the estimate are given in Fig. 5, along with the data of a molecular-dynamics calculation.[13] It can be seen that the qualitative change of the dispersion curves as γ increases corresponds to the tendency of the values of ω_k to decrease, observed in a computer experiment.

The largest divergence between curve *2* and the points obtained from the data of Ref. 13 is observed at large k. This divergence can be associated with the fact that curve *2* and the points in Fig. 5 relate to different values. The points correspond to the maxima of the dynamic structure factor, i.e., to the statistical estimate of $k^2A(\omega,k)$. As far as curves *1* and *2* are concerned, they are the real parts of the solutions

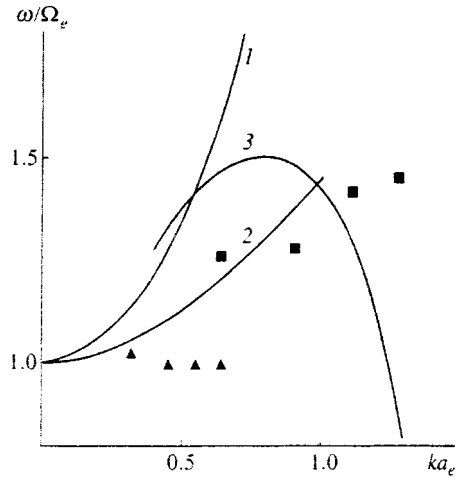

FIG. 5. Dispersion of plasma waves in a nonideal plasma. Calculation: *1*—Refs. 30 and 31, *2*—according to Eq. (31), $\gamma=1.56$, *3*—$\omega_{\max,k}$. Molecular-dynamics data:[13] squares $\gamma=0.39$, triangles $\gamma=1.56$.

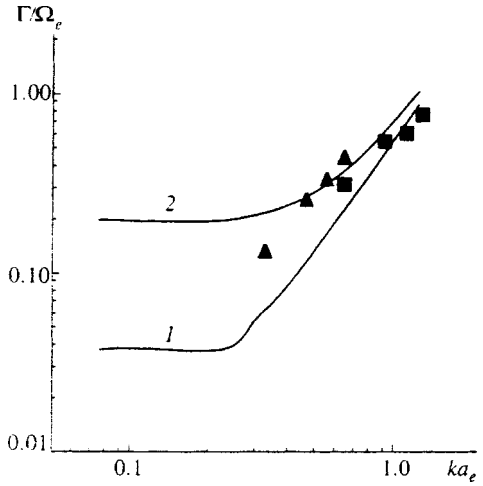

FIG. 6. Damping rate of plasma waves in a nonideal plasma: *1* and *2*—Eq. (12) using an estimate of the Coulomb logarithm from the minimum and the maximum. Molecular-dynamics data:[13] squares $\gamma=0.39$, triangles $\gamma=1.56$.

ω_k' of the dispersion equation $\varepsilon(k,\omega'+i\omega'')=0$ for an ideal and a nonideal plasma. The position of the maximum of $A(\omega,k)$ in ω coincides with the solution ω_k' of the dispersion equation only if the damping ω_k'' of the plasma waves is assumed to be infinitely small. Only then [as in Eq. (26)] is the function $A(\omega,k)$ a singular term of Sokhotskiĭ's formula, and the support of this function reduces to the solution of equation $\varepsilon(k,\omega)=0$.

If the damping is nonzero, this function,

$$A(\omega,k)=-\frac{\operatorname{Im}\varepsilon}{|\varepsilon(\mathbf{k},\omega)|^2\omega},$$

is regular, and its maximum approximately corresponds to the maximum of the first factor or to the minimum of the function $|\varepsilon(\mathbf{k},\omega)|^2=\varepsilon'^2+\varepsilon''^2$ (ε' and ε'' are the real and imaginary parts of the permittivity). Close to the solution of the dispersion equation we have $\varepsilon'\approx A(\omega-\omega_k')$ and $\varepsilon''\approx B+C(\omega-\omega_k')$, and it is possible to write $\varepsilon'^2+\varepsilon''^2\approx(A^2+C^2)(\omega-\omega_k')^2+2BC(\omega-\omega_k')+B^2$. The minimum of such a quadratic trinomial occurs at $\omega_{\max,k}=\omega_k'-BC/(A^2+C^2)$. For small damping, the values of B and C are infinitesimal. As k increases, the B and C values increase while C remains positive, and $\omega_{\max,k}$ shifts toward smaller ω_k' values.

Figure 5 shows the dependence of $\omega_{\max,k}$ for an ideal plasma, calculated numerically from curves of the form shown in Fig. 4. It can be seen that the $\omega_{\max,k}$ curves are below curve *1* in Fig. 5 when $ka_e>0.4$. The same effect with respect to curve *2* is also expected for a nonideal plasma. An estimate of the maximum displacement (for $k\approx q_0$) gives a value of order $\Omega_e\gamma^{5/2}$.

The determination of the correlation energy in Sec. 5 and the influence of the equation of state on the $\omega(k)$ dependence makes it possible to connect the approach developed in this paper with the approach in which the main effect comes from computing the contribution of the short-range (quantum) interaction to the thermodynamic functions (see, for example, Ref. 28 and the references therein). First, the corresponding equations of state can be used in calculating r; in this case, however, the correlation energy of Eq. (25) should be supplemented by the short-range contribution, as is done for a Debye plasma. Second, it can be expected that there are short-range effects on the dispersion characteristics for k values close to $k_{\max}$, as occurs for the phonon spectra.

The dependence of the damping rate on the wave number is shown in Fig. 6. The two analytical curves correspond in their minimum and maximum to the estimates of the Coulomb logarithm in Eq. (12), made in connection with the indeterminacy of the numerical coefficients mentioned in Sec. 3.

The estimate from Sec. 3 agrees fairly well with the results of the numerical simulation of Ref. 13 in the region of k where Landau damping dominates. Unfortunately, in the region of small k, where the main contribution must be from collisional damping, there are no results of a molecular-dynamics calculation.

Table I shows the results of an experimental observation of plasma waves. References 6–8, 41, and 42 only recorded that plasma oscillations exist for some values of the charge density and temperature. Values of $\omega(k)$ for the conditions of Ref. 8 were found in Ref. 9 by interpreting the results of electrical conductivity measurements. It follows from Table I that plasma waves are reliably observed even in the region of values $\gamma\simeq 0.2-1$ for the coupling constant, where, according to Sec. 3, the total damping rate must reach a maximum.

7. INDIVIDUAL DEGREES OF FREEDOM

By differentiating Eq. (15) with respect to p and q, we can find the equations of motion for the quasiparticles (electrons) corresponding to the individual degrees of freedom:

TABLE I. Values of the coupling constant for which plasma oscillations were experimentally observed.

	$\gamma=0.2$	$\gamma=0.5$	$\gamma=0.2-1$	$\gamma=0.5$	$\gamma=0.25$
ω/Ω_e	–	–	1.0 ± 0.1	–	–
Refs.	6	7	8	41	42

$$\frac{dp_i}{dt} = \nabla \sum_j \Phi(q_{ij}) + f_{\text{in-os}}. \tag{32}$$

Here $\Phi(q_{ij})$ is the potential energy of the interaction of electron i with charge j (including ions), and the force f is created by the electric field of the collective modes. We recall that, in a nonideal plasma, the potential Φ in Eq. (32) should be considered the short-range core of the Coulomb potential, and scattering off it is strong but effectively pairwise. Collective scattering at large distances, on the other hand, is described by the second term in Eq. (32).

In the simplest form, the forces in Eq. (32) can be specified as follows: When the collective variables are separated out, the long-wavelength part of the Coulomb interaction forms a field of plasma waves.[4,21] In a Debye plasma, the limiting vector that separates the collective fields from the short-wavelength interaction ($\sim\kappa$) was not sufficiently well-defined. This, however, was compensated by the fact that the actual contribution came only from the regions $k \ll \kappa$ and $k \gg \kappa$, and the plasma was correctly described.

For significantly strong coupling, the situation changes. Now the limiting vector is defined unambiguously—it is q_0, and the integral

$$\Phi(r) = 4\pi \int \frac{d^3k}{k^2} \exp(-i\mathbf{k}\cdot\mathbf{r}),$$

taken over $k > q_0$ can be used as the potential Φ.

Both force terms in Eq. (32) correspond to certain scattering cross sections and the electron relaxation times or their effective scattering frequencies defined in terms of these cross sections. The considerations explained in Sec. 3 in estimating the collisional damping of plasma waves are valid for the collisional part of the effective frequency ν_c, and ν_c differs from Γ_c from Eq. (13) only by a coefficient:[1,43]

$$\nu_c = (3\pi/16\alpha_e)\Gamma_c,$$

where, in a singly ionized low-density plasma, the coefficient $\alpha_e = 0.582$, according to Ref. 43, allows for the electron–electron interaction. We shall use the same value of α_e for the estimates.

The scattering cross section for the interaction of electrons with plasma waves was first computed in Ref. 44 and equalled the Rutherford cross section but without the Coulomb logarithm. The interaction frequency, however, is given by the expression $\nu_{\text{os}} = \xi\Omega_e$. The total effective frequency of the ''collisions'' can be written in the form

$$\nu_{\text{eff}}/\Omega_e = \nu_c/\Omega_e + \nu_{\text{os}}/\Omega_e. \tag{33}$$

Note that a formula of the form Eq. (33) was first used to estimate the electrical conductivity of a nonideal plasma in Ref. 36. For small γ, both terms in Eq. (33) increase as $\gamma^{3/2}$, whereas, for large degrees of nonlinearity, the first term decreases as $\gamma^{-1/2}$ while the second approaches a constant $\approx 1/3$ along with the total effective frequency. The results of an estimate for the two expressions for ξ from Eqs. (28) and (30) are shown in Fig. 7, in comparison with the data of a molecular-dynamic experiment of Ref. 45 with regard to electronic relaxation times. It can be seen that taking plasma waves into account strongly changes the estimates of the collision frequency and, along with it, the electronic transport properties of the plasma. The latter question deserves a separate treatment.

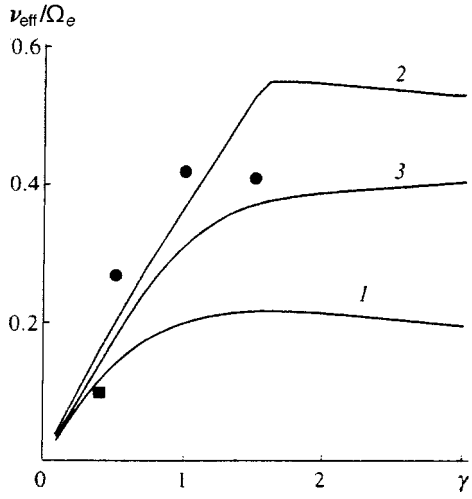

FIG. 7. Effective frequency of electron collisions: *1*—collisional contribution, *2* and *3*—the sum of Eq. (33), where ξ is estimated from Eqs. (28) and (30), respectively. Molecular-dynamics results: points from Ref. 45, square from Ref. 13 (obtained for a potential smaller than in Ref. 45).

8. CONCLUSION

The construction of a theory of a nonideal plasma at the same level of rigor as for a Debye plasma is made possible by the smallness of the damping Γ/Ω_e. However, this theory cannot be constructed by expanding in the small parameter Γ/Ω_e, since no analytical representations have yet been found that describe the properties of the plasma as a function of this parameter.

The identification of the contribution of the collective degrees of freedom in the Hamiltonian is associated with the most general properties of the motion of charged particles; i.e., the topology of the isoenergetic hypersurface in the phase space of the system must be equivalent to the topology of a torus. The dimension of this torus determines the number of oscillational degrees of freedom. The validity of these assertions would support the correctness of the approach of Ref. 37 in the general case, as well as the possibility of using it for a nonideal plasma.

The shape of the phase hypersurface needs to be studied using computer simulation. The topological properties of the phase trajectory must then be studied by modern methods of differential geometry and mathematical statistics. This will confirm the validity of the assumed approach to the theory of a nonideal plasma.

We thank A. A. Rukhadze and M. V. Netsina for interest in the work and for support.

*)E-mail: henry@aha.ru

1)The Debye radius determines only the asymptotic limits of the pairwise correlation functions at distances greater than $n^{-1/3}$.

[1] E. M. Lifshitz and L. P. Pitayevsky, *Physical Kinetics* (Nauka, Moscow, 1979; Pergamon Press, Oxford, 1981).

[2] Yu. L. Klimontovich, *Kinetic Theory of Nonideal Gases and Nonideal Plasmas* (Nauka, Moscow, 1975; Pergamon Press, Oxford, 1982).
[3] B. B. Kadomtsev, *Collective Phenomena in a Plasma* (Nauka, Moscow, 1976).
[4] L. A. Artsimovich and R. Z. Sagdeev, *Plasma Physics for Physicists* (Atomizdat, Moscow, 1979).
[5] A. F. Alexandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Principles of Plasma Electrodynamics* (Vysshaya Shkola, Moscow, 1978; Springer-Verlag, Berlin, 1984).
[6] B. Ya'akobi and G. Bekefi, Phys. Lett. **130**, 539 (1969).
[7] M. Skowronek, J. Rous, A. Goldstein, and F. Cabannes, Phys. Fluids **13**, 378 (1970).
[8] I. Ya. Dikhter and V. A. Zeĭgranik, Dokl. Akad. Nauk SSSR **227**, 656 (1976).
[9] É. I. Asinovskiĭ and A. A. Valuev, Teplofiz. Vys. Temp. **18**, 1318 (1980).
[10] G. E. Norman and A. A. Valuev, in *Proceedings of the Twelfth IGPIG*, Eindhoven, 1975, p. 257.
[11] A. A. Valuev, Teplofiz. Vys. Temp. **15**, 1143 (1977).
[12] G. E. Norman and A. A. Valuev, Plasma Phys. **21**, 531 (1979).
[13] J. P. Hansen and I. R. McDonald, Phys. Rev. A **23**, 2041 (1981).
[14] Yu. K. Kurilenkov and A. A. Valuev, Plasma Phys. **24**, 529 (1984).
[15] V. M. Adamyan and I. M. Tkachenko, Teplofiz. Vys. Temp. **21**, 417 (1983).
[16] K. I. Golden, Phys. Rev. A **35**, 5278 (1987).
[17] G. Kalman, in *Physics of Nonideal Plasmas*. Teubner-Texte zur physik, vol. 26, Eds. W. Ebeling, A. Forster, and R. Radtke (Teubner, Stuttgart, 1992), p. 167.
[18] G. Kalman and K. I. Golden, Phys. Rev. A **41**, 5516 (1990).
[19] J. Ortner and I. M. Tkachenko, Phys. Rev. A **46**, 7882 (1992).
[20] F. A. Gutierrez and M. A. Girardeau, Phys. Rev. A **42**, 936 (1990).
[21] D. Bohm, *General Theory of Collective Variables* (Moscow, Mir, 1964).
[22] D. N. Zubarev, Dok. Akad. Nauk SSSR **95**, 757 (1954).
[23] A. S. Kaklyugin, Teplofiz. Vys. Temp. **23**, 217 (1985).
[24] A. S. Kaklyugin, in *Abstracts of Reports of the Fourth All-Union Conference on Gas-Discharge Physics, Makhachkala, 1988* (DGU Press, Makhachkala, 1988), p. 33.
[25] A. A. Valuev, A. S. Kaklyugin, and G. É. Norman, in *Radiation Thermodynamics*, vol. 1, Ed. Yu. S. Protasov (Énergoatomizdat, Moscow, 1991), p. 396.
[26] A. A. Valuev, A. S. Kaklyugin, and G. É. Norman, Matemat. Modelir. **4**, No. 12. 178 (1992).
[27] A. S. Kaklyugin, G. É. Norman, and A. A. Valuev, in *Proceedings of the International Conference on Physics of Strongly Coupled Plasmas*, ed. W. D. Kraeft and M. Schlanges (World Scientific, Singapore, 1996), p. 435.
[28] W. Ebeling, A. Forster, V. E. Fortov, V. K. Gryaznov, and A. Ya. Polishchuk, *Thermophysical Properties of Hot Dense Plasmas*. Teubner-Texte zur physik, vol. 25, (Teubner, Stuttgart, 1992).
[29] V. E. Fortov and I. T. Iakubov, *Physics of Nonideal Plasma* (Énergoatomizdat, Moscow, 1994; Hemisphere, New York, 1990).
[30] G. Bekefi, Ed., *Radiation Processes in Plasmas* (Wiley, New York, 1966; Mir, Moscow, 1971).
[31] M. A. Lieberman and A. Bers, Quarterly Prog. Rept. No. 77, Res. Lab. of Electronics, M. I. T. (1965), p. 141.
[32] S. I. Andreev and N. F. Ivasenko, *Principles of Calculations for Pulsed Xenon Lamps* (Izdvo. Tomsk Univ., Tomsk, 1982), p. 44.
[33] I. T. Iakubov and A. G. Khrapak, in *Transport and Optical Properties of Nonideal Plasma*, Ed. G. A. Kobzev, I. T. Iakubov, and M. M. Popovich (Plenum Press, New York, 1995), p. 1.
[34] J. M. Ziman, *Principles of the Theory of Solids* (Cambridge Univ. Press, London, 1969; Mir, Moscow, 1966).
[35] B. V. Zelener, G. É. Norman, and V. S. Filinov, Teplofiz. Vys. Temp. **12**, 267 (1974).
[36] A. A. Valuev and Yu. K. Kurilenkov, Teplofiz. Vys. Temp. **21**, 591 (1983).
[37] V. I. Tatarskiĭ, Usp. Fiz. Nauk **151**, 273 (1987) [Sov. Phys. Usp. **30**, 134 (1987)].
[38] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, part 1 (Nauka, Moscow, 1995; Pergamon Press, Oxford, 1980).
[39] A. S. Kaklyugin, Dissertation for candidate of physicomathematical sciences, IVTAN, Moscow, 1987.
[40] E. M. Lifshitz and L. P. Pitayevsky, *Statistical Physics*, part 2 (Nauka, Moscow, 1978; Pergamon, Oxford, 1980).
[41] K. H. Finken and U. Ackerman, Physica B **113**, 135 (1982).
[42] M. Ketlitz and R. Radtke, in *Physics of Nonideal Plasmas*. Teubner Text zur physik, vol. 26. Ed. W. Ebeling, A. Forster, and R. Radtke (Teubner, Stuttgart, 1992), p. 233.
[43] L. Spitzer, *Physics of Fully Ionized Gases* (Interscience, New York, 1962; Mir, Moscow, 1966).
[44] B. I. Danilov, in M. A. Leontovich (Ed.), *Plasma Physics and the Problem of Controllable Thermonuclear Reactions* (Izdvo. Akad. Nauk SSSR, Moscow, 1958), p. 77.
[45] A. A. Valuev, I. V. Morozov, and G. É. Norman, in *Abstracts of Reports of the Twenty-Fourth Zvenigorod Conference on the Physics of Plasma and UTS, Zvenigorod, 1997* (RIIS FIAN, Moscow, 1997), p. 236.

Translated by W. J. Manthey

Experiments on two-step heating of a dense plasma in the GOL-3 facility

V. T. Astrelin, A. V. Burdakov,*) V. S. Koĭdan, K. I. Mekler, P. I. Mel'nikov, V. V. Postupaev, and M. A. Shcheglov

G. I. Budker Institute of Nuclear Physics, Siberian Branch, Russian Academy of Sciences, 630090 Novosibirsk, Russia
(Submitted 15 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 897–917 (March 1998)

This paper presents the results of experiments on two-stage heating of a dense plasma by a relativistic electron beam in the GOL-3 facility. A dense plasma with a length of about a meter and a hydrogen density up to 10^{17} cm^{-3} was created in the main plasma, whose density was 10^{15} cm^{-3}. In the process of interacting with the plasma, the electron beam (1 MeV, 40 kA, 4 μs) imparts its energy to the electrons of the main plasma through collective effects. The heated electrons, as they disperse along the magnetic field lines, in turn reach the region of dense plasma and impart their energy to it by pairwise collisions. Estimates based on experimental data are given for the parameters of the flux of hot plasma electrons, the energy released in the dense plasma, and the energy balance of the beam–plasma system. The paper discusses the dynamics of the plasma, which is inhomogeneous in density and temperature, including the appearance of pressure waves. © *1998 American Institute of Physics.* [S1063-7761(98)01003-8]

1. INTRODUCTION

Studies in the physics of the interaction of powerful relativistic electron beams with a plasma have, besides general plasma-physics purposes, the task of developing a method of heating the plasma for a reactor based on a multiple mirror confinement system.[1–3] As is well known, such a reactor must have a plasma with a density of 10^{17}–10^{18} cm^{-3} in order to possess acceptable technical parameters.[3] At the same time, because of collective effects (mainly because of the development of Langmuir turbulence), for existing parameters of relativistic electron beams, energy transfer from a beam to a plasma attains high efficiency when the plasma density does not substantially exceed 10^{15} cm^{-3}.[4–7] When the plasma reaches a density of even $(3-5)\times10^{15}$ cm^{-3}, the beam imparts virtually none of its energy to it. This is because, as the plasma density increases, the growth rate of the beam instability becomes less than the collision frequency, and instability does not develop.

To avoid this limitation on the density, so-called two-stage heating of a dense plasma was proposed in Refs. 3 and 8. The essence of this system is that the plasma is divided into a region with a density of about 10^{15} cm^{-3}, located at the center of the facility (efficient relaxation of the beam occurs in it, with the beam energy being mainly imparted to the electrons of the plasma—see, for example, Refs. 9 and 10), and two adjacent regions with a plasma density of $>10^{17}$ cm^{-3} suitable for tandem-mirror confinement. In such a system, the electrons of the low-density plasma heated by the beam disperse along the magnetic field, are retarded by pairwise Coulomb collisions in the dense plasma, and thus heat it.

The GOL-3 facility was created at the Institute of Nuclear Physics, Siberian Branch, Russian Academy of Sciences, to study the heating of a dense plasma by means of a relativistic electron beam and the physics of multiple-mirror confinement. The first phase of this facility is intended for studying the interaction of a microsecond electron beam with a plasma.[11] The first phase, the GOL-3-I, is distinguished from the full-scale facility by a shorter magnetoplasma system and lower beam energy.

A plasma column inhomogeneous in density can be formed in an experiment by several methods. One method is to place thin foils across the plasma column. The material that comprises a foil heats up, quickly vaporizes, is ionized under the action of the flux of plasma electrons, and forms an expanding gas–plasma cloud with high initial density. A flux of plasma electrons having an energy of 1–5 keV was recorded at the surface of the end foil in model experiments with nanosecond electron beams;[12] heating and expansion of the surface layer of this foil was observed. Experiments in which thin organic films were used as a target were carried out in the GOL-3 facility. These experiments are briefly described in Ref. 13, where it is shown that the beam efficiently heats the dense plasma of the foil in a two-stage manner.

In this paper, to form a dense plasma bunch, a cloud of hydrogen with a given length and density was admitted into the discharge chamber. This modification of the two-stage heating method is shown schematically in Fig. 1. Such a method of obtaining a region with a dense plasma is convenient in allowing various parameters to be measured immediately inside the cloud. Moreover, it is also more efficient, since there is no need to replace the combustible foils. Preliminary summaries of parts of these experiments have been published in Ref. 14.

2. DESCRIPTION AND OPERATING REGIME OF THE GOL-3-I FACILITY

The layout of the facility is shown in Fig. 2, and a detailed description of it is given in Ref. 11. We recall briefly

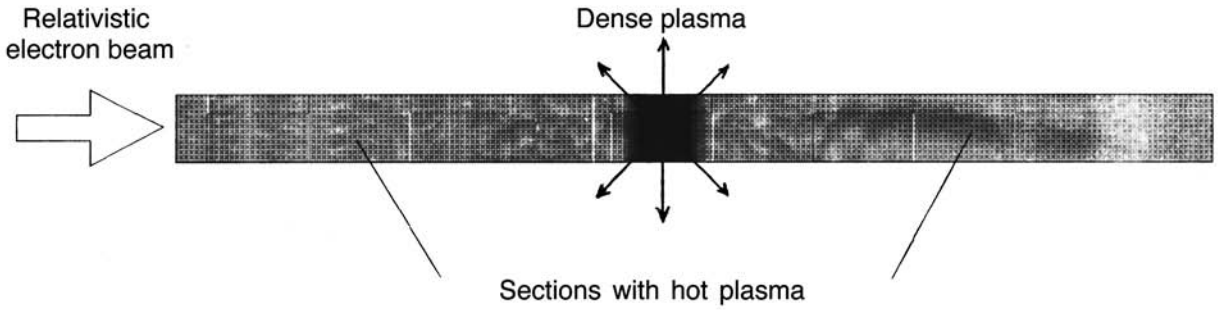

FIG. 1. Two-stage heating system.

that the GOL-3-I facility consists of a U-3 electron-beam generator, a plasma chamber inside a solenoid with a uniform field of up to 6 T at a length of 7 m and 12 T at the individual mirrors at the ends, a 10-MJ capacitor bank for supplying the solenoid, and systems for control, monitoring, and diagnostics. In the experiments discussed in this article, the facility was operated with the following parameters: The magnetic field was 5.5 T in the uniform part of the solenoid and 11 T at the mirrors, and the initial hydrogen plasma column was 7 m long and 8 cm in diameter. The electron beam had an energy of 0.8–0.9 MeV, the maximum current density in the plasma was about 1 kA/cm^2, the beam diameter in the plasma was 6 cm, the pulsewidth was 4 μs, and a typical energy content in the beam was 70 ± 20 kJ (Fig. 3).

The following beam parameters were measured in the experiments: the currents, the voltage on the cathode, and the total energy; moreover, a magnetic analyzer of the beam spectrum was installed at the output of the plasma chamber in certain experiments. Optical interferometers with a working wavelength of 3.39 and 0.63 μm, diamagnetic sensors, and detectors of soft x rays and VUV radiation were used to measure the parameters of the plasma and of the gas cloud. The broadening of the H_α line profile was also measured by means of a polychromator with a dissector tube. Two systems of Thomson scattering of the laser radiation were used in the experiments. The diagnostic facility is described in more detail in Ref. 7. In the rest of this paper, the z coordinate along the magnetic field is measured from the middle of the input mirror.

3. DESIGN OF THE EXPERIMENTS AND FORMING THE GAS CLOUD

To directly model the two-stage heating of a dense plasma on the GOL-3-I facility, a series of experiments was performed in which a high-density hydrogen cloud was used as a target for slowing down fast plasma electrons. Here and below in the text, the term *cloud* will relate to a section of increased density before the beam is injected, and we shall call this object in the process of heating and subsequent dispersal a *plasma bunch*. The cloud in the experiments described here was created by means of a local pulsed gas puffing. The gas pulse was formed by means of positive-action electromagnetic valves,[15] optimized to have short operating times and high gas throughput. Two assemblies of four valves each were mounted on the facility. One assembly was placed at a distance of $z=12.40$, or 270 cm from the input foil, and the other at $z=575$ cm. The valves could be operated separately or together. The structure of the valve made it possible to obtain up to 10^{21} hydrogen molecules per pulse from each assembly.

The experimental scenario was as follows: Hydrogen with a concentration corresponding to the required plasma density in the uniform part was first admitted into the vacuum chamber of the facility through a palladium inlet valve. The concentration of the hydrogen thus admitted could be varied over the range 10^{13}–10^{16} cm^{-3}. Most of the experiments were performed with a uniform plasma density of $(3-5)\times 10^{14}$ cm^{-3}, at which a relativistic electron beam

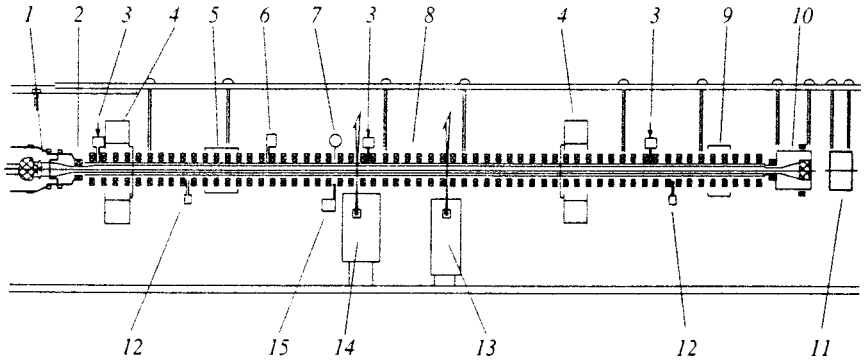

FIG. 2. Layout of experiment GOL-3-I and placement of the diagnostic facility. *1*—U-3 beam-generator diode, *2*—input mirror (coordinate $z=0$), *3*—pulsed valves, *4*—interferometers, *5*—x-ray detectors, *6*—pyroelectric bolometer, *7*—VUV detector, *8*—solenoid with vacuum chamber, *9*—electrooptic detector of soft x rays, *10*—exit module (beam detector, beam-spectrum analyzer), *11*—electrooptic detector of hard x rays, *12*—lightguides to the spectrum analyzer, *13*, *14*—Thomson-scattering systems, *15*—VUV sensor.

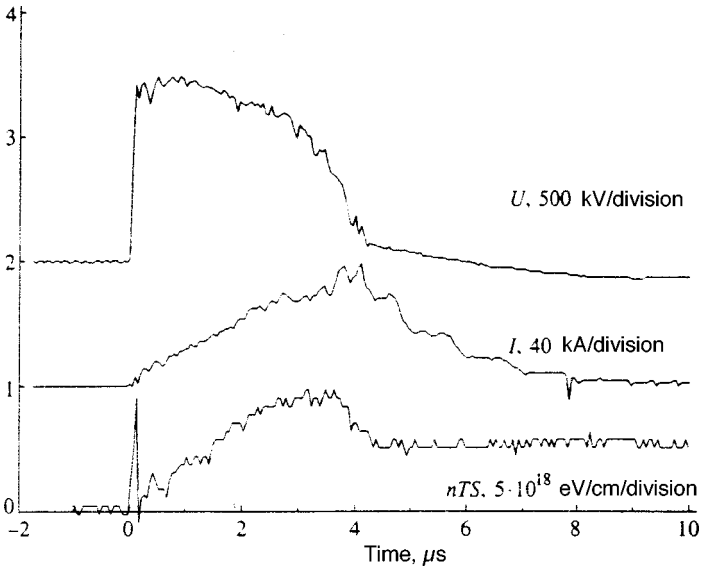

FIG. 3. Typical oscilloscope tracings of the cathode voltage U, the beam current I, and the signal nTS of the diamagnetic sensor.

efficiently interacts with a plasma. After the chamber was filled with hydrogen, the magnetic field was switched on, and then the pulsed gas valves were opened, forming a dense cloud. At the proper time, a forward discharge was triggered, which formed the initial plasma channel. After the preliminary plasma was formed, an electron beam was injected into the chamber. The length of the gas cloud and its density and location could be varied.

In experiments with a dense gas bunch, the forward-discharge operating regime differed somewhat from the standard regime, described earlier in Ref. 7. The discharge current flowing through the plasma was decreased by a small factor. This had the effect that the hydrogen was incompletely ionized even in the homogeneous part. The resulting plasma density was nevertheless sufficient for normal injection of the beam. Ionization is subsequently quickly completed, partially by the electron beam, but mainly by the heated plasma electrons.

To accurately determine the parameters of the hydrogen cloud, the expansion of the gas through the chamber was studied on a special test stand that simulated a 2-m segment of the plasma chamber of the facility. The gas density was measured on the test stand and in the GOL-3 facility by an interferometer operating at a wavelength of 0.63 μm. The dependence of the hydrogen density on time is shown in Fig. 4. The cloud reaches its maximum density at time 1.3 ms at the site where the gas is injected. As the initial gas pressure is reduced, the hydrogen concentration decreases linearly in the valve. The integral of the density along the axis of the facility was also measured. The scatter of the experimental points in measurements over a large series of valve triggerings was 10%, which corresponds to the error in determining the density from interferometric measurements. The accuracy in measurements by means of the interferometer was determined by the detector noise and by vibrations, and therefore the gas-density profile in regions with a density below 2×10^{15} cm^{-3} was determined by solving the one-dimensional self-similar problem of the flow of gas along a tube.[16,17]

4. HEATING A UNIFORM PLASMA

We shall briefly present the main results of a study of the interaction of a relativistic electron beam with a uniform plasma in the GOL-3-I facility (they are explained in Ref. 7), which are required in discussing and interpreting the results of experiments on the two-stage heating of a dense plasma. High transfer efficiency of the beam energy to the plasma is obtained in the experiments when the density is as much as $(1-2)\times10^{15}$ cm^{-3}. As the plasma density is increased above this value, a decrease was observed in the relaxation efficiency of the relativistic electron beam.

During heating, the energy content of the plasma increases approximately linearly until the instant that the power of the injected beam is sharply decreased. Figure 3 shows a typical diamagnetic signal, measured by a sensor at a distance of $z=40$ cm from the input foil. Inhomogeneity of the energy release over the length of the facility is observed and persists during the entire time that the beam is being injected because of anomalously low longitudinal electronic thermal conductivity.

Measurements of the energy content and the energy spectrum of the beam at the output from the plasma show that the beam loses as much as 25% of its energy as a result of collective interaction with the plasma. The main, thermal-

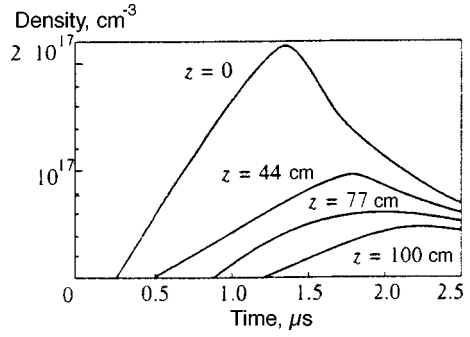

FIG. 4. Molecular density of hydrogen vs time, measured across the vacuum chamber at various distances from the site where the valves are installed.

ized plasma component, with a density of about 10^{15} cm^{-3}, can have a temperature of about 1 keV in the region of the maximum energy release (Thomson diagnostics located at $z=270$ cm were used to measure a plasma temperature of 0.6 keV at the maximum of the heating at the indicated density). Besides this, there is a group of superthermal electrons that contains a significant fraction of the energy lost by the beam. The characteristic energy of these electrons exceeds at least 10 keV, and their instantaneous density at the end of the heating pulse is several percent of the plasma density. The total power of the beam of superthermal electrons at the end has a scale of 10 MW/cm^2, scaled to a magnetic field of 5 T.

The total energy content of the plasma column can reach 3.5–4 kJ, which is $\approx 5\%$ of the total energy of the beam under optimum conditions. The remaining energy lost by the beam escapes at the ends of the facility during the heating pulse and can be used for heating the dense plasma bunch.

Cooling of the plasma after the beam injection ceases is well described by classical electron thermal conductivity to the ends. The measured transverse energy losses from the plasma are insignificant.

An important difference between the physics of experiments in the GOL-3-I facility and experiments with nanosecond beams[4–6] is the increased role of slower, including macroscopic, processes (motion of the plasma, heat transfer, stability, charge exchange processes), which in some cases substantially determines what processes occur.

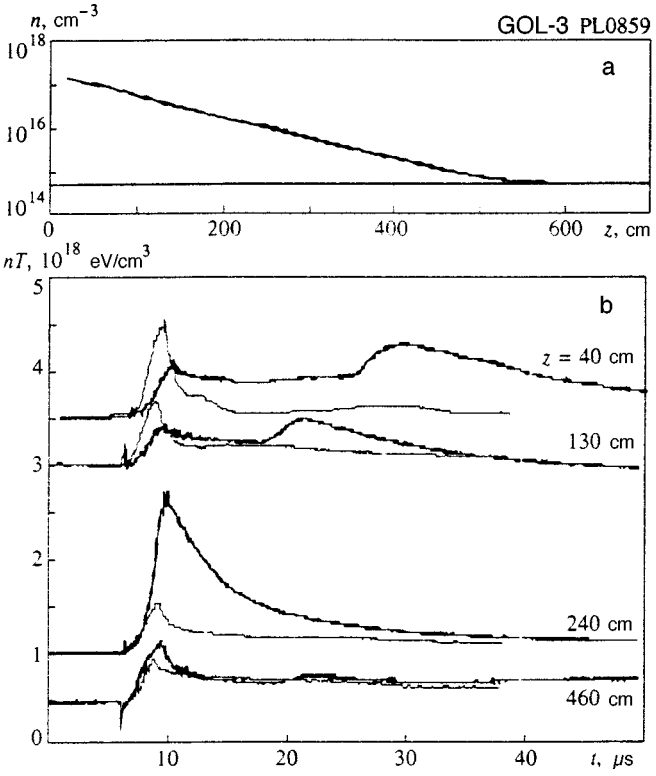

FIG. 5. (a) Distribution of the atomic density of hydrogen over the length of the facility when the beam injection is delayed by 1.75 ms relative to triggering of the valve. (b) Plasma pressure at several points along the length of the plasma column. The thin curve indicates injection into a uniform plasma.

5. RESPONSE TO HEATING OF A DENSE EXTENDED BURST

The first cycle of experiments was devoted to a study of the process of heating a cloud whose maximum density of $(0.2–2)\times 10^{17}$ cm^{-3} occurs at the point where the gas is admitted. The pulsed valves were placed close to the input foil (at $z=12$ or 40 cm). The following quantities were varied in the experiments: the initial pressure in the valves, in the range 2–15 atm (determines the maximum density for a constant density profile along the length); the plasma density in the homogeneous part, in the range $3\times 10^{13}–10^{16}$ cm^{-3} (determines the efficiency with which the beam interacts with the plasma and to a significant extent the spectrum of the hot plasma electrons); the delay of the triggering of the beam with respect to the time that the valves are switched on, in the range 0.5–2.5 ms (determines the length of the cloud for a weakly varying maximum density). These experimental parameters correspond to a length of the cloud equal to 0.5–4 m at the instant when the beam is injected (the length of the cloud is conventionally defined by the region where the density exceeds 5×10^{15} cm^{-3}; we recall that, for such a plasma density, the beam no longer interacts with it directly).

The two-stage heating effect of a dense bunch is illustrated in Fig. 5. The gas-density distribution for our experiment is also shown in this figure. With a valve-triggering delay of 1.75 ms, the gas cloud is about 300 cm long. It can be seen that the plasma pressure nT at the point $z=240$ cm (close to the limit of the bunch) is a factor of 3–4 higher than in the case of a uniform plasma. The plasma pressure decreases somewhat in the depth of the bunch but is always much higher than the pressure obtained when the beam directly interacts with a plasma whose density equals the local density of the bunch. A substantial increase of the plasma cooling time is also observed in the depth of the bunch, which is evidently associated with the decreased role of electronic thermal conductivity to the ends as the plasma density increases and its temperature decreases.

As the length of the cloud decreases, the region of maximum plasma pressure follows its boundary. Figure 6 shows a set of signals from the diamagnetic sensors for a delay of 0.5 ms (the cloud is about 50 cm long). It can be seen that the plasma pressure now sharply increases at $z=40$ cm. For this size of bunch, the maximum plasma pressure is $nT=2.8\times 10^{18}$ eV/cm^3 for a local plasma density of about 10^{16} cm^{-3}. Note that the pressure peak is always close to the boundary of the bunch and that the length of the increased-pressure region depends on the gas-pressure distribution over the length of the facility.

Starting from the diamagnetic measurements and measurements of the plasma density and temperature by means of a laser-scattering system, the plasma parameters were determined in a dense bunch and in a uniform plasma. Figure 7 shows the results of such measurements. It can be seen that, for the case shown here, at the point of the laser measurements ($z=270$ cm), the plasma temperature at the maximum of the heating reaches 0.18 keV when the density is 6×10^{15} cm^{-3}.

As the delay is increased, there is a time when the gas

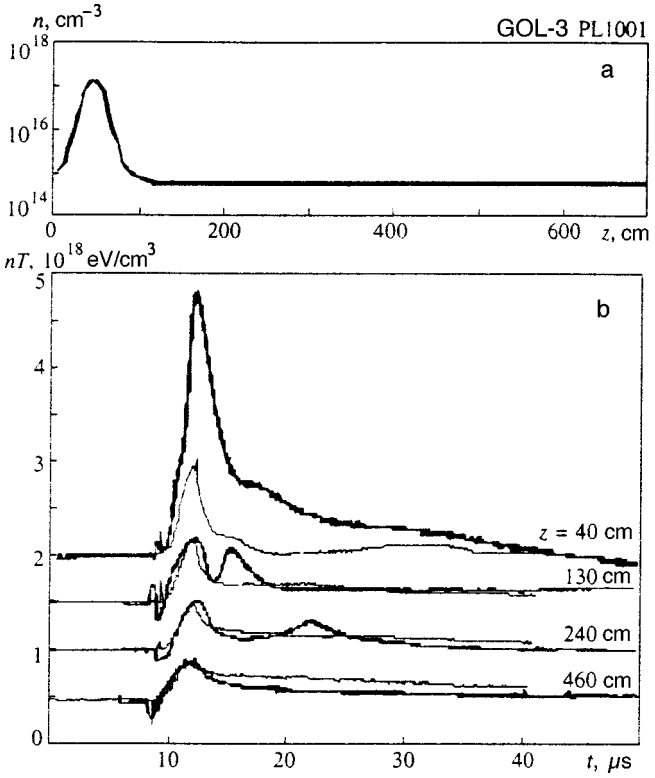

FIG. 6. (a) Distribution of the atomic density of hydrogen over the length of the facility when the beam injection is delayed by 0.5 ms relative to triggering of the valve. (b) Plasma pressure at several points along the length of the plasma column. The thin curve indicates injection into a uniform plasma.

cloud fills a substantial length of the facility. In this case, the region of efficient interaction becomes small, the heating efficiency drops, and hence a smaller fraction of the beam energy is transferred to the bunch of hot plasma electrons. When the density of the homogeneous plasma exceeds 10^{15} cm^{-3}, the efficiency with which the beam interacts with the homogeneous plasma decreases, and the energy transport into the dense bunch accordingly decreases sharply.

6. ABSORPTION OF FAST ELECTRONS IN THE BURST

The set of experiments that was carried out showed that the parameter that determines the character of the energy release in the bunch (while identical conditions are maintained with respect to the relaxation of the beam in the homogeneous part of the plasma) is the quantity

$$\langle nl\rangle = \int (n(l)-n_0)dl,$$

where $n(l)$ and n_0 are the local density of the cloud and the density of the homogeneous plasma, and l is the distance measured from the uniform plasma into the interior of the bunch. For cases with different distributions of the hydrogen concentration over the length of the facility, the plasma parameters by the end of the heating pulse are identical for points with an identical value of $\langle nl\rangle$. Therefore, even though the sensors of the plasma pressure are located at a significant distance from one another, the set of experiments makes it possible to study the dependence of the plasma parameters on $\langle nl\rangle$ with far better spatial resolution.

The results of diamagnetic measurements were used to find the character of the distribution of the absorbed energy as a function of the depth of the bunch. Each measurement corresponds to a definite value of the gas density n at the site where a sensor is placed; accordingly, the energy E_a per starting atom was computed at this point. Since the absolute temperatures are not very large in the densest part of the bunch, the hydrogen ionization energy E_i^a (taken from Ref. 18) was allowed for in the calculations.

Let us first consider how the temperatures of the electrons and ions in the bunch are connected as a function of E_a. In the part of the bunch adjacent to the homogeneous plasma, the diamagnetic signal

$$W=(3/2)n(T_e+T_i)S$$

is caused only by the electrons ($T_e \gg T_i$), since the ions are not susceptible to heating. For $n=7\times10^{15}$ cm^{-3} and $T_e=200$ eV ($\langle nl\rangle\sim4\times10^{17}$ cm^{-2}) the energy-exchange time between the electrons and ions is 10 μs, whereas, for $n=2\times10^{16}$ cm^{-3} and $T_e=50$ eV ($\langle nl\rangle\sim10^{18}$ cm^{-2}) it decreases to 0.2 μs. The plasma can therefore conventionally be assumed to be isothermal for $\langle nl\rangle>10^{18}$ cm^{-2}. In this case we have $T_e=T_i$. This relationship remains valid until $\langle nl\rangle\sim10^{19}$ cm^{-2}, when the temperature decreases so much that the plasma becomes incompletely ionized. It can be found by solving the Saha equation that the degree of ionization is still fairly high at $\langle nl\rangle=10^{19}$ cm^{-2} and amounts to 95%, whereas it decreases to $\approx$50% when $\langle nl\rangle\approx1.5\times10^{19}$ cm^{-2}.

The dependence of E_a on $\langle nl\rangle$ was next plotted (Fig. 8). As can be seen, the energy per electron–ion pair is about 1 keV at the boundary of the bunch and then decreases with increasing $\langle nl\rangle$. The character of the falloff is associated

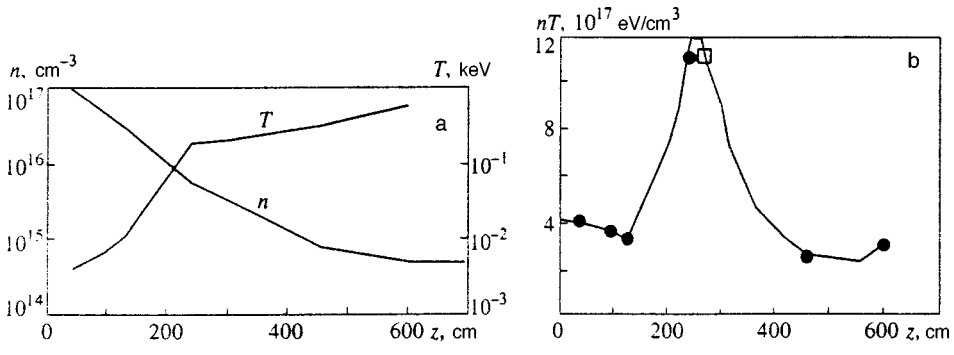

FIG. 7. Distribution of density n, temperature T, and pressure nT over length for a long cloud, compared with the data of diamagnetic measurements (circles) and Thomson scattering (squares).

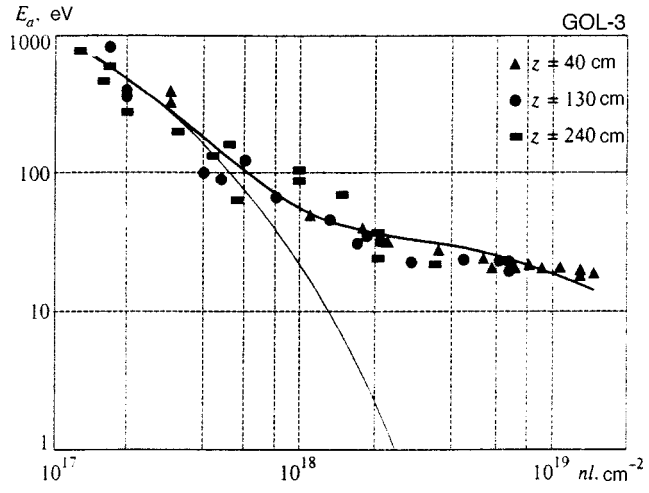

FIG. 8. Mean energy per hydrogen atom in a cloud, imparted by time 3.2 ms, vs the depth of the cloud, taking into account the ionization energy. The points correspond to measurements with the diamagnetic sensors, and the curves correspond to a calculation of the character of the energy release from an electron flux having a model distribution function (light curve—only Maxwellian electrons, heavy curve—supplemented by superthermal electrons).

with the parameters of the plasma flux that heats the cloud, in which the electron component can conventionally be divided into two parts: Maxwellian electrons with a temperature of about 1 keV, and superthermal electrons with a mean energy greater than 10 keV. The energy contribution from these groups of electrons is distributed over the depth $\langle nl \rangle$ of the bunch in accordance with their collision length. The following model was used to calculate the energy release. The beam passes through a region with a dense cold plasma without interacting and is then incident on a disperse plasma, which it heats up, losing as it does a significant part of its energy (under conditions close to optimum for interaction). The fast electrons that result from the interaction were assumed to be isotropic, and therefore about 15% of them escape to the exit mirror, while the remaining 85% are incident on the cloud and are thermalized there. We consider the conditions under which the fast electrons in the bunch are retarded to be constant during the entire pulse. Besides the fast electrons, the thermal electrons of the homogeneous plasma also contribute to the heating of the bunch because of longitudinal electronic thermal conductivity. The dependence shown in Fig. 8 relates to a beam with an initial energy of about 50 kJ (pulsewidth about 3.2 μs).

The same figure shows the calculated absorption curve of the plasma electrons, which have a temperature of 1 keV and an energy reserve of 2 kJ in the electronic component (the technique for calculating the absorption curve is given in Ref. 14). It can be seen that the curve of the observed energy release in the cloud cannot be explained except by retardation of the thermal electrons, especially at depths with $\langle nl \rangle > 10^{18}$ cm^{-2}.

The figure also shows the results of model calculations for a two-temperature distribution function of the fast plasma electrons (with components of 1 keV, 2 kJ and 10 keV, 5 kJ). As can be seen, besides the absorption of the hot electrons of the main component (with the density of the uniform plasma), absorption of a group of superthermal electrons with a characteristic energy of $E_f \sim 10$ keV is observed. A certain fraction of the hot electrons passes all the way through the cloud, depositing part of its energy, and then is absorbed in the input foil or is reflected from the magnetic mirror. Note that, from the results of the calculation, the contribution from fast electrons depends weakly on the form of their distribution function when the mean energy is conserved.

The density n_f of fast electrons with energy 10–20 keV can be obtained directly from the diamagnetic measurements (from the sharp pressure falloff in the interior of the bunch immediately after the beam ends). These estimates give $n_f \sim 10^{13}$ cm^{-3}.

Since there is virtually no heat dissipation from the interior of the cloud, the energy contained in the fast electrons is accumulated in the cloud during the beam injection τ_b. The relationship

$$n_f E_f L/\tau_f \sim \int n((3/2)T + E_i^a)dl/\tau_b ,$$

should then obviously be satisfied, where τ_f is the lifetime of the fast electrons in the trap. From this estimate, the fast electrons have lifetimes of $\tau_f \sim 0.14$ μs, which virtually coincides with the time of flight of the fast electrons through the trap.

7. ENERGY BALANCE

On the basis of the ''universal'' energy-release curve of the hot electrons in the bunch, obtained by processing a large number of experiments (Fig. 8), the resulting model function of the flux distribution of hot electrons can be used to determine the energy-deposition distribution over the length of the bunch in each specific shot. For example, we show the data for a shot corresponding to Fig. 5. The input energy of the beam in this case was 77 kJ. The total energy release in the plasma and the cloud by the end of the pulse is 5.6 ± 0.5 kJ. Of this energy, about 0.9 kJ is contained in the ''homogeneous'' plasma, with a density up to 10^{15} cm^{-3}, 3 ± 0.2 kJ is contained in the dense bunch, with $n > 10^{16}$ cm^{-3} (on the assumption that the mean energy of the fast electrons equals 10 keV), and the remaining energy is imparted to the intermediate-density plasma. The energy losses of the beam in this shot amount to 10–15 kJ; i.e, a significant fraction of the energy lost by the beam is transferred to the dense plasma. A part of the energy lost by the beam during the pulse escapes into the exit mirror, where there is no bunch of dense plasma.

8. OBTAINING A DENSE PLASMA BURST WITH HIGH TEMPERATURE

The main feature of the experiments discussed in this section is the use of a positive-action system of gas-pulsing, which made it possible to obtain a plasma bunch with a characteristic size of about 0.5 m at the time the beam was injected. The maximum density in the cloud was smaller than that in the experiments described in Sec. 5, and equalled $(0.5–2) \times 10^{16}$ cm^{-3}. This was done so that the plasma tem-

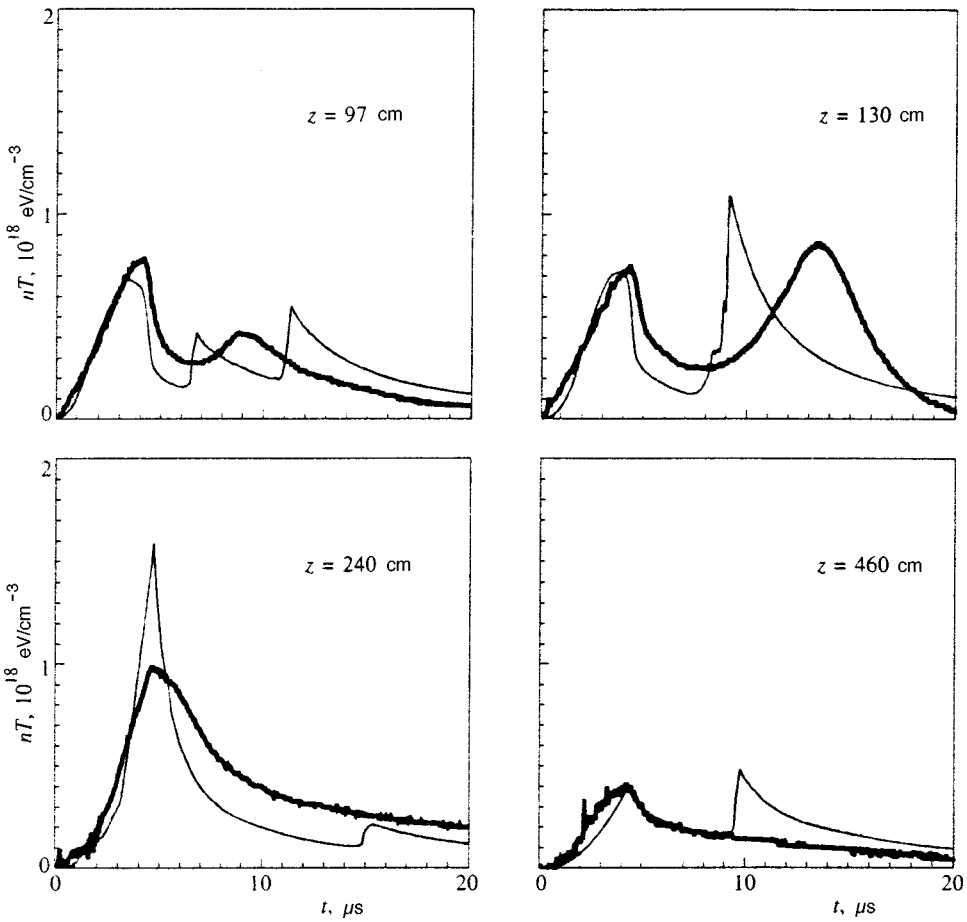

FIG. 9. Experimental (heavy curves) and calculated (light curves) values of the plasma pressure at various points of the system.

perature even in its densest part would significantly exceed the ionization energy of hydrogen. The site where the new gas-admission system was installed coincided with that of the Thomson scattering diagnostics, so that, unlike the preceding experiments, the plasma parameters could be measured in the densest central part of the cloud ($z=270$ cm). The plasma density in the homogeneous part was 5×10^{14} cm^{-3}.

Figure 9 shows typical signals from diamagnetic sensors located at various points of the plasma column. It can be seen that the plasma pressure close to the center of the bunch ($z=270$ cm) increases monotonically during the heating pulse and then gradually decreases. The two-stage heating effect consists of the fact that the amplitude of the plasma pressure in the cloud significantly exceeds (by a factor of 3–5) the value obtained at the same point when a homogeneous plasma of optimum density is heated.

In the experiments described in Sec. 5, the thickness $\langle nl\rangle$ of the cloud was about 10^{20} cm^{-2}, and therefore the center of the cloud was heated only by the superthermal electrons, while the Maxwellian electrons left their energy at the periphery of the cloud. In the experiments described here, the density of the cloud was chosen so that the thermal electrons of the homogeneous plasma heat a significant part of it. Therefore, by the end of the heating, the pressure distribution along the length of the system has one clearly expressed maximum at the site of formation of the bunch. The pressure of the plasma in its homogeneous part has a characteristic second maximum (see Fig. 9), which is associated with the arrival of a pressure wave at the recording point. The pressure in the wave in this case is comparable with the maximum pressure of the plasma at the time the beam is injected. In this regime, the density and temperature of the plasma were measured at the center of the bunch during the heating and a certain time after it ended. As a whole, the Thomson measurements agree with the diamagnetic measurements (allowing for the fact that the closest diamagnetic zone is 30 cm from the center of the cloud). A typical final temperature at the center of the bunch is 60–70 eV at a density of about 10^{16} cm^{-3} and up to 200 eV at a density of 7×10^{15} cm^{-3}. Note that the plasma density measured after the heating ends is less than the initial value, which can be explained by propagation of the short bunch along the magnetic field under the action of the pressure gradient. The plasma temperature obtained at the center of the bunch is several times as large as the ionization energy, which makes it possible to carry out a more accurate analysis of the energy balance, although the low value of $\langle nl\rangle$ makes it impossible to effectively confine a significant fraction of the superthermal electrons.

9. DYNAMICS OF A DENSE PLASMA BURST

After the gas cloud begins to be heated, the rate at which it expands changes substantially. Moreover, because the temperature and pressure are distributed over the depth of the

plasma column, a complex plasma-flow pattern is established in it, and pressure waves are formed. We shall consider the case with the pulsed-inlet valves located at $z=12$ cm (Figs. 5 and 6).

We turn our attention to the features of the shape of the signals from the diamagnetic sensors. Some time after the end of the beam, secondary maxima are observed on the diamagnetic signals. They are associated with the arrival of a pressure wave at the measurement point. The signals from the sensors of the visible, VUV, and soft x rays from the plasma show the same features. Broadening of the H_α line at the time of arrival of the wave is observed, which is evidence that the ion temperature and the plasma density increase at that instant (the spectral resolution of the diagnostic facility makes it possible to distinguish the Stark broadening of the wings of the line and the Doppler broadening of its core). The place of origin of these waves and their propagation velocity can be traced by varying the experimental conditions.

An analysis of the data shows that there are several waves in the dense plasma. A wave propagates from the region with a pressure peak both into the interior of the bunch (into increasing density) and into the region of homogeneous plasma (into decreasing density). Moreover, a pressure wave is formed close to the input foil (such a wave is also observed when the beam is injected into the homogeneous plasma[7]). The pressure in the oncoming wave can be twice as large as that of the plasma before the arrival of this wave (Figs. 6 and 9). The velocity of the wave depends on the parameters of the plasma in which it propagates. For example, as the density of the gas bunch varies, its temperature and consequently the wave velocity varies. As the density in the gas bunch decreases, the wave velocity increases.

10. MODELLING THE DYNAMICS OF THE DENSE PLASMA

The heating of the plasma by the beam and the processes of heat transfer and motion of the plasma were numerically modelled, using the hydrodynamic model described in detail in Ref. 19. The model uses phenomenological parameters that characterize these phenomena and are determined by comparing the calculations with the experimental results. The model was developed to study the heating and dynamics of the dense plasma bunchs of variable atomic composition in an inhomogeneous magnetic field, and some of its capabilities are therefore not used for the calculations carried out in this paper.

The starting equations

The dynamics of the plasma are described by the equation of motion and the continuity equation:

$$\frac{\partial n}{\partial t}+H\frac{\partial}{\partial s}\left(\frac{nV}{H}\right)=0, \tag{1}$$

$$\frac{\partial V}{\partial t}+\frac{H}{n}V\frac{\partial}{\partial s}\left(\frac{nV}{H}\right)+\frac{1}{Mn}\frac{\partial(nT+q)}{\partial s}=0. \tag{2}$$

Here n and V are the density and velocity of the plasma, M is the mass of the ions, s is the coordinate along a line of the magnetic field H, $T=T_e+T_i$,

$$q=-\lambda^2 nM\frac{\partial V}{\partial s}\left|\frac{\partial V}{\partial s}\right|$$

is the artificial viscosity introduced in the standard way to damp out the instability of the solution that arises when the wave breaks in the plasma, $\lambda\sim 3-10$ cm is the characteristic length of the wavefront at which the viscous addition becomes substantial, and t is time.

The heat distribution in the system is determined by the heat-balance equations for the electrons and ions:

$$\frac{3}{2}\frac{\partial nT_e}{\partial t}+H\frac{\partial}{\partial s}\left(\frac{3}{2}\frac{nT_eV}{H}\right)+nT_eH\frac{\partial}{\partial s}\left(\frac{V}{H}\right)=\frac{\partial}{\partial s}\left(\kappa_e\frac{\partial T_e}{\partial s}\right)+Q_e, \tag{3}$$

$$\frac{3}{2}\frac{\partial nT_i}{\partial t}+H\frac{\partial}{\partial s}\left(\frac{3}{2}\frac{nT_iV}{H}\right)+nT_iH\frac{\partial}{\partial s}\left(\frac{V}{H}\right)=\frac{\partial}{\partial s}\left(\kappa_i\frac{\partial T_i}{\partial s}\right)+Q_i. \tag{4}$$

Here T_e and T_i are the temperatures of the electrons and ions in the plasma, while the longitudinal heat conductivities are given by

$$\kappa_e=F_e(Z_{\text{eff}})\frac{nT_e\tau_e/\zeta}{m},\quad \kappa_i=F_i(Z_{\text{eff}})\frac{nT_i\tau_i}{M}, \tag{5}$$

where the collision times are defined as

$$\tau_e=\frac{3.5\times 10^4 T_e^{3/2}}{(\Lambda/10)Z_{\text{eff}}^2 n},\quad \tau_i=\frac{2.1\times 10^6 T_i^{3/2}}{(\Lambda/10)Z_{\text{eff}}^2 n}\sqrt{\frac{M}{M_p}}. \tag{6}$$

The temperature in Eq. (6) is expressed in electron volts, and the other quantities are in cgs units; M_p is the proton mass, the coefficient $F_e(Z_{\text{eff}})$ is taken from Ref. 20, $F_i(Z_{\text{eff}})=3.9$, and ζ is the flux-limiting coefficient of the electron thermal conductivity. This coefficient depends on the beam power and is determined by the turbulence level in the plasma; it is no greater than 100–1000.[7]

The quantities $Q_{e,i}$ on the right-hand sides of Eqs. (3) and (4) are represented as follows:

$$Q_e=\frac{P\eta_T}{SL}A\frac{\max(0.1\log[n_c/n(s)])}{(\Delta\vartheta^2+s/l_0)^{1/2}}+\frac{\partial E}{\partial t}-\frac{\partial}{\partial t}(E_{\text{ion}}-E_0)+n\nu_\varepsilon(T_i-T_e), \tag{7}$$

$$Q_i=n\nu_\varepsilon(T_e-T_i), \tag{8}$$

where

$$\nu_\varepsilon=3.2\times 10^{-9}\frac{nZ_{\text{eff}}^2\Lambda}{T_e^{3/2}}\sqrt{\frac{M_p}{M}}$$

is the characteristic frequency of the electron–ion energy exchange,

$$E_0=\frac{3}{2}\,n(T_{e0}+T_{i0})$$

is the initial energy of the plasma, $T_{e0}=T_{i0}=T_0$, and $E_{\rm ion}(n,T)$ is the ionization energy of hydrogen, calculated from the electron temperature (see Ref. 18). The first term in Eq. (7) describes the energy release of the beam in the plasma, $n_c=(3-5)\times10^{15}$ cm^{-3} is the limiting value of the plasma density, above which no turbulence develops, A is a normalizing factor, $P(t)$ is the beam power, S is the cross section of a force tube, η_T is the fraction of the beam energy that goes to heat the Maxwellian part of the plasma distribution function, l_0 is the relaxation length of the beam in the plasma, and $\Delta\vartheta\sim0.2$ is the initial angular spread of the electrons of the beam over velocities.

The quantity $\partial E/\partial t$ in Eq. (7) describes the variation of the energy density of the plasma because the tail of fast plasma electrons is retarded in it (see Ref. 14),

$$\frac{\partial E(\xi,T_h,t)}{\partial t}=\frac{P(t)\,\eta_h}{T_hS(\xi)}\int_{\varepsilon'_{\min}}^{\infty}\psi(\xi,T_h\varepsilon')\exp(-\varepsilon')\,d\varepsilon', \tag{9}$$

where $\varepsilon'=\varepsilon/T_h$ is the dimensionless energy of the superthermal electrons, η_h is the fraction of the beam energy that goes to heat the superthermal electrons, $\varepsilon'_{\min}$ is the energy at which mean free path R_0 of the electron in the plasma equals

$$\langle ns\rangle=\int n(s)\,ds,$$

$\xi=\langle ns\rangle/R_0$, and the electron absorption function in the plasma target is approximated by

$$\psi(\xi,\varepsilon)=3.39\varepsilon(\xi+0.01)^{0.25}\exp(-6\xi^{2.5}).$$

Initial conditions

The initial conditions correspond to spatially inhomogeneous filling of the system with hydrogen, which is partially ionized by a longitudinal discharge at the initial temperature $T_0\sim1$ eV. The initial density is given by

$$n(s)=n_0+n_d\exp(-(s-s_d)^2/l_d^2)+n_f\times\exp(-(s+s_f)^2/l_f^2). \tag{10}$$

The second term in Eq. (10) describes the gas cloud in the system, and the third term describes the plasma at the surface of the input foil, whose parameters are taken from the measurements of Ref. 21. The initial velocity of the plasma is assumed equal to zero.

Boundary conditions

The boundary conditions at the ends of the system have the form

$$T_1=T_2=T_0,\quad V_1=V_2=0,\quad \frac{\partial(nT_1)}{\partial s}=\frac{\partial(nT_2)}{\partial s}=0. \tag{11}$$

which corresponds to experimental conditions in the presence of an input foil and an exit calorimeter on whose surfaces a plasma appears whose temperature is relatively low but increases with time when a beam passes through.

Results of modelling, comparison with experiment

Our calculations use initial conditions close to the experiments described in Sec. 8. In this case, a cloud was formed at a distance of $z=270$ cm. The density of the main homogeneous plasma was 10^{15} cm^{-3}. The transport efficiency of the beam energy into the main component of the plasma was assumed in the calculation to be equal to 4% and that into the superthermal electrons was assumed to be 10%.

The calculated distribution of the plasma parameters over length is given in Fig. 10 for different times (the beam lasted 4.4 μs in this shot). At time 1 μs after the beginning of the beam, the density distribution over length virtually coincides with the initial distribution. The length of the bunch with a maximum density of 7×10^{15} cm^{-3} is about 30 cm. The temperature in the bunch during the action of the beam remains lower than in the homogeneous plasma. The pressure in the bunch is lower than in the surroundings at the beginning of heating because of the large expenditure to ionize the hydrogen, increases during the pulse more rapidly than in the homogeneous part of the plasma, and exceeds the pressure in the homogeneous plasma by a factor of 2–4 by the end of the beam.

An interesting feature of the process of heating a dense bunch located in a hot plasma is the ballistic compression of the dense plasma at a certain stage of the heating. In this case, pressure peaks appear at the edges of the bunch and move toward the center of the bunch. Under the experimental conditions in the GOL-3 facility, this effect can increase the density and pressure at the center of the bunch. In the case described here, the density increases by $\approx30\%$. Besides compression of the central part of the bunch, expansion of its periphery is observed at the same time during heating and subsequent cooling.

The pressure gradient of electrons (and ions) formed in the bunch during expansion causes a large-amplitude wave with a phase velocity of $\approx2\times10^7$ cm/sec to develop at times $t>7$ μs. The electron temperature at the wavefront is virtually constant over its length and equals $T_e\sim180-130$ eV at times $t\sim7-11$ μs, while the ion temperature increases from $T_i\sim11-17$ eV in front of the wavefront to 30–40 eV behind the wavefront. Since the ions satisfy $T_i\ll T_e\approx$ const, assuming a constant polytrope, $\gamma\sim1$,[19] we get the following calculated ion-sound velocity:

$$C_s=(\gamma(T_e+T_i)/m_i)^{0.5}\approx10^7\ \text{cm/sec}.$$

This means that a large-amplitude shock-induced ion-sound wave with a Mach number of $M\sim2$ is formed in the system. Because of the high electron thermal conductivity, the wave parameters are not described by the Hugoniot adiabat. The plasma density at the wavefront increases by a factor of ≈3, while the velocity of the ions exceeds the ion-sound velocity, so that the wave formation and propagation process rather corresponds to the expansion of a plasma into a vacuum and collective acceleration of the ions by electron pressure. Note

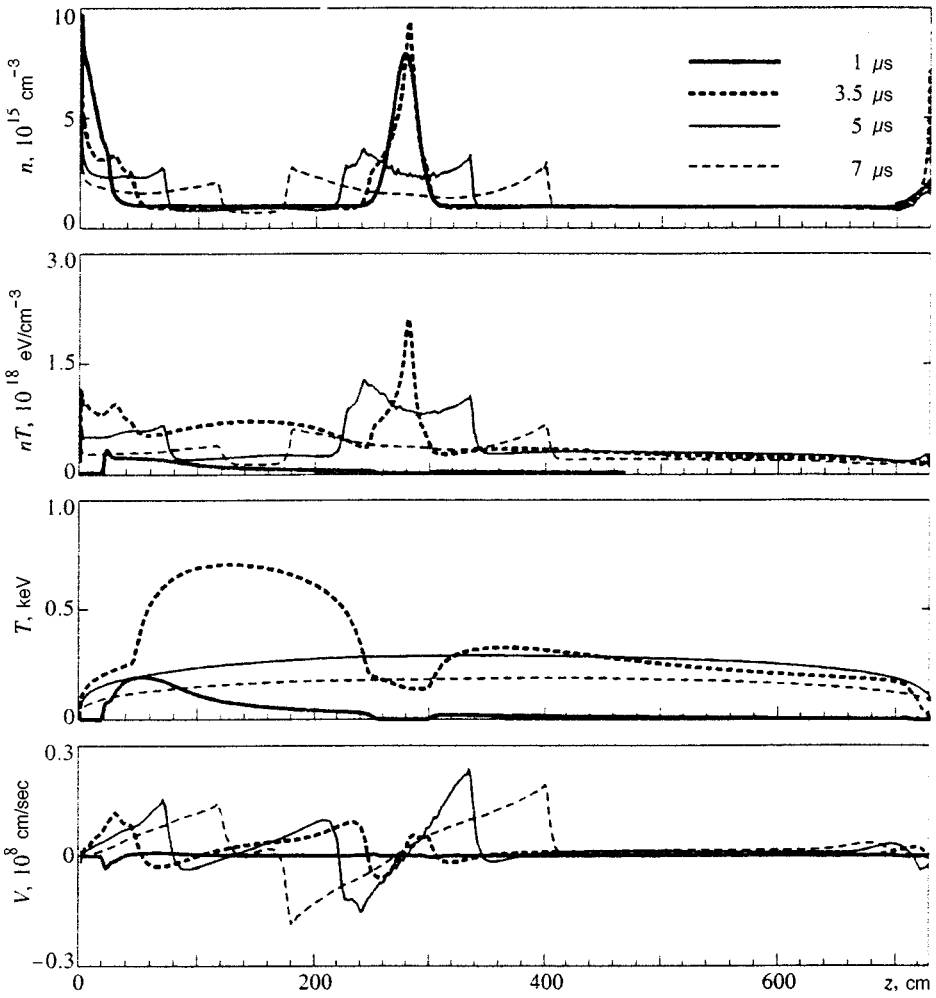

FIG. 10. Calculated distributions of the plasma characteristics along the length of the system at various times: n is the plasma density, nT is the pressure, T is the temperature of the electrons, and V is the velocity of the plasma.

that this mechanism of increasing the ion energy can be useful for heating a plasma, for example, when plasma streams formed in such a way collide.

A comparison of the calculated plasma pressure at various points of the system with the value measured by the diamagnetic sensors (Fig. 9) shows that the increase in time and the longitudinal pressure distribution at the stage of beam heating agree well with experiment, as does the falloff in temperature because the classical thermal conductivity is restored when the beam is turned off. However, the experimentally observed wave in this case is satisfactorily described by the model close to the cloud and damps out in amplitude significantly faster than in the model as one goes away from the cloud. Moreover, no steep pressure fronts are measured in the experiment. This is partially explained by the finite spatial resolution of the dynamic sensors, which smoothes out the signal in time from the moving wavefront. However, the wavefront damping is most probably associated with radial inhomogeneity of both the hot plasma and the dense bunch. As a result, the expansion rate of the bunch is different at different points over the cross section of the plasma. The faster damping of the wave in experiment is presumably associated with losses of the fast ions of the bunch in the charge transfer that occurs while the wave is moving.

Besides the compression wave associated with the expansion of the dense bunch, a pressure wave generated in the dense gas-plasma formation that appears during the explosive vaporization of the input foil (see Ref. 7) propagates through the plasma. A complex pattern of propagation and interference of pressure waves results from this. This is illustrated by Fig. 11, which shows the space–time dynamics of the plasma pressure. The propagation of waves from the input foil ($z=0$) and from the site at which the cloud is admitted ($z=270$ cm) can be seen. These pressure waves can intersect at a definite instant. A similar picture is observed in experiment.

11. THE CLOUD AT THE END OF THE APPARATUS: TWO CLOUDS

In order to more fully absorb the beam energy released in the main plasma, a series of experiments was carried out in which, in addition to the cloud at $z=12$ cm, a second assembly of pulsed valves was mounted on the facility at $z=575$ cm. Some of the experiments with this assembly were performed with the valves close to the entrance foil switched off. The general character of the energy release is similar to that observed in the cloud formed at the beginning of the facility. An exception is that there are virtually no pressure waves in the cloud. This is most probably because the region where the beam most intensely interacts with the plasma lies

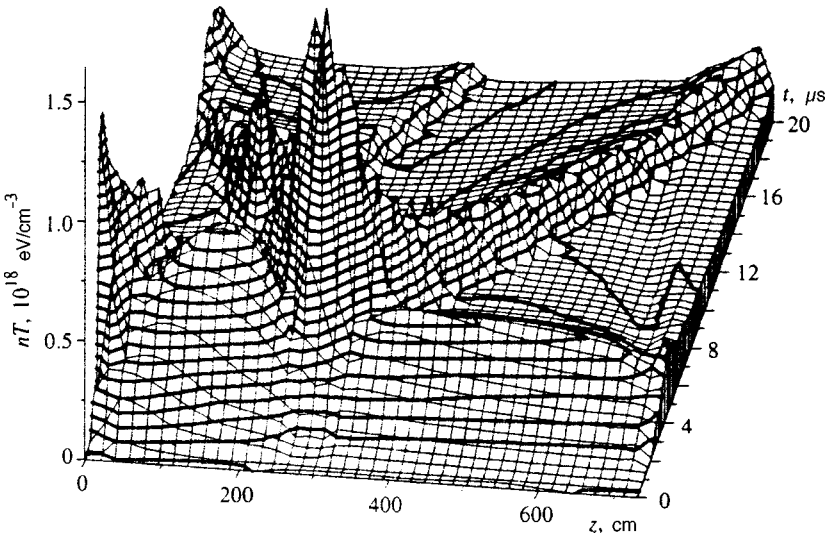

FIG. 11. Calculated dynamics of the plasma pressure for a short cloud.

close to the entrance foil, i.e., far from the boundary of the cloud. However, when the cloud is located at the input, the beam can begin to intensely relax even within the cloud in the region where the local plasma density exceeds the density at the homogeneous section only by a small factor. Moreover, the temperature of the main component of the plasma is higher at the input, and the specific energy contribution is accordingly greater at the boundary of the cloud. For this reason, when the cloud gets to the end of the plasma column, the distribution of the absorbed energy over the depth of the cloud changes appreciably, large pressure gradients do not appear, and the conditions for the formation of shock waves do not arise.

The conditions for two-stage heating of the plasma are completely met when the valves at both ends of the facility are actuated simultaneously. When this is done, a region with low density, in which the beam relaxes efficiently, is formed at the center of the plasma column. At the ends of the facility, this region is surrounded on two sides with dense plasma bunchs. The density profile of the gas and the distribution of the specific energy content (taking into account the ionization energy) over the length in this regime is shown in Fig. 12. At the time the beam ends, the maximum signals are observed on the sensors located in the region of the homogeneous plasma. The homogeneous plasma then transfers energy comparatively quickly to the plasma bunchs at the ends, which slowly cool. Because the length of the low-density plasma, in which relaxation of the beam occurs, was not very large in these experiments, the absolute parameters of the dense plasma are inferior to those shown above for other configurations.

12. CONCLUSION

The main results of the experiments on two-stage heating of a dense plasma in the GOL-3 facility are the following:

1. The feasibility in principle of a system for two-stage heating of a dense plasma has been shown experimentally for the first time. Energy was concentrated in a dense plasma cloud. The efficiency of energy transfer from the beam to an initial plasma with a density of 5×10^{14} cm^{-3} by collective interaction was about 30%. Approximately half of this energy was transferred into a bunch of dense plasma in this case.

2. The dense plasma is heated both by the flow of Maxwellian electrons with a temperature of a about 1 keV and by superthermal electrons with a characteristic energy of >10 keV.

3. Gas-dynamic expansion of the heated dense plasma bunch is observed, with development of waves having phase velocities close to that of ion-sound waves and identified in experiment as shock-induced ion-sound waves. These waves, on one hand, can serve as an additional source to heat the plasma in the depth of the dense bunch and, on the other

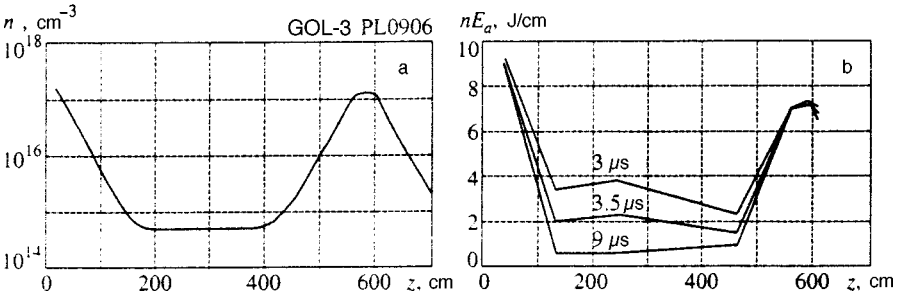

FIG. 12. Experiment with two bunchs of a dense plasma: (a) Distribution of the atomic density of hydrogen over the length of the facility with a delay of 0.5 ms in triggering the valves. (b) Distribution of the specific energy content of the plasma over the length of the facility (beam duration 3 μs).

hand, can produce more efficient heating of the ions in the region of the low-density plasma.

4. When a cloud with a characteristic length of 0.5 m is heated, a temperature of up to 75 eV at the center is attained by the end of the heating pulse, with a density of 2×10^{16} cm^{-3}. In this case, the energy consumed in heating the dense plasma is several times as great as that spent on ionization.

5. The dynamics of the expansion of the dense plasma bunch has been numerically investigated, and a hydrodynamic model has been developed to describe this process. Efficient pumping of energy into the longitudinal component of the velocity occurs in a number of regimes. The results of the numerical modelling are found to be in reasonable agreement with experiment.

6. The specific parameters of the dense plasma can be enhanced by increasing the energetics of the beam. This is planned for the second phase of the GOL-3 facility, along with an additional slowdown of the dispersal of the dense plasma (by using magnetic well configurations or multimirror geometry).

The authors are sincerely grateful to D. D. Ryutov and É. P. Krugyakov for interest in the work and to S. G. Voropaev, A. F. Guber, V. V. Konyukhov, S. V. Lebedev, A. G. Makarov, A. A. Nikiforov, S. S. Perin, and V. V. Chikunov for collaboration and help in preparing and performing the described experiments. This work was carried out with the financial support of the Ministry of Science of the Russian Federation, Grants RFFI No. 96-02-19436 and MNF No. NQF000 and No. NQF300.

*)E-mail: burdakov@inp.nsk.su

[1] G. I. Budker, V. V. Mirnov, and D. D. Ryutov, JETP Lett. **14**, 212 (1971).
[2] B. G. Logan, A. J. Lichtenberg, M. A. Lieberman, and A. Makijani, Phys. Rev. Lett. **28**, 144 (1972).
[3] D. D. Ryutov, VANT, Ser. Termoyad. Sintez Nos. 1–2, 96 (1978).
[4] A. V. Arzhannikov, A. V. Burdakov, V. S. Burmasov *et al.*, in *Proceedings of the Third International Conference on High-Power Electron and Ion-Beam Research and Technology, Novosibirsk*, 1979 (IYaF SO Akad. Nauk SSSR, Novosibirsk, 1979), vol. 1, p. 29.
[5] V. S. Burmasov, L. N. Vyacheslavov, I. V. Kandaurov *et al.*, VANT, Ser. Termoyad. Sintez No. 2, 31 (1987).
[6] V. Babicky, V. Bohaček, M. Člupek *et al.*, in *Proceedings of the Eighth International Conference on High-Power Particle Beams, Novosibirsk, 1990* (World Scientific, Singapore, 1991), vol. 1, p. 225.
[7] A. V. Burdakov, S. G. Voropaev, V. S. Koĭdan *et al.*, Zh. Éksp. Teor. Fiz. **109**, 2078 (1996) [JETP **82**, 1120 (1996)].
[8] A. V. Arzhannikov, B. N. Breizman, A. V. Burdakov *et al.*, in *Proceedings of the Seventh International Conference on Plasma Physics and Controlled Nuclear Fusion Research, Innsbruck*, 1978 (IAEA, Vienna, 1979), vol. 2, p. 623.
[9] A. V. Arzhannikov, A. V. Burdakov, O. I. Meshkov *et al.*, in *Proceedings of the International Conference on Plasma Physics, Lausanne, Switzerland*, 1984, Ed. M. Q. Tran and R. J. Verbeek, Invited Papers Com. of the European Community, Brussels, 1984, vol. 1, p. 285.
[10] B. N. Breĭzman, in *Reviews of Plasma Physics*, Ed. B. B. Kadomtsev (Énergoatomizdat, Moscow, 1987; Consultants Bureau, New York, 1990), Vol. 15, p. 55.
[11] A. V. Arzhannikov, A. V. Burdakov, V. A. Kapitonov *et al.*, Plasma Phys. Controlled Fusion **30**, 1571 (1988).
[12] A. V. Burdakov, V. S. Koĭdan, V. V. Postupaev, Fiz. Plazmy **11**, 944 (1985) [Sov. J. Plasma Phys. **11**, 548 (1985)].
[13] A. V. Burdakov, A. V. Karyukin, V. S. Koidan *et al.*, in *Proceedings of the Seventeenth European Conference on Controlled Fusion and Plasma Heating, Amsterdam*, 1990, Ed. G. Brifford, A. Nijsen-Vis, and F. C. Schuller, Europ. Phys. Soc. **14B**, 614 (1990).
[14] A. V. Burdakov, S. G. Voropaev, A. F. Guber *et al.*, Preprint IYaF SO RAN 92–12 (1992).
[15] G. E. Derevyankin, V. G. Dudnikov, P. A. Zhuravlev, Prib. Tekh. Eksp. **5**, 168 (1975).
[16] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Nauka, Moscow, 1988; Pergamon Press, Oxford, 1987).
[17] K. P. Stanyukovich, *Unsteady Motion of Continuous Media* (Nauka, Moscow, 1971; Plenum, New York, 1960).
[18] A. V. Nedospasov and M. Z. Tokar', in *Reviews of Plasma Physics*, Ed. B. B. Kadomtsev (Énergoatomizdat, Moscow, 1990; Consultants Bureau, New York, 1993), Vol. 18, p. 83.
[19] V. T. Astrelin and A. V. Burdakov, Preprint IYaF SO RAN 94–91 (1991).
[20] S. N. Braginskiĭ, in *Reviews of Plasma Physics*, Ed. M. A. Leontovich (Gosatomizdat, Moscow, 1963; Consultants Bureau, New York, 1963), Vol. 1, p. 183.
[21] A. V. Burdakov, V. V. Filippov, V. S. Koidan *et al.*, J. Nucl. Mater. **212–215**, 1345 (1994).

Translated by W. J. Manthey

Theory of a weakly nonideal Bose gas in a magnetic field

A. I. Akhiezer, S. V. Peletminskiĭ, and Yu. V. Slyusarenko*)

Kharkov Physicotechnical Institute, National Scientific Center, 310108 Kharkov, Ukraine
(Submitted 14 August 1997)
Zh. Éksp. Teor. Fiz. **113**, 918–929 (March 1998)

This paper investigates Bose–Einstein condensation of an ideal gas of finite-spin bosons in an external magnetic field. We generalize Bogolyubov's theory of a weakly nonideal Bose gas to the case where the gas of finite-spin bosons is located in an external magnetic field. We find the corresponding quasiparticle spectrum and formulate the superfluidity criterion for the boson gas. The magnetization of the weakly nonideal Bose gas is also determined. Finally, we specify a method of studying kinetic processes that take place in a weakly nonideal Bose gas. © *1998 American Institute of Physics.* [S1063-7761(98)01103-2]

1. INTRODUCTION

Lately intricate experimental investigations have been carried out[1] that prove the existence of a remarkable phenomenon, Bose–Einstein condensation, predicted by Albert Einstein in 1925. The thoroughness and precision of the experiments, conducted at temperatures as low as 2×10^{-8} K, suggest that it is possible to study the effect of an external magnetic field on Bose–Einstein condensation, which should exhibit well-defined features. In this connection it would be interesting to study the phenomenon of Bose–Einstein condensation in the case where the bosons have finite spins and an external magnetic field is applied to the system.

In this paper we consider the problem for an ideal Bose gas and for a nonideal Bose gas. Note that if the bosons carry an electric charge, we must allow not only for the paramagnetism of the gas but also for the Landau diamagnetism, which is a reflection of quantization of motion in a magnetic field. However, in an ideal Bose gas, quantization of motion makes Bose–Einstein condensation in the presence a magnetic field impossible (see below). Hence we assume from here on that the particles are electrically neutral, and allow only for the paramagnetism of the atoms.

We start with Bose–Einstein condensation for an ideal gas in a constant magnetic field. Then we introduce a weakly nonideal Bose gas of finite-spin particles in a magnetic field and generalize Bogolyubov's results[2] that refer to a weakly nonideal Bose gas of spin-zero particles.

2. BOSE–EINSTEIN CONDENSATION FOR AN IDEAL GAS IN A MAGNETIC FIELD

We begin our investigation of Bose–Einstein condensation for an ideal gas in a magnetic field, a gas consisting of electrically neutral spin-S particles, by introducing the boson distribution function

$$N_{S_{ZP}}(\beta,\zeta)=\{\exp[\beta(\varepsilon_p-\mu HS_Z-\zeta)]-1\}^{-1}, \tag{2.1}$$

where $\varepsilon_p=p^2/2m$ is the kinetic energy of a particle with momentum $\mathbf{p}$, S_Z is the projection of the particle spin in the direction of the magnetic field $\mathbf{H}$, $S\mu>0$ is the magnetic moment, β is the reciprocal temperature, and ζ is the chemical potential. Since the distribution function is positive, $\varepsilon_p-\mu HS_Z-\zeta>0$ is always positive, and the chemical potential ζ can be determined from the expression (cf. Ref. 3)

$$n=\frac{N}{\mathscr{V}}=\frac{1}{\mathscr{V}}\sum_{S_Z=-S}^{S}\sum_p\{\exp[\beta(\varepsilon_p-\mu HS_Z-\zeta)]-1\}^{-1}, \tag{2.2}$$

where n is the gas density. As the temperature decreases, each term in (2.2) decreases. Since n is a fixed quantity and ζ is always smaller than $-\mu HS$ ($\zeta<-\mu HS$), ζ increases as the temperature drops.

The Bose–Einstein condensation temperature T_0 is defined as the temperature at which ζ reaches $-\mu HS$ ($\zeta=-\mu HS$), and can by found by solving the equation

$$n=\frac{1}{\mathscr{V}}\sum_{S_Z=-S}^{S}\sum_p\{\exp[\beta_0(\varepsilon_p-\mu H(S_Z-S))]-1\}^{-1}. \tag{2.3}$$

When $T<T_0$, the distribution function of the above-condensate particles is given by

$$N_{S_{ZP}}(\beta,-\mu HS)=\{\exp[\beta(\varepsilon_p-\mu H(S_Z-S))]-1\}^{-1}, \tag{2.4}$$

with the particles with the largest projection of spin S forming the condensate. The condensate particle density n_0 is

$$n_0=n-\frac{1}{\mathscr{V}}\sum_{S_Z=-S}^{S}\sum_p\{\exp[\beta(\varepsilon_p-\mu H(S_Z-S))]-1\}^{-1}. \tag{2.5}$$

We also give the formulas for the magnetization of the boson gas:

$$m=\begin{cases}\dfrac{1}{\mathscr{V}}\mu\displaystyle\sum_{S_Z=-S}^{S}S_Z\sum_p\left\{\exp\left[\beta\left((\varepsilon_p-\mu HS_Z-\zeta\right)\right]-1\right\}^{-1}, & T>T_0,\\ \mu n_0S+\dfrac{1}{\mathscr{V}}\mu\displaystyle\sum_{S_Z=-S}^{S}S_Z\sum_p\left\{\exp\left[\beta\left(\varepsilon_p-\mu H(S_Z-S)\right)\right]-1\right\}^{-1}, & T<T_0,\end{cases} \tag{2.6}$$

Next we study two limits, $\mu H\ll T$ and $\mu H\gg T$ ($T<T_0$). In the first limit we are dealing with ordinary Bose condensation (with small corrections ignored), and the magnetization is given by

$$m=\mu n_0S. \tag{2.7}$$

In the second limit the magnetization is

$$m=\mu n_0S+\mu S\frac{1}{\mathscr{V}}\sum_p\{\exp(\beta\varepsilon_p)-1\}^{-1},$$

or, with allowance for (2.3),

$$m=\mu nS. \tag{2.8}$$

Thus, in weak magnetic fields the magnetization is determined by the density of just the condensate particles, while in strong magnetic fields the magnetization is determined by the total particle density.

Now we write the expressions for the transition temperature T_0 in weak and strong magnetic fields. In weak magnetic fields, the transition temperature is determined by the usual formula

$$T_0=\frac{1}{\mathscr{V}}\left\{\frac{4\pi^2n\hbar^3}{(2S+1)\Gamma(3/2)\tilde{\zeta}(3/2)}\right\}^{2/3},\quad T_0\gg\mu H, \tag{2.9}$$

where $\Gamma(x)$ and $\tilde{\zeta}(x)$ are the gamma function and Riemann zeta function, respectively. In strong magnetic fields we have

$$T_0=\frac{1}{\mathscr{V}}\left\{\frac{4\pi^2n\hbar^3}{\Gamma(3/2)\tilde{\zeta}(3/2)}\right\}^{2/3},\quad T_0\ll\mu H. \tag{2.10}$$

Clearly, the Bose condensation temperature in strong magnetic fields is higher than it is in the absence of a field.

If the bosons carry an electric charge, we must allow for Landau quantization of the boson motion across the magnetic field. The particle density in this case is

$$n=\frac{eH}{c(2\pi\hbar)^2}\sum_{k=0}^{\infty}\sum_{S_Z=-S}^{S}\int dp_z\left\{\exp\left[\beta\left(\frac{p_z^2}{2m}+\left(k+\frac{1}{2}\right)\omega_H-\mu HS_Z-\zeta\right)\right]-1\right\}^{-1},$$

with $\omega_H=eH/2mc$ the Larmor frequency. Since the term in the sum corresponding to $k=0$, with $\zeta=\mu HS+\omega_H/2$, is divergent for small values of p_z, there can be no Bose condensation in this case. In other words, for Bose condensation in an ideal gas in the presence of a magnetic field to occur, the bosons must be electrically neutral (see Ref. 4).

3. PRECIPITATION OF THE CONDENSATE AND THE EFFECTIVE HAMILTONIAN OF A WEAKLY NONIDEAL BOSE GAS IN A MAGNETIC FIELD

Now let us study the behavior of a nonideal Bose gas in a magnetic field. The kinetic energy of the bosons and their interaction with the magnetic field **H** are governed, in the second-quantization scheme, by the formula

$$\mathscr{H}_0=\frac{1}{2m}\sum_\alpha\int d^3r\nabla\psi_\alpha^\dagger(\mathbf{r})\nabla\psi_\alpha(\mathbf{r})-\mu\mathbf{H}\cdot\sum_{\alpha,\beta}\int d^3r\psi_\alpha^\dagger(\mathbf{r})\mathbf{S}_{\alpha\beta}\psi_\beta(\mathbf{r}), \tag{3.1}$$

where $\mathbf{S}_{\alpha\beta}$ is the boson spin matrix, and $\psi_\alpha^\dagger(\mathbf{r})$ and $\psi_\alpha(\mathbf{r})$ are the boson creation and annihilation operators at point **r** with spin-projection α, which are related to the operators $a_{p\alpha}^\dagger$ and $a_{p\alpha}$ of boson creation and annihilation with momentum **p** and spin-projection α:

$$\psi_\alpha^\dagger(\mathbf{r})=\mathscr{V}^{-1/2}\sum_p a_{p\alpha}^\dagger\exp(-i\mathbf{p}\cdot\mathbf{r}),$$

$$\psi_\alpha(\mathbf{r})=\mathscr{V}^{-1/2}\sum_p a_{p\alpha}\exp(i\mathbf{p}\cdot\mathbf{r}). \tag{3.2}$$

In terms of $a_{p\alpha}^\dagger$ and $a_{p\alpha}$, the operator $\mathscr{H}_0$ assumes the form

$$\mathscr{H}_0=\sum_\alpha\sum_p\varepsilon_\alpha(\mathbf{p})a_{p\alpha}^\dagger a_{p\alpha},\quad \varepsilon_\alpha(\mathbf{p})=\frac{p^2}{2m}-\mu H\alpha. \tag{3.3}$$

We must also allow for the boson–boson interaction. We begin with the expression for the interaction energy of two bosons,

$$V=V_1(\mathbf{r}_1-\mathbf{r}_2)+V_2(\mathbf{r}_1-\mathbf{r}_2)\mathbf{S}_1\cdot\mathbf{S}_2+V_3(\mathbf{r}_1-\mathbf{r}_2)(\mathbf{n}\cdot\mathbf{S}_1)(\mathbf{n}\cdot\mathbf{S}_2), \tag{3.4}$$

with $\mathbf{r}_1$ and $\mathbf{r}_2$ the radius vectors of the two bosons, $\mathbf{S}_1$ and $\mathbf{S}_2$ the boson spins, and $\mathbf{n}=(\mathbf{r}_1-\mathbf{r}_2)/|\mathbf{r}_1-\mathbf{r}_2|$. The functions $V_1(\mathbf{r}_1-\mathbf{r}_2)$, $V_2(\mathbf{r}_1-\mathbf{r}_2)$, and $V_3(\mathbf{r}_1-\mathbf{r}_2)$ depend only on $|\mathbf{r}_1-\mathbf{r}_2|$. The first term on the right-hand side of Eq. (3.4) represents the usual potential interaction between the bosons. The second term describes the spin–spin exchange interaction. Finally, the third term is a reflection of possible tensor forces. Magnetic-dipole interaction is also described by the Hamiltonian (3.4).

In the second-quantization scheme, the interaction Hamiltonian of a Bose gas is

$$V=\frac{1}{2}\sum_{\alpha_1\cdots\alpha_4}\int d^3r_1\int d^3r_2\psi_{\alpha_1}^\dagger(\mathbf{r}_1)\Psi_{\alpha_2}^\dagger(\mathbf{r}_2)\times V_{\alpha_1\cdots\alpha_4}(\mathbf{r}_1-\mathbf{r}_2)\Psi_{\alpha_3}(\mathbf{r}_2)\psi_{\alpha_4}(\mathbf{r}_1), \tag{3.5}$$

where $V_{\alpha_1\cdots\alpha_4}(\mathbf{r}_1-\mathbf{r}_2)$ is given by (3.4). In terms of the operators $a_{p\alpha}^\dagger$ and $a_{p\alpha}$, this Hamiltonian can be written

$$V=\frac{1}{2\mathscr{V}}\sum_{p_1\cdots p_4}\sum_{\alpha_1\cdots\alpha_4}\Phi_{\alpha_1\cdots\alpha_4}(\mathbf{p}_1,\mathbf{p}_4)\times\delta_{p_1+p_2;p_3+p_4}a_{p_1\alpha_1}^\dagger a_{p_2\alpha_2}^\dagger a_{p_3\alpha_3}a_{p_4\alpha_4}, \tag{3.6}$$

where

$$\Phi_{\alpha_1\cdots\alpha_4}(\mathbf{p}_1,\mathbf{p}_4)=U_1(\mathbf{p}_1-\mathbf{p}_4)\delta_{\alpha_1\alpha_4}\delta_{\alpha_2\alpha_3}+U_2(\mathbf{p}_1-\mathbf{p}_4)S^i_{\alpha_1\alpha_4}S^i_{\alpha_2\alpha_3}+U_4(\mathbf{p}_1-\mathbf{p}_4)\times\frac{(\mathbf{p}_1-\mathbf{p}_4)_i(\mathbf{p}_1-\mathbf{p}_4)_j}{(\mathbf{p}_1-\mathbf{p}_4)^2}S^i_{\alpha_1\alpha_4}S^j_{\alpha_2\alpha_3}, \quad (3.7)$$

with the functions $U_l(\mathbf{p})$, $l=1,2,3$, related to the Fourier transforms

$$V_l(\mathbf{p})=\int d^3r V_l(\mathbf{r})\exp(-i\mathbf{p}\cdot\mathbf{r}) \quad (3.8)$$

of the $V_l(\mathbf{r})$ as follows:

$$U_1(\mathbf{p})=V_1(\mathbf{p}),$$

$$U_2(\mathbf{p})=V_2(\mathbf{p})-\frac{1}{2}\left\{\frac{\partial^2V_3(\mathbf{p})}{\partial p_l\partial p_l}-\frac{p_ip_j}{p^2}\frac{\partial^2V_3(\mathbf{p})}{\partial p_i\partial p_j}\right\}, \quad (3.9)$$

$$U_3(\mathbf{p})=\frac{1}{2}\frac{\partial^2V_3(\mathbf{p})}{\partial p_l\partial p_l}-\frac{3}{2}\frac{p_ip_j}{p^2}\frac{\partial^2V_3(\mathbf{p})}{\partial p_i\partial p_j}.$$

The total Bose-gas Hamiltonian $\mathscr{H}$ is given by (see Eqs. (3.3) and (3.6)–(3.9)

$$\mathscr{H}=\mathscr{H}_0+V, \quad (3.10)$$

which contains quadratic terms and fourth-order terms in the boson creation and annihilation operators.

The equilibrium properties of a Bose gas are described by the Gibbs distribution

$$w=\exp[\Omega-\beta(\mathscr{H}-\zeta N)], \quad (3.11)$$

where Ω is the thermodynamic potential, N is the particle number operator, ζ is the chemical potential, and β is the reciprocal temperature.

According to Bogolyubov,[2] in the operator $\mathscr{H}-\zeta N$ we must isolate the terms containing the creation and annihilation operators of bosons with minimum energy, i.e., bosons with zero momentum and projection of spin S. These operators, $a^\dagger_{0S}$ and a_{0S}, correspond to a macroscopically large occupation number n_{0S} (a fact that is a reflection of Bose condensation) and hence can be replaced by C-numbers:

$$a^\dagger_{0S}\to\sqrt{n_{0S}}, \quad a_{0S}\to\sqrt{n_{0S}}, \quad (3.12)$$

where n_{0S} is the macroscopically large number of condensate particles. As a result, the operator $\mathscr{H}-\zeta N$ becomes

$$\mathscr{H}-\zeta N\equiv H=H_2+H_3+H_4, \quad (3.13)$$

where the Hamiltonian H_2 is quadratic in the boson creation and annihilation operators:

$$H_2(n_{0S})=f(n_{0S},H)+\frac{\partial f(n_{0S},H)}{\partial n_{0S}}\left\{\sum_{p\neq0}a^\dagger_{pS}a_{pS}+\sum_{\alpha\neq S}\sum_p\alpha^\dagger_{p\alpha}a_{p\alpha}\right\}+\sum_\alpha\sum_p\varepsilon(p)a^\dagger_{p\alpha}a_{p\alpha}+\left\{\frac{n_{0S}}{S\mathscr{V}}[U_S(0)-U_1(0)]-\mu H\right\}\times\sum_\alpha(\alpha-S)\sum_p a^\dagger_{p\alpha}a_{p\alpha}+S\frac{n_{0S}}{\mathscr{V}}\times\sum_{p\neq0}\left\{U_2(\mathbf{p})+\frac{1}{2}U_3(\mathbf{p})\sin^2\vartheta\right\}a^\dagger_{pS-1}a_{pS-1}+\sum_{p\neq0}\kappa(\mathbf{p})\{a^\dagger_{pS}a^\dagger_{-pS}+a_{pS}a_{-pS}+2a^\dagger_{pS}a_{pS}\}+\sum_{p\neq0}\{[t^{(+)}(\mathbf{p})(a^\dagger_{pS-1}a^\dagger_{-pS}+a^\dagger_{-pS}+a^\dagger_{pS}a^\dagger_{-pS-1}+2a^\dagger_{pS-1}a_{pS})]+\text{H.c.}\}+\sum_{p\neq0}\{q^{(+)}(\mathbf{p})a^\dagger_{pS-1}a^\dagger_{-pS-1}+\text{H.c.}\}. \quad (3.14)$$

Here, to simplify subsequent calculations, we introduce the notation (see (3.8) and (3.9))

$$f(n_{0S},H)\equiv\frac{1}{2}\frac{U_S(0)}{\mathscr{V}}n^2_{0S}-\mu SHn_{0S},$$

$$U_S(0)\equiv U_1(0)+S^2\left[U_2(0)+\frac{1}{3}U_3(0)\right],$$

$$\kappa(\mathbf{p})\equiv\frac{1}{2}\frac{n_{0S}}{\mathscr{V}}\{U_1(p)+S^2[U_2(p)+U_3(p)\cos^2\vartheta]\}, \quad (3.15)$$

$$t^{(\pm)}(p)\equiv\frac{S\sqrt{2S}}{4}\frac{n_{0S}}{\mathscr{V}}\frac{p_z(p_x\pm ip_y)}{p^2}U_3(p),$$

$$q^{(\pm)}(p)\equiv\frac{S}{4}\frac{n_{0S}}{\mathscr{V}}\frac{(p_x\pm ip_y)^2}{p^2}U_3(p),$$

where ϑ is the angle between the momentum $\mathbf{p}$ and the magnetic field $\mathbf{H}$ (in deriving (3.14) and (3.15) we used the explicit form of the spin matrices $S^i_{\alpha\beta}$ and assumed that the z axis is directed along $\mathbf{H}$).

The Hamiltonians H_3 and H_4 in (3.12) are, respectively, of third and fourth order in the boson creation and annihilation operators. The Hamiltonian H_2 describes a gas of free quasiparticles and is known as the effective Hamiltonian of a weakly nonideal Bose gas, while the Hamiltonians H_3 and H_4 describe the interaction between quasiparticles. Allowing for H_3 and H_4 is important when one studies various kinetic processes in a Bose gas. Studies of such processes lie outside the scope of the present paper, however, so that we will not write the expressions for the operators H_3 and H_4 explicitly.

The quantity n_{0S} introduced in (3.12) is a well-defined function of temperature and chemical potential. However, in the Gibbs distribution with a specified condensate,

$$w(\beta,\zeta,n_{0S})=\exp\{\Omega(\beta,\zeta,n_{0S})-\beta H(n_{0S})\}, \quad (3.16)$$

in which the thermodynamic potential Ω can be found from the normalization condition

$$\Omega(\beta,\zeta,n_{0S})=\ln\,\mathrm{Tr}\,\exp[-\beta H(n_{0S})], \quad (3.17)$$

n_{0S} is a free parameter. Actually this quantity should be found by minimizing the thermodynamic potential Ω at fixed β and ζ (see Ref. 5):

$$\left\{\frac{\partial\Omega}{\partial n_{0S}}\right\}_{\beta,\zeta}=0. \tag{3.18}$$

In the adopted approximation, where almost all particles are in the condensate (low temperatures), according to (3.14) and (3.15) this condition becomes

$$\zeta\approx\frac{\partial f(n_{0S},H)}{\partial n_{0S}}=U_S(0)\frac{n_{0S}}{\mathscr{V}}-\mu SH. \tag{3.19}$$

Now let us explain how the concept of a quasiparticle is introduced. According to Ref. 2, we must diagonalize H_2. This is achieved by introducing a unitary transformation U that ''mixes'' the operators a_{pS}, $a^{\dagger}_{-pS}$, a_{pS-1}, and $a^{\dagger}_{-pS-1}$:

$$\begin{aligned}Ua_{pS}U^{\dagger}&=\phi_1(p)a_{pS}+\phi_2^*(p)a^{\dagger}_{-pS}+\psi_1(p)a_{pS-1}\\&\quad+\psi_2^*(p)a^{\dagger}_{-pS-1},\\Ua_{pS-1}U^{\dagger}&=\phi_3(p)a_{pS}+\phi_4^*(p)a^{\dagger}_{-pS}+\psi_3(p)a_{pS-1}\\&\quad+\psi_4^*(p)a^{\dagger}_{-pS-1}.\end{aligned} \tag{3.20}$$

For this transformation to be unitary, i.e., for it to preserve the commutation relations between the boson creation and annihilation operators, the amplitudes ϕ_i and ψ_i $(i=1,2,3,4)$ must satisfy the relations

$$\begin{aligned}&|\phi_1|^2-|\phi_2|^2+|\psi_1|^2-|\psi_2|^2=1,\\&|\phi_3|^2-|\phi_4|^2+|\psi_3|^2-|\psi_4|^2=1\end{aligned} \tag{3.21}$$

and relations of the type

$$\phi_1\phi_3^*-\phi_2\phi_4^*+\psi_1\psi_3^*-\psi_2\psi_4^*=0. \tag{3.22}$$

The unitary transformation (3.20) is a generalization of Bogolyubov's unitary transformation. In addition to (3.21) and (3.22), to determine the amplitudes of the unitary transformation we need H_2 be to a diagonal operator, i.e.,

$$UH_2U^{\dagger}\equiv H_{2q}=\sum_{\alpha}\sum_{p\neq 0}\omega_{p\alpha}a^{\dagger}_{p\alpha}a_{p\alpha}+E_0, \tag{3.23}$$

where the $\omega_{p\alpha}$ comprise the energy spectrum of quasiparticles (E_0 is the ''energy'' of the ground state of the operator $H-\zeta N$). Equations (3.21)–(3.23) provide a complete solution for the energy spectrum of quasiparticles and the amplitudes ϕ_i and ψ_i, which relate particle states and quasiparticle states.

4. SPECTRA OF QUASIPARTICLES OF A WEAKLY NONIDEAL BOSE GAS IN A MAGNETIC FIELD

Now we deal directly with the diagonalization of the Hamiltonian H_2 (see (3.13)–(3.15)).

According to (3.23),

$$[H_{2q},a^{\dagger}_{pS}]=\omega_{pS}a^{\dagger}_{pS},\quad [H_{2q},a^{\dagger}_{pS-1}]=\omega_{pS-1}a^{\dagger}_{pS-1}. \tag{4.1}$$

Allowing for the fact that $H_2=U^{\dagger}H_{2q}U$, we obtain

$$\begin{aligned}&[H_{2q},Ua^{\dagger}_{pS}U^{\dagger}]=U[H_2,a^{\dagger}_{pS}]U^{\dagger},\\&[H_{2q},Ua^{\dagger}_{pS-1}U^{\dagger}]=U[H_2,a^{\dagger}_{pS-1}]U^{\dagger}.\end{aligned} \tag{4.2}$$

Employing Eqs. (3.20) and calculating the commutators on the left-hand sides of Eqs. (4.2), we arrive at a system of homogeneous linear equations for the amplitudes $\phi_i(\pi)$,

$$\begin{aligned}&(\varepsilon+2\kappa-\omega_{pS})\phi_1+2\kappa\phi_2+2t^{(-)}\phi_3+2t^{(+)}\phi_4=0,\\&2\kappa\phi_1+(\varepsilon+2\kappa+\omega_{pS})\phi_2+2t^{(-)}\phi_3+2t^{(+)}\phi_4=0,\\&2t^{(+)}\phi_1+2t^{(+)}\phi_2+(\varepsilon+\beta-\omega_{pS})\phi_3+2q^{(+)}\phi_4=0,\\&2t^{(-)}\phi_1+2t^{(-)}\phi_2+2q^{(-)}\phi_3+(\varepsilon+\beta+\omega_{pS})\phi_4=0,\end{aligned} \tag{4.3}$$

and a system of homogeneous equations for the amplitudes $\psi_i(p)$,

$$\begin{aligned}&(\varepsilon+2\kappa-\omega_{pS-1})\psi_1+2\kappa\psi_2+2t^{(-)}\psi_3+2t^{(+)}\psi_4=0,\\&2\kappa\psi_1+(\varepsilon+2\kappa+\omega_{pS})\psi_2+2t^{(-)}\psi_3+2t^{(+)}\psi_4=0,\\&2t^{(+)}\psi_1+2t^{(+)}\psi_2+(\varepsilon+\beta-\omega_{pS})\psi_3+2q^{(+)}\psi_4=0,\\&2t^{(-)}\psi_1+2t^{(-)}\psi_2+2q^{(-)}\psi_3+(\varepsilon+\beta+\omega_{pS})\psi_4=0.\end{aligned} \tag{4.4}$$

Note that the coefficients in the second set of equations can be obtained from those in the first by replacing ω_{pS} with ω_{pS-1}. The quantities $\kappa(p)$, $t^{(\pm)}(p)$, and $q^{(\pm)}(p)$ in Eqs. (4.3) and (4.4) are determined by (3.15) and (3.9); we have also introduced the notation

$$\begin{aligned}\beta(p)=\mu H+S\frac{n_{0S}}{\mathscr{V}}\Bigg[&U_2(p)-U_2(0)\\&+\frac{1}{2}U_3(p)\sin^2\vartheta-\frac{1}{3}U_3(0)\Bigg].\end{aligned} \tag{4.5}$$

Since the determinants of the systems of equations (4.3) and (4.4) are identical, we actually have one biquadratic dispersion equation,

$$\omega_p^4-[\chi(p)+\alpha(p)]\omega_p^2+\alpha(p)\chi(p)-4\eta(p)=0, \tag{4.6}$$

where

$$\begin{aligned}&\alpha(p)=\varepsilon_p[\varepsilon_p+4\kappa(p)],\\&\chi(p)=[\varepsilon_p+\beta(p)]^2-4q^2(p),\\&\eta(p)=16q(p)l(p)\varepsilon_p[\varepsilon_p+\beta(p)-2q(p)],\end{aligned} \tag{4.7}$$

with

$$\begin{aligned}&q(p)=\frac{S}{4}\frac{n_{0S}}{\mathscr{V}}U_3(p)\sin^2\vartheta,\\&l(p)=\frac{S^2}{2}\frac{n_{0S}}{\mathscr{V}}U_3(p)\cos^2\vartheta.\end{aligned} \tag{4.8}$$

The dispersion equation (4.6) has two roots for ω_p^2. The first we identify with

$$\omega_{pS}^2=\tfrac{1}{2}[\chi(p)+\alpha(p)]-\tfrac{1}{2}\{[\chi(p)-\alpha(p)]^2+4\eta(p)\}^{1/2}, \tag{4.9}$$

while the second we identify with

$$\omega_{pS-1}^2=\tfrac{1}{2}\,[\chi(p)+\alpha(p)]+\tfrac{1}{2}\,\{[\chi(p)-\alpha(p)]^2+4\,\eta(p)\}^{1/2}. \tag{4.10}$$

Note that ω_{pS} is a nonactivation frequency (a Goldstone frequency, $\omega_{0S}=0$), while ω_{pS-1} is an activation frequency. The quantities ω_{pS}^2 and ω_{pS-1}^2 must be positive. This imposes constraints on the shape of the interaction potentials at low momenta and field intensities, i.e., $\omega_{pS}>0$ and $\omega_{pS-1}>0$ if

$$\mu H\geqslant\frac{1}{3}\,S\,\frac{n_{0S}}{\mathscr{V}}U_3(0),\quad U_3(0)\geqslant 0,$$

$$\mu H\geqslant S\frac{n_{0S}}{\mathscr{V}}\left[U_2(0)+\frac{1}{3}U_3(0)\right], \tag{4.11}$$

and the interaction amplitudes (see Eqs. (3.9)) satisfy the conditions

$$U_1(0)+S^2U_2(0)\geqslant S^2U_3(0),\quad U_3(0)\geqslant 0. \tag{4.12}$$

We see that in the adopted theory, we cannot pass to the limit $H\to 0$ and derive formulas similar to those in Bogolyubov's theory.

Solution of Eqs. (4.3) and (4.4) with allowance for the normalization conditions (3.21) leads to the following expressions for $\phi_i(p)$ and $\psi_i(p)$:

$$\phi_1(p)=\frac{\varepsilon_p+\omega_{pS}}{2\sqrt{\varepsilon_p\omega_{pS}}}\Delta_{S-1}(p),\quad \phi_2=\frac{\varepsilon_p-\omega_{pS}}{2\sqrt{\varepsilon_p\omega_{pS}}}\Delta_{S-1}(p),$$

$$\phi_3(p)=2\sqrt{\frac{\varepsilon_p}{\omega_{pS}}}\,\frac{t^{(+)}(p)\{2q(p)-[\varepsilon_p+\beta(p)+\omega_{pS}]\}}{\chi(p)-\omega_{pS}^2}\times\Delta_{S-1}(p), \tag{4.13}$$

$$\phi_4(p)=2\sqrt{\frac{\varepsilon_p}{\omega_{pS}}}\,\frac{t^{(-)}(p)\{2q(p)-[\varepsilon_p+\beta(p)-\omega_{pS}]\}}{\chi(p)-\omega_{pS}^2}\times\Delta_{S-1}(p),$$

$$\psi_1(p)=\frac{\varepsilon_p+\omega_{pS-1}}{2\sqrt{\varepsilon_p\omega_{pS-1}}}\Delta_S(p),\quad \psi_2=\frac{\varepsilon_p-\omega_{pS-1}}{2\sqrt{\varepsilon_p\omega_{pS-1}}}\Delta_S(p),$$

$$\psi_3(p)=2\sqrt{\frac{\varepsilon_p}{\omega_{pS-1}}}\,\frac{t^{(+)}(p)\{2q(p)-[\varepsilon_p+\beta(p)+\omega_{pS-1}]\}}{\chi(p)-\omega_{pS-1}^2}\times\Delta_S(p),$$

$$\psi_4(p)=2\sqrt{\frac{\varepsilon_p}{\omega_{pS-1}}}\,\frac{t^{(-)}(p)\{2q(p)-[\varepsilon_p+\beta(p)-\omega_{pS-1}]\}}{\chi(p)-\omega_{pS-1}^2}\times\Delta_S(p), \tag{4.14}$$

where

$$\Delta_S(p)\equiv\eta^{1/2}(p)\{\eta(p)+[\chi(p)-\omega_{pS}^2]^2\}^{-1/2},$$

$$\Delta_{S-1}=\Delta_S|_{S\to S-1}. \tag{4.15}$$

Using Eqs. (4.13) and (4.14), we can easily see that conditions like (3.22) are automatically satisfied.

We now return to the Gibbs distribution (3.16). According to (3.13), (3.20), and (3.23),

$$Uw(\beta,\zeta,n_{0S})U^\dagger\equiv w_q=\exp\{\Omega-\beta[H_{2q}+H_{3q}+H_{4q}]\}, \tag{4.16}$$

where H_{2q} is given by

$$H_{2q}=E_0+\omega_{pS}a_{pS}^\dagger a_{pS}+\omega_{pS-1}a_{pS_{-1}}^\dagger a_{pS-1}+\sum_{\alpha=-S}^{S-2}\omega_{p\alpha}a_{p\alpha}^\dagger a_{p\alpha}. \tag{4.17}$$

Here

$$\omega_{p\alpha}\Big|_{\substack{\alpha\neq S\\ \alpha\neq S-1}}\equiv\omega_{p\tilde{\alpha}}=\varepsilon_p+(\alpha-S)\left\{-\mu H+S\frac{n_{0S}}{\mathscr{V}}U_2(0)\right\} \tag{4.18}$$

(we do not write the expression for E_0). The operator H_{2q} is the Hamiltonian of an ideal gas of quasiparticles, while H_{3q} and H_{4q} describe triple and quadruple interactions between quasiparticles ($H_{3q}\equiv UH_3U^\dagger$ contains terms like aaa and $a^\dagger aa$ and $H_{4q}\equiv UH_4U^\dagger$, terms like $aaaa$, $a^\dagger aaa$, and $a^\dagger a^\dagger aa$).

Equation (4.17) shows that there are $2S+1$ species of quasiparticles, of which the branch ω_{pS} is a Goldstone branch (i.e., ω_{pS} vanishes at $p=0$) and the remaining $2S$ are activation branches. These branches constitute a modification of the energy spectrum of the particles of an ideal Bose gas in a magnetic field:

$$\omega_{p\alpha}=\varepsilon_p+\alpha\mu H,\quad \alpha=-S,\ldots,S. \tag{4.19}$$

The mean value of an arbitrary physical quantity b is

$$\mathrm{Tr}\ w\hat{b}=\mathrm{Tr}\ w_qU\hat{b}(n_{0S})U^\dagger, \tag{4.20}$$

where $\hat{b}(n_{0S})$ is the operator $\hat{b}$ in which the creation and annihilation operators of bosons with zero momentum and spin projection S are replaced by $\sqrt{n_{0S}}$ (see Eqs. (3.12)).

In particular, the diagonal elements of the one-particle density matrix are

$$\hat{b}\equiv\hat{f}_{\alpha\alpha}(p)=a_{p\alpha}^\dagger a_{p\alpha}\equiv\hat{f}_{p\alpha},$$

with the result that

$$\begin{aligned}f_{p\alpha}=n_{0S}\delta_{p0}\delta_{\alpha S}+N_{p\tilde{\alpha}}\delta_{\alpha\tilde{\alpha}}+\delta_{\alpha S}[|\phi_1|^2N_{pS}+|\phi_2|^2\\ \times(1+N_{-pS})+|\psi_1|^2N_{pS-1}+|\psi_2|^2(1+N_{-pS})]\\ +\delta_{\alpha S-1}[|\phi_3|^2N_{pS}+|\phi_4|^2(1+N_{-pS})\\ +|\psi_3|^2N_{pS-1}+|\psi_4|^2(1+N_{-pS})],\end{aligned} \tag{4.21}$$

where $\tilde{\alpha}=\{-S,-S+1,\ldots,S-2\}$, and $N_{p\alpha}$ is the quasiparticle distribution function:

$$N_{p\alpha}=\mathrm{Tr}\ w_qa_{p\alpha}^\dagger a_{p\alpha}. \tag{4.22}$$

We now examine the superfluidity of a weakly nonideal gas in a magnetic field. To this end we introduce the term $-\beta\mathbf{u}\cdot\mathbf{P}$ into the Gibbs distribution, where $\mathbf{u}$ is the velocity of the weakly nonideal Bose gas (the velocity of the normal component), and

$$\mathbf{P}=\sum_\alpha\sum_p\mathbf{p}a_{p\alpha}^\dagger a_{p\alpha}$$

is the system's momentum operator. It is easy to prove that $U\mathbf{P}U^{\dagger}=\mathbf{P}$. Hence the formulas for the distribution function are still expressions of type (4.21) in which the quasiparticle distribution functions $N_{p\alpha}$ ($\alpha=S,S-1,\widetilde{\alpha}$) are given by (see Eq. (4.22))

$$N_{p\alpha}=\{\exp[\beta(\omega_{p\alpha}-\mathbf{p}\cdot\mathbf{u})]-1\}^{-1},\quad \alpha=S,S-1,\widetilde{\alpha}. \tag{4.23}$$

These distribution functions must be positive. For low momenta, the distribution functions N_{pS-1} and $N_{p\widetilde{\alpha}}$ are sure to be positive because the activation energies ω_{pS-1} and $\omega_{p\widetilde{\alpha}}$ are as well (see Eqs. (4.9), (4.10), and (4.18)). However, the Goldstone-quasiparticle distribution function N_{pS} is positive for low momenta only if the velocity u is smaller than a certain critical value u_c, which is determined by requiring that ω_{pS}/p be at its minimum for small p:

$$u<u_c=\min_p \frac{\omega_{pS}}{p}; \tag{4.24}$$

this is Landau's superfluidity condition.

We also write the expression for the magnetization of a weakly nonideal Bose gas at $\mathbf{u}=0$. This magnetization is given by the diagonal part of the one-particle density matrix:

$$m=\mu\sum_{\alpha}\ \alpha\frac{1}{\mathscr{V}}\sum_{p}\ f_{p\alpha}. \tag{4.25}$$

Inserting this into (4.21), we obtain

$$\begin{aligned} m=\mu S\frac{n_{0S}}{\mathscr{V}}+\mu\sum_{\alpha=-S}^{S-2}\alpha\frac{1}{\mathscr{V}}\sum_{p}\ N_{p\alpha}\\ +\mu\frac{1}{\mathscr{V}}\sum_{p}\ \{N_{pS}[S|\phi_1|^2+(S-1)|\phi_3|^2]+(1+N_{pS})\\ \times[S|\phi_2|^2+(S-1)|\phi_4|^2]+N_{pS-1}[S|\psi_1|^2\\ +(S-1)|\psi_3|^2]+(1+N_{pS-1})[S|\psi_2|^2\\ +(S-1)|\psi_4|^2]\}, \end{aligned} \tag{4.26}$$

where the quasiparticle distribution functions $N_{p\alpha}$ are given by (4.23) with $\mathbf{u}=0$.

In conclusion, we note that the quasiparticle energy depends on the magnetic field strength and condensate density. If these parameters vary slowly with time, the emerging modulation of the energy spectrum can be interpreted as an external field in which the Bose gas has been placed. Because of such modulation, the system entropy $\mathscr{S}$ also changes. Using the method developed in Ref. 6, we can formulate the H theorem and calculate the dissipative function $T\dot{\mathscr{S}}$. To this end we must study the kinetic equation for the distribution function of quasiparticles in a magnetic field. That equation allows for the various quasiparticle interaction processes described by the Hamiltonians H_3 and H_4. Obviously, quadruple processes correspond to scattering, while triple processes correspond to the fusion of two quasiparticles into one and fission of one quasiparticle into two.

The H theorem implies that the dissipation function $T\dot{\mathscr{S}}$ is proportional to the square of the frequency ω of modulation of external parameters. The absorption (by the system) of the energy of a variable external field and of an acoustic field capable of propagating in the quasiparticle gas also varies according to this law. The law is valid as long as ωt is much less than unity, with τ the relaxation time, i.e., the time that it takes static equilibrium to set in in the quasiparticle gas. Here, however, we will not calculate $T\dot{\mathscr{S}}$ and will not establish its dependence on the temperature T of the gas.

Financial support for this work was provided by the Ukrainian Foundation for Basic Research. The authors would also like to express their gratitude to the International Soros Science Education Program (ISSEP).

*)E-mail: slusarenko@kipt.kharkov.ua

[1] M. H. Anderson, I. R. Ensher, M. R. Mattews *et al.*, Science **269**, 198 (1995).
[2] N. N. Bogolyubov, Izv. Akad. Nauk SSSR, Ser. Fiz. **11**, 77 (1947).
[3] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Part 2, 3rd ed., Pergamon Press, Oxford (1980).
[4] J. Daicic, N. E. Frankel, R. M. Gailis, and V. Kovalenko, Phys. Rep. **237**, 65 (1994).
[5] I. A. Akhiezer and S. V. Peletminskiĭ, *Methods of Statistical Physics*, Pergamon Press, Oxford (1980).
[6] A. I. Akhiezer, Zh. Éksp. Teor. Fiz. **8**, 1318 (1938).

Translated by Eugene Yankovsky

Localization of vibrational modes in a layered lattice with diagonal and off-diagonal disorder

A. P. Zhernov*)

Kurchatov Institute Russian Scientific Center, Institute of Superconductivity and Solid-State Physics, 123182 Moscow, Russia

E. P. Chulkin

Physicotechnical Institute, 426000 Izhevsk, Russia
(Submitted 24 April 1997)
Zh. Éksp. Teor. Fiz. **113**, 930–954 (March 1998)

This paper discusses the possibility of weak localization of the phonon modes in a nonideal harmonic crystal lattice whose separate layers are weakly coupled to each other. An expression is obtained for the diffusivity tensor. The role of inverse coherent scattering processes is studied. It is established that, when such processes occur under conditions of strong diagonal disorder in the region of relatively low frequencies, where the dispersion law of the phonon modes exhibits two-dimensional properties, substantial renormalization of the diffusivity can take place. The paper analyzes the situation that occurs under conditions of off-diagonal disorder when resonant-scattering impurity centers are present in the lattice. The question of the possible nature of the low-temperature plateau in the thermal conductivity of the complex crystals BSCCO and BSYCO is discussed. © *1998 American Institute of Physics.* [S1063-7761(98)01203-7]

1. INTRODUCTION

The problem of the propagation and scattering of quasiparticles in disordered media when multiple coherent-scattering processes are important has been widely discussed in the literature. Most of the interest here is caused by phenomena associated with coherent inverse scattering. This effect was predicted theoretically and was detected at the beginning of the 1980s for electrons in disordered metals. We emphasize that the localization problem for electrons has been quite widely studied, and interest in it has been maintained for a very long time (see, for example, Altshuler *et al.*'s monograph[1] and the reviews of Kramer and MacKinnon[2] and Olemskiĭ[3]).

The phenomenon of localization of electromagnetic waves propagating in disordered media has recently been proposed as an independent topic. See, for example, Kuz'min and Romanov's review[4] relative to the appearance of coherent phenomena accompanying the scattering of light. A discussion appeared comparatively recently in connection with the experiments of the Amsterdam group.[5] They investigated radiation transport velocity in samples with powdered TiO_2 in air, and a substantial difference was detected between the energy transport velocity v_E and the phase velocity. Similar results were obtained in Ref. 6 for the case of porous glasses. Concepts concerning the existence of randomly located resonant-scattering impurity centers in the medium have been used in a number of papers to explain these facts.[7,8]

As far as problems of weak localization of phonon modes and sound waves in disordered systems are concerned, the situation is as follows: Studies have been carried out for standard three-dimensional lattices. The number of papers is small. There is appreciable interest in systems with resonant-scattering impurity centers. The case of a harmonic crystal with isotopic defects has been considered in the approximation of a single-site, coherent potential under conditions of crossover splitting of the spectrum.[9,10] In our papers,[11,12] we analyzed the situation in the general case of off-diagonal disorder, taking into account the interaction of the impurities. Reference 13 discussed the question of the Ioffe–Regel criterion. Moreover, several papers have analyzed the question of renormalization of the phonon spectrum because of the appearance of coherent inverse scattering effects in anharmonic processes. Reference 14 was the first to consider such features of the spectrum of quasiparticles close to localization thresholds, and a certain hypothetical situation was studied. We have studied the temperature and concentration dependence of spectra close to the thresholds[15] as well as in the ultrasound limit[16,17] for systems with resonant-scattering defects. According to our results, coherent inverse scattering processes in the ultrasound limit in a standard three-dimensional anharmonic lattice result in substantial renormalization of the group velocity and the phonon lifetime. This relaxation mechanism can predominate over the Rayleigh mechanism and conventional anharmonic relaxation mechanisms.

In the case of phonons, the situation has been studied for a long time, but not so fully as for electrons and electromagnetic radiation. We emphasize that, until now, the question of how coherent effects accompanying the propagation and multiple scattering of phonon modes affect the frequency and low-temperature behavior of the kinetic coefficients in low-dimension compounds has gone virtually undiscussed in

the literature. Localization phenomena must be more pronounced in low-dimensional systems than in standard three-dimensional compounds. It therefore is of interest to carry out theoretical studies of vibrational-mode localization in layered and quasi-one-dimensional compounds. The treatment of this problem for layered crystals is the goal of the present paper.

Below, we discuss the possibility of weak localization of vibrational excitations in certain nonideal strongly anisotropic harmonic lattices. We assume that the separate layers of such lattices are weakly coupled with each other and that two-dimensional properties appear. We consider the cases of strong diagonal and off-diagonal disorders. A scalar model of the lattice is used for the sake of simplicity. Specifically, the frequency behavior of the diffusivity tensor has been studied. In determining the diffusivity for phonon quasiparticles, we employ a rigorous expression of the Kubo type for the lattice thermal conductivity. Using such an expression, we can obtain an analog of the Bethe–Salpeter equation for the two-particle Green's function, in terms of which the diffusivity is expressed. We emphasize that coherent inverse scattering processes are taken into account when the two-particle lattice Green's function averaged over impurity configurations is determined.

Note that, in layered lattices with diagonal disorder (i.e., with heavy defects), virtually no well-defined quasilocal modes appear in the region where the system displays quasi-two-dimensional properties.[18–20] In this case, the renormalization of the group velocity of the phonons (but not of the lifetime) is comparatively weak. A standard type of quasilocal modes appears, but under conditions of off-diagonal disorder; i.e., there is drastic weakening of the local force parameters for the impurity atom.[20–28] Only in this case it is possible to speak of a system of resonant-scattering impurity centers.

The experimental data on the nonstandard temperature dependence of the thermal conductivity of single crystals of layered systems of the high-temperature superconductor (HTSC) $Bi_2Sr_2CaCu_2O_{8+y}$ (BSCCO) and its insulating analog $Bi_2Sr_2YCu_2O_8$ (BSYCO) are discussed in terms of the theory that is developed.

2. MODEL OF THE CRYSTAL LATTICE

Let us first of all discuss a model of a crystal lattice whose separate layers are weakly coupled to each other. A lattice with tetragonal symmetry is used as the model in this paper. It is assumed that only the nearest-neighbor atoms dynamically interact and that the effective interaction between the atoms in the base plane xy is much stronger than along the fourth-order symmetry axis z.

In this model of the lattice, the dispersion law for a vibrational mode with quasimomentum $\mathbf{k}$ is written as

$$\omega^2(\mathbf{k})=\omega_\parallel^2\left(\sin^2\frac{ak_x}{2}+\sin^2\frac{ak_y}{2}\right)+\omega_\perp^2\sin^2\frac{bk_z}{2}. \qquad (1)$$

Here $\{a,b\}$ are the components of a translational lattice vector. The symbols $\omega_\parallel^2$ and $\omega_\perp^2$ are used to denote the quantities

$$\omega_\parallel^2=\frac{4\gamma_\parallel}{M_0},\quad \omega_\perp^2=\frac{2\gamma_\perp}{M_0},$$

where M_0 is the mass of an atom of the lattice–matrix. By definition, the force parameters $\gamma_\parallel$ and $\gamma_\perp$ ($\gamma_\parallel,\gamma_\perp>0$) characterize the interaction in the xy plane and along the z axis. According to the discussion above, $\gamma_\parallel>\gamma_\perp$.

In the long-wavelength limit, when the conditions

$$ak_{x(y)}\ll 1,\quad bk_z\ll 1, \qquad (2)$$

are satisfied, Eq. (1) simplifies to

$$\omega^2(\mathbf{k})=\dot{v}_\parallel^2k_\parallel^2+\dot{v}_\perp^2k_\perp^2,\quad k_\parallel^2=k_x^2+k_y^2.$$

The quantities $\dot{v}_\parallel$ and $\dot{v}_\perp$ that appear here are the sound velocities, with

$$\dot{v}_\parallel=\frac{a\omega_\parallel}{2},\quad \dot{v}_\perp=\frac{b\omega_\perp}{\sqrt{2}}.$$

In the situation under consideration, when the dispersion law corresponds to the case of a three-dimensional lattice, the spectral function of the square density of phonon states, $g(\omega^2)$, is defined by the chain of equations

$$g(\omega^2)=\frac{1}{N}\sum_{\mathbf{k}}\delta(\omega^2-\dot{v}_\parallel^2k_\parallel^2-\dot{v}_\perp^2k_\perp^2)=\frac{a^2b}{(2\pi)^2}\frac{\omega}{\dot{v}_\parallel^2\dot{v}_\perp^2}$$
$$=\frac{\sqrt{2}}{\pi^2}\frac{\omega}{\omega_\parallel^2\omega_\perp}. \qquad (3)$$

At relatively low frequencies,

$$0<\omega^2\leqslant 2\omega_\perp^2,$$

a situation is possible in which a condition having the form of the inequalities (2) is still satisfied for the quantity $k_{x(y)}a$, but it breaks down for the components of the wave vector along a direction perpendicular to the layers. We then have

$$\omega^2(\mathbf{k})=\dot{v}_\parallel^2k_\parallel^2+2\omega_\perp^2\sin^2\frac{k_\perp b}{2}. \qquad (1')$$

Using Eq. (1′), we represent the density of states $g(\omega^2)$ in the form

$$g(\omega^2)=\frac{a^2b}{2\pi^2}\frac{1}{b\omega_\perp^2}\int dk_\parallel k_\parallel\int\frac{d\cos(bk_\perp)}{\sin(bk_\perp)}$$
$$\times\delta\left(\cos(bk_\perp)-\frac{\omega_\perp^2-\omega^2+\dot{v}_\parallel^2k_\parallel^2}{\omega_\perp^2}\right).$$

From this, after elementary transformations, we get

$$g(\omega^2)=\frac{1}{\pi^2\omega_\parallel^2}\arccos\frac{\omega_\perp^2-\omega^2}{\omega_\perp^2}.$$

We turn our attention to the following fundamental fact. Namely, if the frequency of a phonon mode lies in the interval

$$2\omega_\perp^2<\omega^2<\omega_\parallel^2.$$

then it depends weakly on k_z, as a consequence of which the dispersion law given by Eq. (1′) must approximately correspond to the case of a two-dimensional lattice. In this case, we have

$$g(\omega^2)=a^2b\int \frac{dk_\perp}{2\pi}\int \frac{d(k_\parallel^2\dot{v}_\parallel^2)}{2(2\pi)\dot{v}_\parallel^2}\,\delta\Big(\omega^2-\dot{v}_\parallel^2k_\parallel^2 -2\omega_1^2\sin^2\frac{bk_1}{2}\Big)=\frac{1}{\pi\omega_\parallel^2}. \tag{4}$$

Thus the density of the square phonon states, $g(\omega^2)$, is independent of frequency in the frequency interval under consideration.

Note that it is assumed in this paper that there is strong anisotropy of the interatomic interaction force. It is assumed at the same time that the unit cell parameters can have different orders of magnitude. We shall consider both cases: $a\approx b$ and $a\ll b$. For more detail concerning the model, see, for example, Kosevich's monograph.[21]

3. LATTICE THERMAL CONDUCTIVITY AND DIFFUSIVITY IN AN ANISOTROPIC IRREGULAR HARMONIC CRYSTAL

3.1. Diffusivity

As is well known, the component of the lattice thermal conductivity tensor $\kappa_{\alpha\alpha'}$ is expressed in terms of the correlation function of the energy-flux operators **f**. We have[22]

$$\kappa_{\alpha\alpha'}=\frac{1}{2T^2V}\int_{-\infty}^{\infty}dt\langle\langle f_\alpha(t),f_{\alpha'}(0)\rangle\rangle_c, \tag{5}$$

where T is the temperature and V is the volume of a unit cell of the lattice. The symbols $\langle ...\rangle$ and $\langle\ldots\rangle_c$ denote, respectively, the operations of averaging over the equilibrium thermodynamic distribution with the Hamiltonian of a harmonic crystal and over the implemented impurity configurations. The quantity f_α that appears in Eq. (5) is defined as

$$f_\alpha=\frac{1}{2}\sum_{\mathbf{s},\mathbf{s}'}\Phi^{\mathbf{ss}'}_{\alpha\alpha'}\mathbf{u}_\mathbf{s}^\alpha\frac{p_{\mathbf{s}'}^{\alpha'}}{M_{\mathbf{s}'}}R_\alpha^{\mathbf{ss}'},\quad R_\alpha^{\mathbf{ss}'}=R_\mathbf{s}^\alpha-R_{\mathbf{s}'}^\alpha. \tag{6}$$

Here the quantities $u_\mathbf{s}^\alpha$ and $p_\mathbf{s}^\alpha$ are the Cartesian components of the displacement and momentum operators of the **s** atom (**s** is the site index), $M_\mathbf{s}$ is the mass of the atom at the **s** site, and $\Phi^{\mathbf{ss}'}_{\alpha\alpha'}$ are the matrix elements of the second-order force parameters. The symbol $R_\mathbf{s}^\alpha$ denotes the component of the radius vector of the **s**th lattice site.

Using Eq. (6), Eq. (5) can be represented as

$$\kappa=\kappa_1+\kappa_2, \tag{7}$$

$$\kappa_1=\frac{1}{6\pi T^2}\sum_{\substack{\mathbf{s},\mathbf{s}'\\ \mathbf{s}_1,\mathbf{s}_1'}}(\Phi^0R)_{\mathbf{ss}'}(\Phi^0R)_{\mathbf{s}_1\mathbf{s}_1'}\int_0^\infty d\omega\;\omega^2n(\omega) \times[n(\omega)+1]\langle G^+_{\mathbf{ss}_1}(\omega)G^-_{\mathbf{s}_1',\mathbf{s}'}(\omega)\rangle_c, \tag{8}$$

$$\kappa_2=\frac{1}{3\pi T^2}\sum_{\substack{\mathbf{s},\mathbf{s}'\\ \mathbf{s}_1,\mathbf{s}_1'}}[(2\Phi^0(R)_{\mathbf{ss}'}\langle(\Delta\Phi R)_{\mathbf{s},\mathbf{s}_1'} +(\Delta\Phi R)_{\mathbf{ss}'}(\Delta\Phi R)_{\mathbf{s}_1\mathbf{s}_1'}]\int_0^\infty d\omega\;\omega^2n(\omega)[n(\omega]+1) \times\langle\mathrm{Im}\,G^+_{\mathbf{ss}_1}(\omega)\mathrm{Im}\,G^+_{\mathbf{s}_1'\mathbf{s}'}(\omega)\rangle_c. \tag{9}$$

Here the symbols $G^\pm_{ss'}$ denote the retarding and advancing single-particle lattice Green's functions (G functions). These G functions are ''assembled'' on the operators of the dynamic atomic displacements $\mathbf{u}_\mathbf{s}$.

In the expression for κ given by Eq. (7), the first term κ_1, given by Eq. (8), describes the contribution to the thermal conductivity of the standard phonon modes, which are scattered at the dynamically perturbed regions. As shown in Appendix A, the second term κ_2, given by Eq. (9), is associated with the impurity modes. In the low-temperature region, the behavior of κ is determined in this case by κ_1.

Keeping in mind what has been said, we represent κ as

$$\kappa_{\alpha\alpha'}\approx\frac{1}{3\pi T^2}\int_0^\infty d\omega\omega^2n(\omega)[n(\omega)+1] \times\sum_j g_j(\omega)D^j_{\alpha\alpha'}(\omega), \tag{10}$$

where by definition g_j and D^j are the density of vibrational states and the diffusivity of the jth vibrational mode; $n(\omega)=[\exp(\omega/T)-1]^{-1}$.

We shall consider a layered crystal. Its lattice possesses axial symmetry. The tensor κ in this case has two principal values, which we denote by $\kappa_\parallel$ and $\kappa_\perp$. For simplicity, we restrict ourselves to the single-mode approximation and assume that the matrices of the force parameters are diagonal in the Cartesian indices. For the sake of brevity, the set of site **s** and Cartesian α indices are denoted as s. We then have

$$D_{\alpha\alpha}(\omega)=\frac{1}{\pi g(\omega)}\sum_{\mathbf{s},\mathbf{s}'}\sum_{\mathbf{s}_1,\mathbf{s}_1'}R_\alpha^{\mathbf{ss}'}R_\alpha^{\mathbf{s}_1\mathbf{s}_1'}\Phi^0_{\mathbf{s},\mathbf{s}'}\Phi^0_{\mathbf{s}_1,\mathbf{s}_1'} \times\langle G^+_{\mathbf{ss}_1}(\omega)G^-_{\mathbf{s}_1'\mathbf{s}'}(\omega)\rangle_c. \tag{11}$$

Note that the quantity $G_2=\langle G^+G^-\rangle_c$ is a two-particle lattice Green's function.

Let us convert Eq. (11) from the site representation to the momentum representation. After a number of transformations, we get

$$D_{\alpha\alpha}(\omega)=\frac{1}{\pi g(\omega)}\sum_{\mathbf{k}\mathbf{k}'}\dot{v}_\alpha(\mathbf{k})\dot{v}_\alpha(\mathbf{k}')\omega(\mathbf{k})\omega(\mathbf{k}') \times G_2(\mathbf{k};\mathbf{k}';\omega), \tag{12}$$

$$G_2(\mathbf{k},\mathbf{k}',\omega)=\lim_{\Omega\to0}\langle G^+_{\mathbf{k},\mathbf{k}'}(\omega)G^-_{\mathbf{k},\mathbf{k}'}(\omega-\Omega)\rangle_c. \tag{13}$$

We recall that $\omega(\mathbf{k})$ and $v(\mathbf{k})=\partial\omega(\mathbf{k})/\partial\mathbf{k}$ are the dispersion law and the group velocity of the phonon mode with quasimomentum **k**. See also Ref. 11 for an expression for the diffusivity in a form like Eqs. (12) and (13).

In the case of insulators in the region of the limiting low temperatures, the phonon mean free paths $l_{\rm ph}$ are determined by the geometrical size of the sample. For comparatively high temperatures, the value of $l_{\rm ph}$ is dictated by anharmonic Umklapp processes. In the intermediate region, the flight path is sensitive to defects.[23] Equations (10), (12), and (13) describe the behavior of κ in just this intermediate temperature interval with reasonable accuracy.

3.2. The two-particle Green's function and the Bethe–Salpeter equation

In the simplest approximation, the impurities are isotopic defects. However, an impurity actually interacts differently with its surroundings than do the atoms of the matrix. We are interested in the qualitative picture. We shall therefore assume that the perturbation of the interaction force parameters between an individual impurity and its surroundings extends only to the nearest coordination spheres. The defect concentration c is assumed to be low. The total perturbation is represented as the sum of the contributions from the individual defects. It can be stated that specific quasimolecules appear in a crystal, each of which is formed by an impurity and the matrix atoms that dynamically interact with it. The perturbation operator in the site representation has the form

$$\widetilde{V}_{\mathbf{s}_1\mathbf{s}_2}=\sum_{\boldsymbol{\mu}} c_{\boldsymbol{\mu}} V^{\boldsymbol{\mu}}_{\mathbf{s}_1\mathbf{s}_2},$$

$$V^{\boldsymbol{\mu}}_{\mathbf{s}_1\mathbf{s}_2}=\sum_{\mathbf{L},\mathbf{L}'} \delta_{\mathbf{s}_1,\boldsymbol{\mu}+\mathbf{L}_{\boldsymbol{\mu}}}\delta_{\mathbf{s}_2,\boldsymbol{\mu}+\mathbf{L}'_{\boldsymbol{\mu}}} V_{\mathbf{L}\mathbf{L}'\boldsymbol{\mu}}. \tag{14}$$

$$V_{\mathbf{L}\mathbf{L}'}=M_0\epsilon\omega^2\delta_{\mathbf{L}\mathbf{L}'}+\Delta\Phi_{\mathbf{L}\mathbf{L}'}, \quad \epsilon=1-M_d/M_0. \tag{15}$$

In Eq. (14), V^{μ} denotes the perturbation introduced by the individual defect located at the μth site; $\mathbf{L}$ and $\mathbf{L}'$ are summed over the sites occupied by the atoms of the quasimolecule, i.e., over the site $\mu=\mathbf{d}$ and over the sites adjacent to it.

In the momentum representation, an equation of the form

$$\begin{aligned} G_2(\mathbf{k},\mathbf{k}',\omega)&=\lim_{\Omega\to 0}\langle G^+_{\mathbf{k},\mathbf{k}'}(\omega)G^-_{\mathbf{k},\mathbf{k}'}(\omega-\Omega)\rangle_c \\ &=\lim_{\Omega\to 0}\overline{G}^+_{\mathbf{k}}(\omega)\overline{G}^-_{\mathbf{k}}(\omega-\Omega)\Big(\delta_{\mathbf{k}\mathbf{k}'} \\ &\quad+\sum_{\mathbf{k}_1} U(\mathbf{k}\cdot\mathbf{k}_1;\omega,\Omega) \\ &\quad\times\langle G_2(\mathbf{k}_1,\mathbf{k}',\omega,\Omega)\rangle_c\Big), \end{aligned} \tag{16}$$

can be used in the general case for a nonideal harmonic crystal lattice, where $\overline{G}^+_{\mathbf{k}}(\omega)$ is the Fourier component of the single-particle function $G_{\mathbf{s}\mathbf{s}'}=\langle G_{\mathbf{s}\mathbf{s}'}\rangle_c$ for a nonideal harmonic system, and $U(\mathbf{k}\cdot\mathbf{k}_1;\omega,\Omega)$ is the vertex part. Equations like Eq. (16) are called Bethe–Salpeter equations.

We assume here and in what follows that the contribution of only the fan diagrams that describe inverse coherent scattering processes and determine the weak-localization regime is introduced into the vertex part $U(\mathbf{k},\mathbf{k}';\omega,\Omega)$.[4,11,14] We have

$$U(\mathbf{k},\mathbf{k}';\omega,\Omega)=\frac{\Gamma}{N}\left[1-\frac{\Gamma}{N}\sum_{\mathbf{k}_1}\overline{G}^+_{\mathbf{k}_1}(\omega)\overline{G}^-_{\mathbf{k}_1-\mathbf{q}}(\omega-\Omega)\right]^{-1},$$

$$\mathbf{q}=\mathbf{k}+\mathbf{k}'. \tag{17}$$

In Eq. (17), the bare vertex is $\Gamma=ct^+t^-$; i.e., it is defined in terms of single-site t scattering matrices.

Let us transform to the single-particle Green's function. We set

$$\overline{G}^+_{\mathbf{k}}(\omega)=\left[\omega^2-\omega^2_{\mathbf{k}}-P(\omega)-i\,\frac{\omega}{\tau_i(\omega)}\right]^{-1}, \tag{18}$$

where $P(\omega)$ and $\omega/\tau_i(\omega)$ are the real and imaginary parts of the mass operator Σ, which corresponds to the elastic interaction of phonons with the defects.

In specific applications, for $G^+_{\mathbf{k}}(\omega)$, instead of Eq. (18), we use below a representation of the form

$$\overline{G}^+_{\mathbf{k}}(\omega)=Q(\omega)\left[\omega^2-\widetilde{\omega}^2_{\mathbf{k}}-i\,\frac{\omega}{\tau'_i(\omega)}\right]^{-1}. \tag{19}$$

Here $\widetilde{\omega}_{\mathbf{k}}$ is the renormalized dispersion law, for which

$$\widetilde{\omega}^2_{\mathbf{k}}-\omega^2_{\mathbf{k}}-P(\widetilde{\omega}_{\mathbf{k}})=0. \tag{20}$$

The factor $Q(\omega)$ is given by

$$Q^{-1}(\omega)=1-\cos\Phi\,\frac{\partial|\Sigma(\omega)|}{\partial\omega^2}+\mathrm{Im}\,\Sigma(\omega)\,\frac{\partial\Phi(\omega)}{\partial\omega^2}. \tag{21}$$

In this case, we have for the resonant part of the phase shift

$$\Phi(\omega)=\arctan\frac{\mathrm{Im}\,\Sigma(\omega)}{\mathrm{Re}\,\Sigma(\omega)} \tag{22}$$

[see the derivation of Eqs. (21) and (22) in Appendix B]. Moreover, we set

$$\tau'_i=Q^{-1}\tau_i, \tag{23}$$

where τ'_i is the lifetime of the quasiparticles for a nonideal harmonic system.

In this case, in a narrow frequency interval close to the resonance frequency [where $\mathrm{Re}\,\Sigma(\omega_R)=0$] and at a certain distance from it, instead of Eq. (21), we have approximately

$$Q^{-1}(\omega)\approx\begin{cases} 1+\mathrm{Im}\,\Sigma(\omega)\,\dfrac{\partial\Phi}{\partial\omega^2}, & |\omega^2-\omega^2_R|<\Gamma_R, \\ 1-\dfrac{\partial P(\omega)}{\partial\omega^2}, & |\omega^2-\omega^2_R|>\Gamma_R, \end{cases} \tag{24}$$

where Γ_R is the resonance width.

We emphasize that we have taken into account multiple scattering processes at the impurity site in finding the single-site scattering matrix. When this is done, the real part of the mass operator in the immediately vicinity of resonance has strong dispersion. When the single-particle phonon Green's function was defined in the form of Eq. (19), this dispersion

was taken into account by introducing the Q factor. Furthermore, the scattering cross section in the standard theory is expressed in terms of phase shifts. When this is done, the cross section in the case of resonant scattering is expressed by means of the Breit–Wigner formula. At the same time, the cross section is proportional to the square of the magnitude of the t matrix. There is consequently a connection between the resonant part of the phase shift and the components of the mass operator. The reflection of this fact is the expression of the Q factor close to resonance in terms of the derivative of the resonant part of the phase shift, given by Eq. (24) (see, for example, Taylor's monograph[24]).

Based on the expression for the dispersion law, Eq. (20), we have for the components of the group velocity vector

$$v_{\parallel}^2=\dot{v}_{\parallel}^2 Q^2(\omega)\left[1-\frac{P(\omega)}{\omega^2}-2\,\frac{\omega_{\perp}^2}{\omega^2}\sin^2\frac{k_{\perp}b}{2}\right],$$

$$v_z=v_{\perp}=Q(\omega)\,\frac{\omega_{\perp}^2 b}{2\omega}\sin(k_{\perp}b).$$

Let us turn our attention to the strong angular dispersion of the group velocity.

Since the defect concentration c is assumed to be low (i.e., $c\ll 1$), the method of expansion in powers of concentration, developed in Refs. 25 and 26, is suitable for determining the mass operator $\Sigma(\mathbf{k},\omega)$. In the momentum representation, under conditions of off-diagonal disorder, such an expansion has the form[11]

$$\begin{aligned}\Sigma(\mathbf{k},\omega)\approx\sum_{L_{\mu}L_{\mu_1}} e^{-i\mathbf{k}\mathbf{L}_{\mu}}\Bigg[& c t^{\mu}_{\mathbf{L}_{\mu}\mathbf{L}_{\mu_1}}\delta_{\mu\mu_1}\\ &+c^2\sum_{\mu_1\neq\mu}\left(\frac{t^{\mu}G^{0\mu_1}_{\mu}t^{\mu_1}G^{0\mu}_{\mu_1}t^{\mu}}{1-G^{0\mu_1}_{\mu}t^{\mu_1}G^{0\mu}_{\mu_1}t^{\mu_1}}\right)_{\mathbf{L}_{\mu}\mathbf{L}_{\mu_1}}\delta_{\mu\mu_1}e^{i\mathbf{k}L_{\mu_1}}\\ &+c^2\sum_{\mu_1\neq\mu}\left(\frac{t^{\mu}G^{0\mu_1}_{\mu}t^{\mu_1}}{1-G^{0\mu}_{\mu_1}t^{\mu}G^{0\mu_1}_{\mu}t^{\mu_1}}\right)_{\mathbf{L}_{\mu}\mathbf{L}_{\mu_1}}\\ &\times e^{-i\mathbf{k}(\mathbf{R}_{\mu}-\mathbf{R}_{\mu_1})}\ldots \end{aligned} \tag{25}$$

The t matrix used above for scattering at an individual quasimolecule is defined as

$$t^{\mu}=\frac{V^{\mu}}{1-G^{0\mu}_{\mu}V^{\mu}}. \tag{26}$$

where $G^{0\mu_1}_{\mu}$ is the Green's function of the ideal lattice. The indices $\boldsymbol{\mu}$ and $\boldsymbol{\mu}_1$ mean that the first site index belongs to the quasimolecule μ, while the second belongs to the quasimolecule μ_1. The first term in Eq. (25) in this case describes phonon scattering at isolated impurity quasimolecules, while the second describes the scattering at all groups of two quasimolecules. The omitted terms describe scattering by groups of three or more impurity quasimolecules.

In Ref. 27, the Binet–Cauchy theorem is used to show that the elements of the scattering matrix in the site representation are defined by

$$t_{\mathbf{ss}'}=\frac{V_{\mathbf{ss}'}+\Sigma_{p=1}^{I}(-1)^p\Sigma_{i_i,k_i}V\begin{pmatrix}si_1\ldots i_p\\ s'k_1\ldots i_p\end{pmatrix}\dot{G}\begin{pmatrix}k_1\ldots k_p\\ i_1\ldots i_p\end{pmatrix}}{1+\Sigma_{p=1}^{I}(-1)^p\Sigma_{i_i,k_i}V\begin{pmatrix}i_1\ldots i_p\\ k_1\ldots k_p\end{pmatrix}\dot{G}\begin{pmatrix}k_1\ldots k_p\\ i_1\ldots i_p\end{pmatrix}}.$$

Here $A\left(\begin{smallmatrix}si_1\ldots i_p\\ s'k_1\ldots i_p\end{smallmatrix}\right)$ is the minor of matrix (G,V) of dimension $L\times L$, which is composed of elements lying on the intersection of the rows $L_p=(i_1\ldots i_p)$ and columns $L_p'=(k_1\ldots k_p)$. Each index in the sets L_p and L_p' can be encountered only once. For any permutation of indices, the value of $A\left(\begin{smallmatrix}si_1\ldots i_p\\ s'k_1\ldots i_p\end{smallmatrix}\right)$ changes sign.

For a qualitative description of the phonon spectra in the case $c\ll 1$, it is sufficient in practice to take into account in Eq. (25) terms to second order in the concentration, inclusive. The effect of the interaction of phonons with groups of more than two impurities must be expressed weakly because of the anharmonicity. It is assumed that the decomposition over groups of interacting impurities converges in an asymptotic sense. In other words, such a series is considered convergent if, for its first few terms, each subsequent term is less than the preceding one (the divergence close to singular points can be formally eliminated by taking into account anharmonic damping). We shall return to the question of the convergence of the series when discussing specific models.

In specific calculations, it must be taken into account that the Green's function G^0 possesses the symmetry of the crystal, while the perturbation V possesses the symmetry of the quasimolecule. To do so, it is convenient to use the methods of group-representation theory.

As for the bare vertex Γ, as can be shown, it is determined by a relationship of the form

$$\begin{aligned}\Gamma=c\sum_{\mathbf{L}_{\mu}\mathbf{L}'_{\mu}\mathbf{L}_{\mu^{(1)}}\mathbf{L}'_{\mu^{(1)}}}&\exp(-i\mathbf{k}_1\mathbf{L}_{\mu})\tilde{t}^{(+)\mu}_{\mathbf{L}_{\mu}\mathbf{L}'_{\mu}}\exp(i\mathbf{k}_2\mathbf{L}'_{\mu})\\ &\times\exp(i\mathbf{k}_3\mathbf{L}_{\mu^{(1)}})\tilde{t}^{(-)\mu^{(1)}}_{\mathbf{L}_{\mu^{(1)}}\mathbf{L}'_{\mu^{(1)}}}\exp(-i\mathbf{k}_4\mathbf{L}'_{\mu^{(1)}})\\ &=c\tilde{t}^{(+)}(\mathbf{k}_1,\mathbf{k}_2)\tilde{t}^{(-)}(\mathbf{k}_3,\mathbf{k}_4).\end{aligned} \tag{27}$$

In this case, single-site matrix $\overline{t}$ is expressed in terms of V^{μ} from Eqs. (14) and (15) and $\overline{G}_{\mathbf{k}}(\omega)$ from Eq. (19).

4. THE CASE OF DIAGONAL DISORDER

4.1. Behavior of the t matrix and the Σ operator

We consider the case of diagonal disorder for a layered crystal, in which the force parameters are not perturbed and we have $\Delta\Phi=0$. We shall analyze the behavior of the single-site scattering matrix t, as well as of the mass operator $\Sigma(\omega)$. First, by using the identity

$$\sum_{\mathbf{s}\neq 0}G^{02}_{0\mathbf{s}}(\omega)=-\frac{\partial G^0_{00}}{\partial\omega^2}-G^{02}_{00}(\omega),$$

in place of Eq. (25), we get

$$\Sigma(\omega)=ct(\omega)[1+\Delta_c(\omega)], \tag{28}$$

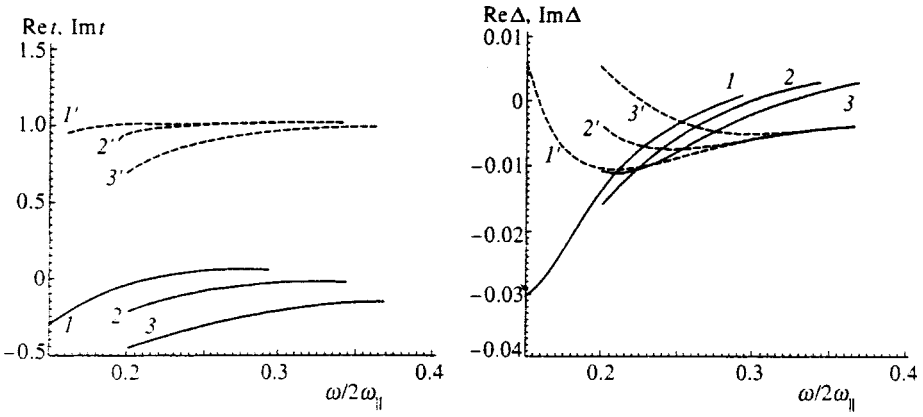

FIG. 1. Re $t(\omega)$, Im $t(\omega)$, Re $\Delta_c(\omega)$, Im $\Delta_c(\omega)$ vs $\omega/2\omega_\parallel$. In the case of the t matrix, $1(1')$, $2(2')$, $3(3')$ are curves corresponding to the values of parameter ϵ^2 equal, respectively, to 5, 7, 10; for the factor Δ_c, $1(1')$, $2(2')$, $3(3')$ are curves corresponding to values of parameter $c\epsilon^2$ equal, respectively, to 0.5, 0.75, and 1. The indices with the prime are used for marking the imaginary parts of t and Δ_c.

$$\Delta_c(\omega)=ct^2(\omega)\left[-\frac{\partial G_{00}^0}{\partial\omega^2}-G_{00}^{02}(\omega)\right]+\frac{c}{2}\sum_{\mathbf{s}\neq 0}G_{0\mathbf{s}}^{02}(\omega)t^2(\omega)\left(\frac{1+e^{i\mathbf{k}\cdot\mathbf{s}}}{1-t(\omega)G_{0\mathbf{s}}^0}-2\right)+\frac{c}{2}\sum_{\mathbf{s}\neq 0}G_{0\mathbf{s}}^{02}t^2(\omega)\,\frac{1-e^{i\mathbf{k}\cdot\mathbf{s}}}{1+t(\omega)G_{0\mathbf{s}}^0}+\ldots \qquad (28')$$

$[G_{00}^0(\omega)=G_{0\mathbf{s}=0}^0(\omega)]$. In this case, the t matrix given by Eq. (26) is

$$t^{-1}(\omega)=\frac{V}{1-VG_{00}^0(\omega^2)},\quad V=+\epsilon\omega^2. \qquad (28'')$$

We should point out that both the sums in Eq. (28) converge rapidly. The convergence criterion of the series for $\Sigma(\omega)$, according to Refs. 18 and 19, is that

$$\left|t^2(\omega)\,\frac{\partial\dot{G}_{00}(\omega)}{\partial\omega^2}\right|\ll 1. \qquad (29)$$

As far as the Green's function $G_0^0(\omega^2)$ of the ideal crystal is concerned, in the frequency region where the lattice displays quasi-two-dimensional properties, in accordance with Eq. (4),

$$\text{Im } G_{00}^0(\omega^2)=\pi g(\omega^2)\approx\frac{1}{\omega_\parallel^2},$$

where $\omega_\parallel^2$ is the square of the maximum frequency of the vibrations in the plane. Using the Kramers–Kronig relation, we get

$$\text{Re } G_{00}^0(\omega^2)=\int_0^\infty\frac{d\omega'^2}{\pi}\,P\,\frac{1}{\omega^2-\omega'^2}\,\text{Im } G_{00}^0(\omega^2)\approx\frac{1}{\pi\omega_\parallel^2}\ln\frac{\omega_\parallel^2}{\omega^2}.$$

Consequently,

$$G_{00}^0(\omega^2)\approx-\frac{1}{\pi\omega_\parallel^2}\ln\frac{\omega_\parallel^2}{\omega^2}-i\pi g(\omega^2)\approx-\frac{1}{\pi\omega_\parallel^2}\ln\frac{\omega_\parallel^2}{\omega^2}-\frac{i}{\omega_\parallel^2}. \qquad (30)$$

Taking into account the explicit form of the Green's function G_{00}^0 given by Eq. (30) and of the diagonal perturbation operator V given by Eqs. (14) and (15), we find that the denominator of the t matrix given by Eq. (28″) is

$$1-VG_0^0(\omega^2)\approx 1+\frac{\epsilon\omega^2}{\pi\omega_\parallel^2}\ln\frac{\omega_\parallel^2}{\omega^2}+i\,\frac{\epsilon\omega^2}{\omega_\parallel^2},$$

whence it can immediately be seen that the real part of the t matrix in the case of extremely heavy impurities, generally speaking, can go to zero for values of the frequency satisfying the equation

$$1+\frac{\epsilon\omega_R^2}{\pi\omega_\parallel^2}\ln\frac{\omega_\parallel^2}{\omega_R^2}=0. \qquad (31)$$

In this case, the broadening of such a level is of the order of magnitude of

$$\Gamma_R\approx|\epsilon|\omega_R^2.$$

Thus, if Eq. (31) is satisfied in the frequency region under consideration for reasonable values of parameter ϵ, the condition $\Gamma_R/\omega_R^2\ll 1$ is not fulfilled, as a result of which it is virtually meaningless to talk about quasilocal vibrations existing in a narrow frequency interval, as well as about resonant scattering at such modes. In confirmation of what has been said, Fig. 1 shows graphs of the real and imaginary parts of the single-site t matrix of Eq. (28″) for three values of the parameter ϵ.

We stress that, if the presence of heavy impurities in weakly anisotropic crystals always results in the appearance of quasilocal modes, it is difficult for such modes to appear in low-dimensional crystals, as first shown in Refs. 18 and 19. However, such modes do appear in the case of weakly coupled impurities (see Sec. 5).

We now discuss the question of the corrections to the mass operator of the phonons which are quadratic in the concentration, i.e., the question of the value of Δ_c in Eq. (28). In the case of strong disorder, when $c\ll 1$ and $\epsilon^2 c\leqslant 1$ simultaneously, on one hand, the relationships

$$\frac{\text{Re } t}{\omega_\parallel^2}\ll 1,\quad \frac{\text{Im } t}{\omega_\parallel^2}\approx 1.$$

are satisfied in the frequency interval $|\omega-\omega_R|/\omega_R<1$. On the other hand, the lattice Green's function, according to Eq. (30), is of the order of $\omega_\parallel^{-2}$, as a result of which we have

$$\Delta_c \approx c\epsilon^2 \frac{\omega^2}{\omega_\parallel^2}(1+i).$$

Since $\omega_R^2 \ll \omega_\parallel^2$ holds, we get $\Delta_c \ll 1$. In confirmation of what has been said, Fig. 1 also shows graphs for the real and imaginary parts of the factor Δ_c for several values of the diagonal-disorder parameter $c\epsilon^2$.

It follows from everything that has been said above in this section, that, even under conditions of strong diagonal disorder, renormalization of the phonon-mode frequencies can be neglected in a certain rather wide region of frequencies close to ω_R. As far as damping is concerned, only the term linear in concentration is needed in the expression for Σ. As a result, the configuration-averaged single-particle Green's function of the phonon mode with quasimomentum $\mathbf{k}$ can be written as

$$\bar{G}_{\mathbf{k}}^{+}(\omega) \approx \left[\omega^2-\omega^2(\mathbf{k})-i\frac{\omega}{\tau_i(\omega)}\right]^{-1}. \tag{32}$$

Here $\omega/\tau_i(\omega)$ is the imaginary part of the polarization operator, i.e., ct. In this case,[23]

$$\frac{1}{\tau_i(\omega)} = \sigma_m \omega^2 g(\omega), \quad \sigma_m = \frac{\pi}{2} c\left(\frac{M_d-M_0}{M_0}\right)^2 = \frac{\pi}{2} c\epsilon^2.$$

In the case under consideration, the bare vertex Γ in the Bethe–Salpeter equation immediately satisfies an identity of the Ward type. Namely,

$$t^+(\omega)-t^-(\omega) = \frac{\Gamma}{N}\sum_{\mathbf{k}} [G_{\mathbf{k}}^+(\omega)-G_{\mathbf{k}}^-(\omega)],$$

whence, using Eqs. (32) and (1′), we get

$$1 = \frac{\Gamma}{N}\sum_{\mathbf{k}} \left[\left(\omega^2-k_\parallel^2 v_\perp^2-2\omega_\perp^2 \sin^2\frac{bk_\perp}{2}\right)^2 + \left(\frac{\omega}{\tau_i(\omega)}\right)^2\right]^{-1}. \tag{33}$$

In the low-frequency region, $0 \leqslant \omega^2 \ll \omega_\parallel^2$, instead of Eq. (33), we have

$$\Gamma = \frac{\omega\omega_\parallel^2}{\tau_i(\omega)} = \frac{\omega}{\pi\tau_i(\omega) g_\parallel(\omega^2)}. \tag{34}$$

Using the explicit expressions for the phonon Green's function of Eq. (32) and the bare vertex Γ of Eq. (34), we now determine the diffusion vertex and afterwards the components of the diffusivity tensor, Eq. (12).

4.2. The diffusion vertex

Let us obtain the expression for the vertex $U(\mathbf{q};\omega;\Omega)$ in the weak-localization regime, where

$$ql \ll 1, \quad \Omega\tau_i \ll 1 \tag{35}$$

[$l=v\tau_i(\omega)$ is the phonon flight path]. We shall limit ourselves to the frequency region for which the phonon dispersion law exhibits two-dimensional properties.

Taking into account what has been said, we consider the expression

$$J = \frac{\Gamma}{N}\sum_{\mathbf{k}} \bar{G}_{\mathbf{k}}^{+}(\omega)\bar{G}_{\mathbf{q}-\mathbf{k}}^{-}(\omega+\Omega). \tag{36}$$

We transform in Eq. (36) from summation over $\mathbf{k}$ to integration over $d\mathbf{k}$. We replace the sum $\Sigma_{\mathbf{k}}$ with

$$\frac{a^2 b}{(2\pi)^3 2\dot{v}_\parallel}\int_{-\pi/b}^{\pi/b} dk_\perp \int_0^{2\pi} d\varphi \int d(\dot{v}_\parallel^2 k_\parallel^2)\ldots \tag{37}$$

We shall use for the Green's functions in expression (37) the representation in Eq. (32) and shall take the integral over $dk_\parallel^2$ by finding the residue. We get

$$J = \Gamma\frac{a^2\tau_i(\omega)}{16\pi^2\omega\dot{v}_\parallel^2}\int_{-\pi/b}^{\pi/b} dk_\perp \int_0^{2\pi} d\varphi \left\{1-i\Omega\tau_i - i\frac{\tau_i(\omega)}{\omega} \times E(\mathbf{q},\mathbf{k},\omega)\right\}^{-1},$$

$$E = q_\parallel \dot{v}_\parallel \sqrt{\omega^2-2\omega_\perp^2\sin^2\frac{bk_\perp}{2}} + \omega_\perp^2 \sin\frac{q_\perp b}{2}\sin bk_\perp.$$

We shall use the inequalities (35), which actually identify coherent inverse scattering processes. If these conditions are satisfied, it is possible to find J by expanding in series the expression under the integral in braces, after which it is easy to integrate over $d\varphi$ and $dk_\perp$. We find

$$J = 1+\tau_i(\omega)[i\Omega - R(\mathbf{q},\omega)],$$

$$R(\mathbf{q},\omega) = \frac{\mathbf{q}_\parallel^2\dot{v}_\parallel^2}{2}\left(1-\frac{\omega_\perp^2}{\omega^2}\right)+\frac{\omega_\perp^4}{2\omega^2}\sin^2\frac{q_\perp b}{2}.$$

The result is that

$$U(\mathbf{q};\omega;\Omega) \approx \frac{\omega\omega_\parallel^2}{\tau_i^2(\omega)}\frac{1}{-i\Omega+R(\mathbf{q},\omega)\tau_i(\omega)}. \tag{38}$$

Note that, when the condition $\omega^2 \ll \omega_\perp^2$ is satisfied, the factor $R\tau_i$ has the standard form for a quasi-three-dimensional system and equals $\dot{v}_\parallel^2 q_\parallel^2 \tau_i(\omega)/2 + \dot{v}_\perp^2 q_\perp^2 \tau_i(\omega)$.

4.3. Determining the diffusivity in the weak localization regime

Let us find the principal values of tensor D in a situation in which the condition given by inequalities (35) is satisfied. Using Eqs. (13) and (16), we have

$$\{D_\parallel, D_\perp\} = \{D_\parallel^{(1)}, D_\perp^{(1)}\} - \{D_\parallel^{(2)}, D_\perp^{(2)}\}. \tag{39}$$

where

$$\{D_\parallel^{(1)}, D_\perp^{(1)}\} = \frac{1}{2\pi g(\omega)}\lim_{\Omega\to 0}\sum_{\mathbf{k}} \{2\dot{v}_\parallel^2 k_\parallel^2, \omega_\perp^4 b^2 \times \sin^2(k_\perp b)\}\bar{G}_{\mathbf{k}}^{+}(\omega)\bar{G}_{\mathbf{k}}^{-}(\omega+\Omega), \tag{40}$$

$$\{D_{\parallel}^{(2)},D_{\perp}^{(2)}\}=\frac{1}{2\pi g(\omega)}\lim_{\Omega\to 0}\sum_{\mathbf{q}}U(\mathbf{q};\omega,\Omega)$$
$$\times\sum_{\mathbf{k}}\{2\dot{v}_{\parallel}k_{\parallel}^2,\omega_{\perp}^4b^2\sin^2(k_{\perp}b)\}$$
$$\times\overline{G}_{\mathbf{k}}^{+}(\omega)\overline{G}_{\mathbf{k}}^{-}(\omega+\Omega)\overline{G}_{\mathbf{k}+\mathbf{q}}^{+}(\omega)\overline{G}_{\mathbf{k}+\mathbf{q}}^{-}(\omega+\Omega). \tag{41}$$

The first term of Eq. (39), given by Eq. (40), is the diffusivity neglecting inverse coherent scattering processes. The second term, given by Eq. (41), determines the change of the diffusivity due to the contribution from similar specific interference scattering processes close to defects. In lattices with a strongly anisotropic interatomic force interaction, the summation over $\mathbf{q}$ is bounded above by two small quantities $q_{\parallel,\perp}\approx\pi/l_{\parallel,\perp}(\omega)$. In this case, if $a\approx b$ holds, the flight path is $l_{\parallel,\perp}=\dot{v}_{\parallel,\perp}\tau_i(\omega)$. However, if the unit cell parameters a and b differ strongly, i.e., for $a\ll b$, a situation can occur in which $q_{\perp}\approx\pi/b$.

We restrict ourselves in what follows to the static case, in which $\Omega\to 0$. Let us first consider the frequency region $0<\omega^2<2\omega_{\perp}^2$, in which the density of states is three-dimensional (it corresponds to index 1). Using Eqs. (32) and (38) and taking into account the explicit form of the dispersion law, Eq. (1′), we have

$$\{D_{\parallel}^{(1)},D_{\perp}^{(1)}\}_1\approx\frac{\omega_{\perp}^2\tau_i(\omega)}{2\pi g(\omega)\omega\omega_{\parallel}^2}\left\{2\dot{v}^2G_{\parallel}\left(\frac{\omega^2}{\omega_{\perp}^2}\right),\dot{v}_{\perp}^2G_{\perp}\left(\frac{\omega^2}{\omega_{\perp}^2}\right)\right\}, \tag{42}$$

where

$$G_{\parallel}(x)=(x-1)\arccos(1-x)+\sqrt{2x-x^2},$$
$$G_{\perp}(x)=\arccos(1-x)-(1-x)\sqrt{2x-x^2}.$$

In the limit $\omega\to 0$, taking into account the definition of $g(\omega\to 0)$ given by Eq. (3) and the relationships

$$\arccos x=\arcsin\sqrt{1-x^2},\quad \arcsin x\simeq x+\frac{x^3}{6},|x|<1,$$

we get for the longitudinal component of coefficient D the usual expression of the form

$$\{D_{\parallel}^{(1)},D_{\perp}^{(1)}\}_1\cong\tau_i(\omega)\{\dot{v}_{\parallel}^2/2,\dot{v}_{\perp}^2\}.$$

As for the interference correction to $D_{\parallel}^{(1)}$, as shown in Ref. 11, it is of order

$$1-\frac{3}{2\pi^2g(\omega)\tau_i^2(\omega)\dot{v}_{\parallel}^3}.$$

In the case of a three-dimensional long-wavelength phonon spectrum, the role of such a correction, generally speaking, is negligible. The situation can be otherwise only when the system includes resonant-scattering impurity centers and crossover splitting of the vibrational branch, when the group velocity is close to zero. We considered this case in detail earlier in Ref. 12.

Let us determine the diffusivity D in the frequency region $2\omega_{\perp}^2\leqslant\omega^2<\omega_{\parallel}^2$, in which the vibrational-mode dispersion law is quasi-two-dimensional (this region corresponds to index 2). In order to do this, first, transforming from summation to integration in Eq. (40), we can integrate over $dk_{\parallel}$ by the method of residues and then over $dk_{\perp}$ directly. We have

$$\{D_{\parallel}^{(1)},D_{\perp}^{(1)}\}_2\approx\frac{1}{2\pi g(\omega)}\frac{\tau_i(\omega)}{\omega\omega_{\parallel}}\{2\dot{v}_{\parallel}^2(\omega^2-\omega_{\perp}^2),\omega_{\perp}^4b^2\}. \tag{43}$$

Let us proceed to the interference term of Eq. (39). After carrying out computations similar to those described above, we get

$$\{D_{\parallel}^{(2)},D_{\perp}^{(2)}\}_2\approx\frac{1}{2\pi g(\omega)}\frac{\tau_i^3(\omega)}{2\omega^3\omega_{\parallel}^2}\sum_{\mathbf{q}\leqslant\mathbf{q}_{\parallel},\mathbf{q}_{\perp}}U(\mathbf{q};\omega;0)$$
$$\times\{2\dot{v}_{\parallel}^2(\omega^2-\omega_{\perp}^2),\omega_{\perp}^4b^2\}.$$

Then, using the explicit form of the diffusion vertex given by Eq. (38), we find

$$\{D_{\parallel}^{(2)},D_{\perp}^{(2)}\}_2\approx\frac{1}{2\pi g(\omega)\omega^2}\sum_{\mathbf{q}\leqslant\mathbf{q}_{\parallel},\mathbf{q}_{\perp}}\frac{1}{R(\mathbf{q},\omega)}$$
$$\times\{4\dot{v}_{\parallel}^2(\omega^2-\omega_{\perp}^2),\omega_{\perp}^4b^2\}. \tag{44}$$

Let us obtain an explicit expression for the quantity $\Sigma_{\mathbf{q}}R^{-1}$ that appears in Eq. (44). In the limit $b\ll l_{\perp}$,

$$\sum_{\mathbf{q}\leqslant\mathbf{q}_{\parallel},\mathbf{q}_{\perp}}\frac{1}{R(\mathbf{q},\omega)}=\frac{a^2b}{(2\pi)^2}\int_{-\pi/l_{\perp}}^{\pi/l_{\perp}}dq_{\perp}\int_0^{\pi/l_{\parallel}}dq_{\parallel}^2\left[q_{\parallel}^2\frac{\dot{v}_{\parallel}^2}{2}\right.$$
$$\left.\times\left(1-\frac{\omega_{\perp}^2}{\omega^2}\right)+\frac{\omega_{\perp}^4\tau_i(\omega)}{2\omega^2}\sin^2\frac{q_{\perp}b}{2}\right]^{-1}$$
$$\approx\frac{2a^2}{\pi^2}\left[\dot{v}_{\parallel}^2\tau_i(\omega)\left(1-\frac{\omega_{\perp}^2}{\omega^2}\right)\right]^{-1}$$
$$\times\frac{b}{l_{\perp}}\ln\left(2\sqrt{2}\frac{l_{\perp}}{b}\frac{\omega}{\tau_i\omega_{\perp}^2}\right) \tag{44′}$$

(here and below $l_{\parallel}=l_{\mathrm{ph},\parallel}$, $l_{\mathrm{ph},\perp}=l_{\perp}$). Let us combine the terms of the standard and interference types, Eqs. (43) and (44). We use the representation for $g(\omega)$ in the form of Eq. (4). In this case we also recall that $\tau_i^{-1}(\omega)=(\pi/2)c\epsilon^2\omega^2g(\omega)$. As a result, for the intermediate frequency region, we get the following formula for the diffusivity tensor:

$$\{D_{\parallel},D_{\perp}\}_2=\left\{\frac{\dot{v}_{\parallel}^2\tau_i(\omega)}{2}\left(1-\frac{\omega_{\perp}^2}{\omega^2}\right)\cdot\frac{b^2\omega_{\perp}^4\tau_i(\omega)}{8\omega^2}\right\}F(\omega), \tag{45}$$

$$F(\omega)=1-\frac{2}{\pi}c\epsilon^2\frac{\omega^2}{\omega_{\parallel}^2}\frac{1}{(1-\omega_{\perp}^2/\omega^2)}\frac{b}{l_{\perp}}$$
$$\times\ln\left(\sqrt{2}\pi\frac{l_{\perp}}{b}c\epsilon^2\frac{\omega^4}{\omega_{\perp}^2\omega_{\parallel}^2}\right) \tag{46}$$

(these equations also remain valid when $l_{\perp}\approx b$).

Let us make some comments on the expressions obtained for diffusivity tensor D. It follows directly from Eqs. (45) and (46) that, in the case of strong off-diagonal disorder,

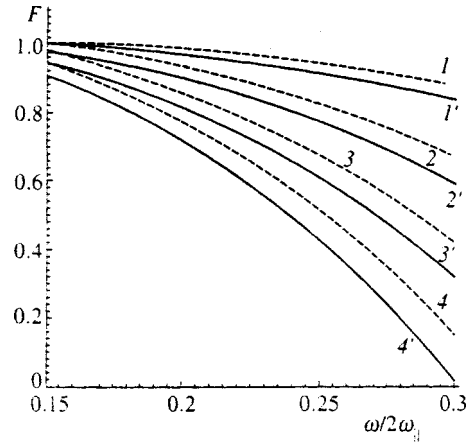

FIG. 2. The factor F vs $\omega/2\omega_{\parallel}$. Curves *1*(*1′*), *2*(*2′*), *3*(*3′*) and *4*(*4′*) correspond to the values of parameter $c\epsilon^2$ equal to 0.25, 0.5, 0.75, and 1, respectively. In the case of the dashed curves, marked with the unprimed indices, the parameter $\omega_{\perp}/\omega_{\parallel}$ equals 0.15. For the solid curves, with the primed indices, $\omega_{\perp}/\omega_{\parallel}=0.1$.

when the parameter $c\epsilon^2$ is of the order of several tens or more, the diffusivity D can vary substantially because of the interference mechanism. We should also emphasize that Eqs. (43), (45), and (46) make it possible to study weak localization processes in the problem of low-temperature thermal conductivity.

In order to illustrate how specific impurity interference processes affect the diffusivity in the weak-localization regime, we numerically calculated the F factors of Eq. (46). We considered the frequency interval in which the frequencies of the vibrational modes depend weakly on the transverse component $k_{\perp}$ of the quasimomentum. It was also assumed that $l_{\mathrm{ph},\perp}\approx b$. The parameter $c\epsilon^2$ was varied from 0.25 to 1 in the calculations. Its value is determined by the measure of disorder. Moreover, the value of the parameter $\omega_{\perp}/\omega_{\parallel}$ was varied. Its value depends on the ratio between the force parameters that characterize the interaction along the z axis and in the xy plane. The results of the calculations are shown in Fig. 2, from which it can be seen that, first, in the case of strong off-diagonal disorder, the partial diffusivities can undergo appreciable variation in regions where the system exhibits quasi-two-dimensional properties. In principle, for $c\epsilon^2>1$ a gap appears in the $D(\omega)$ spectrum. Second, the renormalization of the diffusivity strengthens as the interaction between the layers weakens.

5. THE CASE OF OFF-DIAGONAL ORDER

In this section, we shall consider only the intermediate frequency region $2\omega_{\perp}^2\leqslant\omega^2<\omega_{\parallel}^2$, in which the dispersion law possesses quasi-two-dimensional properties. As pointed out, if off-diagonal disorder occurs in a layered lattice when the local force constants are markedly weakened, a vibrational mode of resonance type arises close to the impurity center. A gap appears in the vibrational spectrum at some critical concentration c_{cr}. We shall analyze the features of the behavior of the diffusivity for $c>c_{\mathrm{cr}}$. We begin with the quasilocal perturbation model.

5.1. Quasilocal perturbation model in the weak-coupling case

We assume that the fluctuational perturbation of the interaction force parameters between an individual substituent impurity and its surroundings extends to the zeroth and first coordination spheres. In the case of a tetragonal lattice considered here, the defect interacts with four atoms (1–4) in the plane of the layer and two atoms (5 and 6) in the adjacent layers. Then the representations $3A_{1g}+2E_u+B_{1g}+A_{2u}$ are available on the occupied impurity site and the neighboring sites 1–6 of the ''impurity quasimolecule.'' The representations $A_{2u}+E_u$ are available on the impurity site itself.

The unitary matrix U for the $A_{2u}+E_u$ representations is defined as

$$U^{+}=\begin{pmatrix}1&0&0&0&0&0&0\\0&w&w&w&w&0&0\\0&0&0&0&0&y&y\end{pmatrix}.$$

It is assumed above that $w=1/\sqrt{4}$ and $y=1/\sqrt{2}$.

Using the explicit form of the matrix U, the matrices V and G^0 can be determined in symmetrized coordinates. First we have

$$V^{+}=\begin{pmatrix}4\tau_{\parallel}\gamma_{\parallel}+2\tau_{\perp}\gamma_{\perp}+M_0\epsilon\omega^2 & -2\tau_{\parallel}\gamma_{\parallel} & 0\\ -2\tau_{\parallel}\gamma_{\parallel} & \tau_{\parallel}\gamma_{\parallel} & 0\\ 0 & 0 & -\tau_{\perp}\gamma_{\perp}\end{pmatrix}. \tag{47}$$

Here it is assumed that

$$\tau_{\parallel}=1-\frac{\gamma_{\parallel}'}{\gamma_{\parallel}},\quad \tau_{\perp}=1-\frac{\gamma_{\perp}'}{\gamma_{\perp}}.$$

We also have

$$\overline{G^0}=\begin{pmatrix}G_{00}^0 & 2G_{01}^0 & 0\\ 2G_{01}^0 & G_{00}^0+G_{12}^0+2G_{13}^0 & 0\\ 0 & 0 & G_{00}^0+G_{56}^0\end{pmatrix}. \tag{48}$$

In order to simplify further, we neglect the interaction between the atomic planes. When this is done, $\gamma_{\perp}=0$. We also use the following relationships between the Green's functions in the case of a planar lattice:

$$4\gamma_{\parallel}(G_{00}^0-G_{01}^0)=1+M_0\omega^2G_{00}^0,$$

$$4\gamma_{\parallel}G_{01}^0-\gamma_{\parallel}(G_{00}^0+G_{12}^0+G_{13}^0)=M_0\omega^2G_{01}^0.$$

In this case, in the low-frequency region $\omega\sim0.1\omega_{\parallel}$,

$$\mathrm{Re}\ G_{00}^0\approx 2\ \mathrm{Re}\ G_{01},\quad \mathrm{Im}\ G_{00}^0\approx \mathrm{Im}\ G_{01}.$$

Substituting Eqs. (47) and (48) into Eq. (26) and using the three last relationships for the Green's functions, we get

$$t_{A_{2u}}=\frac{\hat{t}}{Z},$$

$$\hat{t}=\begin{pmatrix}M_0\omega^2\varepsilon+z\tau\gamma(1-M_0\omega^2\varepsilon G_{01}^0 & \tau\gamma\sqrt{z}(1-M_0\omega^2\varepsilon G_{01}^0)\\ \tau\gamma\sqrt{z}(1-M_0\omega^2\varepsilon G_{01}^0) & -\tau\gamma(1-M_0\omega^2\varepsilon G_{0}^0)\end{pmatrix}$$

$$Z=1-\tau_{\parallel}-M_0\epsilon\omega^2\ \mathrm{Re}\ G_{00}^0(1-\tau_{\parallel}/2)-iM_0\epsilon\omega^2$$

$$\times(1-\tau_{\|})\mathrm{Im}\ G_{00}^{0}. \tag{49}$$

In this case, in the region of low frequencies and small quasimomentum values, we have for the space–time Fourier component of the scattering t matrix given by Eq. (49)

$$t(\mathbf{k},\omega)\approx-\frac{z\gamma_{\|}'\omega^2}{\omega_{\|}^2\gamma_{\|}'/|\epsilon|\,\gamma_{\|}-\omega^2-i\omega^2\sqrt{z}\,\gamma_{\|}'/\gamma_{\|}}+O\left[(k_{\|}a)^4,\epsilon(k_{\|}a)^2\,\frac{\omega^2}{\omega_{\|}^2}\right]. \tag{50}$$

Here z is the number of nearest neighbors in the xy plane.

It can be seen from this that, in the weak-coupling case, in the limit $R=\gamma_{\|}'/\gamma_{\|}\to 0$, the t matrix has a low-frequency resonance. The characteristic frequency is

$$\omega_R=\sqrt{\frac{R}{|\epsilon|}}\,\omega_{\|}.$$

It can also be seen from the formula for the t matrix, Eq. (50), that the condition $\omega_R^2\ll\omega_{\|}^2$ is satisfied and the damping of the level, $\Gamma_R=\sqrt{4}R\omega^2$, is simultaneously small by comparison with the characteristic resonance frequency. The resonance mode is thus well defined.

In what follows, as far as the mass operator is concerned, we recall that the relationship of general form for the t matrix, Eq. (26), includes the matrices V and G. We restrict ourselves to the region of low frequencies and small values of the quasimomentum. In this case, if the indicated matrices are directly multiplied and we transform from the coordinate representation to the momentum representation using Eq. (27), we have (see Ref. 11 for some of the details)

$$\Sigma(\omega,\mathbf{k})=ct(\omega,\mathbf{k})[1+\Delta_c(\omega,\mathbf{k})],$$

$$\Delta_c(\omega,\mathbf{k})=ct^2(\omega,\mathbf{k})\left(-\frac{\partial G_0^0(\omega)}{\partial\omega^2}-G_0^{02}(\omega)\right)+c\sum_{\mathbf{s}\neq 0}\frac{G_{\mathbf{s}}^{02}(\omega,\mathbf{k})t^3(\omega,\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{s}}+G_{\mathbf{s}}^{03}(\omega,\mathbf{k})t^4(\omega,\mathbf{k})}{1-G_{\mathbf{s}}^{02}(\omega)t^2(\omega,\mathbf{k})}\ldots \tag{51}$$

The convergence condition of the series is again given by a relationship of the form of Eq. (29).[19,20]

When the t matrix is of the resonance type, it is necessary to explicitly consider the renormalization of the phonon spectrum and take into account the frequency dependence of factor Q. The configuration-averaged single-particle $G_{\mathbf{k}}^{+}$ function of the $\mathbf{k}$ mode is described by a relationship in the general form of Eq. (19). The Q factor close to and far from resonance was determined above [see Eq. (23)].

5.2. Case of crossover splitting of the vibrational spectrum

In the case of weakly coupled impurity atoms, the t matrix is defined in dimensionless variables as

$$t(x)=-\frac{c\sqrt{z}\Gamma_R}{x_R-x-i\Gamma_R},$$

where

$$\Gamma_R=\sqrt{z}Rx,\quad x=\frac{\omega^2}{\omega_{\|}^2},\quad x_R=\frac{\omega_R^2}{\omega_{\|}^2}.$$

Here $\omega_{\|}$ is the maximum frequency of the vibrations in the plane, and $R=\gamma_{\|}'/\gamma_{\|}\ll 1$.

We shall assume that the defect concentration is relatively high and that the following inequalities are satisfied in some frequency interval close to ω_R:

$$|x-x_R|>\Gamma_R,\quad x_{\mathbf{k}}>|\mathrm{Re}\ ct(x)|. \tag{52}$$

It then turns out that the single-particle Green's function $\overline{G}(\omega)$ in the neighborhood of x_R has not one but two bands of resonance type. In this case, the dispersion curve of the phonons is actually split at frequency x_R, and a band gap appears in the vibrational spectrum.[28,29] Taking into consideration what has been said, the Green's function can be represented as

$$\overline{G}_{1,2}^{+}(\mathbf{k},x)=Q(x)[x-\tilde{x}_{1,2}(\mathbf{k})+i\sqrt{x}/\tau_{R'}]^{-1}.$$

This includes the factor

$$Q(x)=\left(1-\frac{\partial P}{\partial x}\right)^{-1}=\frac{(x_R-x)^2}{(x_R-x)^2+cRx_R}$$

and $\tilde{x}_{1,2}(\mathbf{k})$ denotes the renormalized dispersion law

$$\tilde{x}_{1,2}(\mathbf{k})=\frac{1}{2}\left[x_R+cR+x_{\mathbf{k}}\mp\sqrt{(x_{\mathbf{k}}+x_R+cR)^2-4x_{\mathbf{k}}x_R}\right]$$

$[x(\mathbf{k})=\omega_{\mathbf{k}}^2/\omega_{\|}^2]$. It corresponds to two bands that result from the crossover splitting described above. The upper limit x_- of the first band is close to x_R, while the band gap equals $x_+-x_-=cR$ (x_+ is the lower limit of the second band).

Note that it makes sense to speak of the existence of a gap of $\Delta=cR$ between the two branches if it is greater than the broadening. In other words, the inequalities $\Delta>\Gamma$ or $cR>Rx_R$ must be satisfied. Based on these inequalities, the concentration for which crossover splitting of the spectrum occurs can be estimated. We have

$$c_{cr}\approx x_R=\omega_R^2/\omega_{\|}^2.$$

Note also that, in the case under consideration,

$$\left|t^2(\omega)\,\frac{\partial\dot{G}_{00}(\omega)}{\partial\omega^2}\right|\approx\frac{cR^2x}{(x-x_R)^2}\approx\frac{cR^2x_R}{(cR)^2}=\frac{x_R}{c}=\frac{R}{\epsilon c}.$$

Since we have $c>c_{\mathrm{cr}}$, the condition given by inequality (29) for the convergence of the series in Eq. (51) is satisfied.

Taking into consideration what has been said, when there is simultaneously a well-defined quasilocal state and a gap in the vibrational spectrum, it can be shown that the diffusivity components are determined by a formula like

$$D_{\|,\perp}(\omega)=\dot{D}_{\|,\perp}(\omega)-\sum_{\mathbf{q}}\dot{D}_{\|,\perp}(\omega)\left[-i\Omega+\dot{D}_{\|}^{(l)}(\omega)q_{\|}^2+\dot{D}_{\perp}(\omega)\,\frac{2}{b^2}\sin^2\left(q_{\perp}\,\frac{b}{2}\right)\right]^{-1}, \tag{53}$$

where

$$\dot{D}_{\parallel}(\omega)=\frac{v_{\parallel}^2\tau_{i'}(\omega)}{2},\quad \dot{D}_{\perp}(\omega)=\frac{\omega_{\perp}^4 b^2\tau_{i'}(\omega)}{4\omega^2}\,Q^2(\omega). \tag{53'}$$

The group velocity $v_{\parallel}$ is determined by

$$v_{\parallel}^2(x)=Q^2(x)\left(1-\frac{P(x)}{x}-\frac{x_{\perp}}{x}\right). \tag{54}$$

In this case, the second term of Eq. (53′) in Eq. (53) describes the influence of inverse coherent scattering processes.

Integration can be carried out over dq in Eq. (53). As a result, we get

$$D_{\parallel,\perp}(gq)=\dot{D}_{\parallel,\perp}(\omega)F(\omega). \tag{55}$$

The factor F is given by

$$F(\omega)\approx 1-\frac{a^2}{\pi^2 g(\omega)}\,\frac{1}{v_{\parallel}^2\tau_{i'}(\omega)}\,\ln\frac{\sqrt{2}\pi\omega}{\tau_i(\omega)\omega_{\perp}^2}. \tag{56}$$

Using Eqs. (55) and (56), we now consider the question of localization thresholds. The threshold frequencies ω_{tl} are found by solving the equation

$$F(\omega_{tl})=0.$$

Let us determine the relaxation time, the density of states, and the group velocity that appear in Eq. (56). If we take into account the definition of the t matrix, it is easy to find that

$$\mathrm{Im}\ t(x)=-\frac{zcR^2x^2\sqrt{z}}{(x_R-x)^2+zR^2x^2}.$$

Since we are considering the situation in which inequalities (52) are satisfied, we get, approximately,

$$\mathrm{Im}\ t(x)\approx-\frac{z^{3/2}c(Rx)^2}{(x_R-x)^2}.$$

Now we directly get that

$$\tau_i^{-1}=-\frac{\mathrm{Im}\ t(x)}{\sqrt{z}x}=zcR^2\,\frac{x^{3/2}}{(x_R-x)^2}.$$

The effective relaxation time, taking into account the explicit form of Q, approximately equals

$$\tau_{i'}^{-1}=Q^{-1}(x)\,\tau_i(x)\approx\frac{(x_R-x)^2+cRx_R}{(x_R-x)^2}\,\frac{(x_R-x)^2}{zcR^2x^{3/2}}$$
$$=\frac{(x_R-x)^2+cRx_R}{zcR^2x^{3/2}}. \tag{57}$$

Note that, in the region where quasi-two-dimensional properties are displayed, the correction quadratic in c^2 to the phonon lifetime turns out to be second-degree (and we neglect it), whereas its role is substantial in the three-dimensional case.[11,12] For the density of states in the first region of quasi-two-dimensionality, excluding the gap, when $\omega_{\perp}^2<\omega_R^2\ll 1$, we have approximately

$$\tilde{g}(\omega^2)\simeq\frac{a^2b}{2(2\pi)^2v_{\parallel}^2}\int_{-\pi/b}^{\pi/b}dk_{\perp}\left\{\int_0^{(k_{\parallel}^{\min}v_{\parallel})^2}d(k_{\parallel}\dot{v}_{\parallel}^2)\left[k_{\parallel}^2\dot{v}_{\parallel}^2-\left(\omega^2-\Sigma_R(\omega)-2\omega_{\perp}^2\,\sin^2\frac{k_{\perp}b}{2}\right)\right]\right.$$
$$\left.+\int_{(k_{\parallel}^{\max}v_{\parallel})^2}^{\omega^2}d(k_{\parallel}\dot{v}_{\parallel}^2)\left[k_{\parallel}^2\dot{v}_{\parallel}^2-\left(\omega^2-\Sigma_R(\omega)-2\omega_{\perp}^2\,\sin^2\frac{k_{\perp}b}{2}\right)\right]\right\}\approx\frac{1}{\pi\omega_{\parallel}^2}\left[\eta(\omega_R^2-\omega^2)+\eta(\omega^2-\omega_R^2-cR\omega_{\parallel}^2)\right].$$

By definition, the anomalous dispersion region lies between $k_{\parallel}^{\max}$ and $k_{\parallel}^{\min}$. It corresponds to the forbidden frequency region for the density of vibrational states. Outside this region, in dimensionless variables,

$$\tilde{g}(\sqrt{x})=\frac{\sqrt{x}}{\pi}. \tag{58}$$

Finally, we consider the case $x_R>x_{\perp}=\omega_{\perp}^2/\omega_{\parallel}^2$; i.e., we assume that the frequency ω_R lies in the quasi-two-dimensional region of the spectrum. Therefore, the group velocity is determined from

$$v_{\parallel}^2\approx Q^2(x)\left(1-\frac{P(x)}{x}\right). \tag{54'}$$

We use Eqs. (57), (58), and (54′) for $\tau_{i'}$, $\bar{g}$, and $v_{\parallel}$ and simplify Eq. (56) for F in the frequency region close to the threshold frequency. We have

$$F(x)=1-\frac{z}{\pi}\,\frac{cR^2x}{(x_R-x)^3}\,\frac{(x-x_R)^2+cRx_0}{x_R-x+cR}\times\ln\frac{\sqrt{2}zcR^2x^2}{x_{\perp}(x_R-x)^2}. \tag{59}$$

Using Eq. (59), we show first of all that, far from the gap, i.e., when $x\ll x_R$ and $x\gg x_R+cR$, there is no localization in the case under consideration. Indeed, if $x\ll x_R$ holds, the second term ΔF in the expression for factor F equals

$$|\Delta F|=\left|\frac{zc^2Rx}{\pi x_R^2}\,\ln\frac{zcR^2x^2}{x_{\perp}x_R^2}\right|\ll 1.$$

ΔF obviously goes to zero when $R=\gamma_{\parallel}'/\gamma_{\parallel}\ll 1$. Then for $x\gg x_0+cR$ we get that

$$|\Delta F|=\left|\frac{zc^2R}{\pi x_0}\,\ln\frac{zcR^2}{x_{\perp}}\right|\ll 1.$$

Again $\Delta F\to 0$ when $R\ll 1$. Consequently, the equation $F(\omega_{tl})=0$ is not satisfied in either case.

Let us now analyze the behavior of F close to the edges of the gap in the excitation spectrum. In order to do this, we again transform Eq. (59). We set

$$x_+=x_R+cR+\delta_r,$$

where $\delta_r\ll cR$ is some small frequency interval to the right of the gap. We also set

$$x_- = x_R - \delta_l .$$

Here $\delta_l \ll x_R$ is a small frequency interval to the left of the gap. After some simple computations, we find

$$F(x) = 1 - \begin{cases} \dfrac{zcR^2x^2}{\pi\delta_l^3} \ln \dfrac{zx_-^2 cR}{x_\perp \delta_l^2} \gg 1, & x_- < x_R , \\ \dfrac{zx_+ x_R}{\pi c \delta_r} \ln \dfrac{x_+^2 z}{x_\perp c} > 1, & x_+ > x_R + cR. \end{cases} \tag{60}$$

It immediately follows from the expressions for the diffusivity D given by Eq. (53) and F given by Eq. (60) that regions of localized states appear at the left and right edges of the gap. In this case, the region of localized states at the left edge of the gap is substantially larger that at the right. Note also that, in the case of diagonal disorder, localization effects can in principle appear in a rather wide high-frequency region. If the disorder is off-diagonal, localization occurs close to the edges of the low-frequency gap in the excitation spectrum, as in the case of a three-dimensional lattice.

6. THE NATURE OF THE LOW-TEMPERATURE PLATEAU IN THE THERMAL CONDUCTIVITY OF THE COMPOUNDS BSCCO AND BSYCO

The question of the unusual temperature dependence of the thermal conductivity of a single crystal of BSCCO of complex HTSC systems and its insulating analog BSYCO has been discussed in the literature.[30–32] It is assumed that heat is transported mainly by phonons in these systems in a wide temperature range from 10 to 100 K. The presence of the plateau in the thermal conductivity of BSCCO and BSYCO in the temperature interval $10\ \mathrm{K} \leqslant T \leqslant 30\ \mathrm{K}$ is extremely interesting here [there is no such plateau in the $\kappa(T)$ dependence for 1-2-3 superconductors that are three-dimensional relative to phonon properties, where κ increases monotonically in this temperature range[32]]. In other words, the thermal conductivity has features characteristic of heat transport in amorphous insulators. The authors of an experimental work[33] also point out there is strong cationic disordering in actual crystals because of cationic nonstoichiometry in the layers that contain bismuth and strontium.

At the same time, it is well known that the bismuth cuprates $Bi_2Sr_2Ca_{n-1}Cu_nO_{4+2n}$, where $n = 1,2,3$, possess a crystal lattice with the pseudo-tetragonal symmetry $I4/mmm$ or D_{4h}^{17} and are block structures. They consist of alternating layers of the rock-salt type, BiO–SrO, and the oxygen-defective perovskite Ca–CuO_2. The distinguishing mark of Bi-based compounds is the weak coupling of the Bi–O layers because of the relatively large distance ≈ 30 Å between them. This property makes bismuth compounds similar to mica. In particular, Bush, Kitaev *et al.*[33–35] considered their optical spectra and concluded that the phonon subsystem has a quasi-two-dimensional character. Reference 36, on the basis of an analysis of inelastic incoherent neutron scattering spectra obtained by the isotopic-contrast method, concluded that the interaction between the BiO–SrO and Ca–CuO_2 structural blocks is weak.

It was established in Sections IV and V that substantial renormalization of the diffusivity can occur in low-frequency regions under conditions of strong disorder because of specific interference processes. A plateau can evidently arise in the low-temperature dependence of the thermal conductivity because of such renormalization of the diffusivity.[12]

Thus, taking into account all that has been said, it can be assumed that the phonon subsystem displays two-dimensional properties in the low-temperature region in the heat transport in the compounds BSCCO and BSYCO. Based on the results of this paper, the plateau in the temperature dependence of the thermal conductivity experimentally observed for irregular single crystals of BSCCO and BSYCO can be qualitatively explained as a result of the action of specific impurity interference processes of coherent inverse scattering. The question of the model of the disorder in this case requires special consideration.

7. CONCLUSION

We have discussed the question of the weak localization of the acoustical vibrational modes in a nonideal strongly anisotropic harmonic crystal lattice. It was assumed that the layers of the lattice are weakly coupled to each other and display two-dimensional properties. Based on the Kubo formula for the thermal conductivity, we have obtained an expression for the diffusivity tensor in the weak-localization regime for temperature regions in which the phonon flight paths are determined by elastic scattering at point defects.

The features of the frequency behavior of the diffusivity have been analyzed in the frequency interval where the dispersion law of the acoustic phonon modes possesses two-dimensional properties. Most importantly, we have considered the case of diagonal disorder for which impurity modes of a quasilocal type do not appear in the region where the system manifests quasi-two-dimensional properties. Nevertheless, as it turned out, substantial renormalization of diffusivity D can occur under conditions of strong disorder only because of inverse coherent scattering processes. In this case, renormalization of the diffusivity is strengthened when the interaction between the layers is weakened. When $c\epsilon^2 > 1$, a gap appears in the $D(\omega)$ spectrum at frequencies comparable with the characteristic frequency of the spectrum.

In the case of off-diagonal disorder and weak coupling, there are quasilocal modes in the vibrational spectrum in the region where the system manifests quasi-two-dimensional properties. In other words, a system of resonant-scattering impurity centers appears. We have shown that in this case, if the concentration is greater than a critical value c_{cr}, and a low-frequency gap arises in the vibrational spectrum, regions of localized phonon states arise on the left and on the right edges of such a gap.

We are grateful to N. M. Plakida and H. Schober for support.

APPENDIX A

Above we have discussed questions relating to heat transport associated with standard phonon modes. In the case

of off-diagonal disorder, for $\Delta\Phi\neq 0$, the κ_2 term should also be considered. Keeping in mind what has been said, we consider an expression of the form

$$Y=\sum_{\substack{\boldsymbol{\mu}\mathbf{L}\mathbf{L}'\\ \boldsymbol{\mu}_1\mathbf{L}_1\mathbf{L}_1'}} c_{\boldsymbol{\mu}}c_{\boldsymbol{\mu}_1}(\Delta\Phi\mathbf{R})_{\boldsymbol{\mu}+\mathbf{L},\boldsymbol{\mu}+\mathbf{L}'}(\Delta\Phi\mathbf{R})_{\boldsymbol{\mu}_1+\mathbf{L}_1,\boldsymbol{\mu}_1+\mathbf{L}_1'} \times \mathrm{Im}\, G_{\boldsymbol{\mu}+\mathbf{L},\boldsymbol{\mu}_1+\mathbf{L}_1}\ \mathrm{Im}\, G_{\boldsymbol{\mu}_1+\mathbf{L}_1',\boldsymbol{\mu}+\mathbf{L}'} + \sum_{\substack{\boldsymbol{\mu}\mathbf{L}\mathbf{L}'\\ \mathbf{ss}}} (c_{\boldsymbol{\mu}}\Delta\Phi\mathbf{R})_{\boldsymbol{\mu}+\mathbf{L},\boldsymbol{\mu}+\mathbf{L}'}(\Phi^0\mathbf{R})_{\mathbf{ss}'} \times \mathrm{Im}\, G_{\boldsymbol{\mu}+\mathbf{L}',\mathbf{s}'}\ \mathrm{Im}\, G_{\mathbf{s},\boldsymbol{\mu}+\mathbf{L}}. \tag{A1}$$

For simplicity we restrict ourselves to the single-impurity approximation. We shall assume that the impurity is at the site $\mu=\mathbf{d}$. We multiply the resulting expression by the defect concentration. In this case, instead of Eq. (A1), we have

$$Y=Y_1+Y_2 =c\sum_{\mathbf{L}\ldots\mathbf{L}_1'}(\Delta\Phi\mathbf{R})_{\mathbf{L}\mathbf{L}'}(\Delta\Phi\mathbf{R})_{\mathbf{L}_1\mathbf{L}_1'}\ \mathrm{Im}\,\tilde{G}_{\mathbf{L}\mathbf{L}_1}\ \mathrm{Im}\,\tilde{G}_{\mathbf{L}_1'\mathbf{L}'} + c\sum_{\substack{\mathbf{ss}'\\ \mathbf{L}\mathbf{L}'}}(\Delta\Phi\mathbf{R})_{\mathbf{d}+\mathbf{L},\mathbf{d}+\mathbf{L}'}(\Phi^0\mathbf{R})_{\mathbf{ss}'} \times \mathrm{Im}\,\tilde{G}_{\mathbf{s},\mathbf{d}+\mathbf{L}}\ \mathrm{Im}\,\tilde{G}_{\mathbf{d}+\mathbf{L}',\mathbf{s}'}. \tag{A2}$$

We shall take into account the interaction only between the atoms in the layers, and we restrict ourselves in this case to the case in which only adjacent atoms interact. As far as the first term in Eq. (A2) is concerned, it is possible to directly sum over the sites occupied by the atoms of the impurity quasimolecule. We get

$$Y_1=8c(\tau_\parallel\gamma_\parallel^0 a)^2\ \mathrm{Im}\,\tilde{G}_{\mathbf{dd}}(\mathrm{Im}\,\tilde{G}_{\mathbf{d}+\mathbf{a},\mathbf{d}+\mathbf{a}}-\mathrm{Im}\,\tilde{G}_{\mathbf{d}+\mathbf{a},\mathbf{d}-\mathbf{a}}). \tag{A3}$$

Here a is the lattice constant, and $\mathbf{d}\pm\mathbf{a}$ is the radius vector of one of the nearest neighbors of the impurity. Note that explicit expressions for functions $\bar{G}_{\mathbf{dd}}$ and $\bar{G}_{\mathbf{ds}'}$ can be obtained on the basis of the results given, for example, in Ref. 37.

The expression for term Y_2 is also simplified. Using the explicit form of the matrix $(\Delta\Phi R)_{\mathbf{d}+\mathbf{L},\mathbf{d}-\mathbf{L}}$, it is possible to sum over sites $\mathbf{L}\mathbf{L}'$. We have

$$Y_2=4c\tau_\parallel\gamma_\parallel^0 a\sum_{\mathbf{ss}'}(\Phi^0\mathbf{R})_{\mathbf{ss}'}\ \mathrm{Im}\,\tilde{G}_{\mathbf{ds}'}(\mathrm{Im}\,\tilde{G}_{\mathbf{d}+\mathbf{a},\mathbf{s}}-\tilde{G}_{\mathbf{d}-\mathbf{a},\mathbf{s}}).$$

Since $(\Phi^0\mathbf{R})_{\mathbf{ss}'}\mathrm{Im}\,\bar{G}_{\mathbf{ds}'}\sim(1-\delta_{\mathbf{ss}'})(-1)^{\mathbf{s}'-\mathbf{s}}$, we have $Y_2=0$. The contribution to the thermal conductivity denoted as κ_2 is thus determined by term Y_1 in Eq. (A3).

In the low-temperature limit, only the contribution from low frequencies is substantial. In the quasi-two-dimensional case, as $\omega\to 0$, we have $\mathrm{Im}\,\bar{G}_{\mathbf{dd}}\sim\mathrm{const}$, $\mathrm{Im}[\bar{G}_{\mathbf{d}+\mathbf{a},\mathbf{d}+\mathbf{a}}(\omega)-\bar{G}_{\mathbf{d}+\mathbf{a},\mathbf{d}-\mathbf{a}}(\omega)]\sim\omega^2$; consequently, $Y_1\sim\omega^2$. As a result, we find that

$$\kappa_2\sim ca\dot{v}_\parallel\tau_\parallel^2\gamma_\parallel(T/\theta_\parallel)^3$$

($\theta_\parallel$ is the Debye temperature of a planar lattice). Note that the quantity κ_2 is small by comparison with

$$\kappa_1\sim\frac{\dot{v}_\parallel L}{2}\left(\frac{T}{\theta_\parallel}\right)^2,$$

which is determined by the boundary scattering (L is the characteristic geometrical size of the sample).

APPENDIX B

In general, the expression for the mass operator can be represented as

$$\Sigma(\omega)=\mathrm{Re}\,\Sigma(\omega)+i\ \mathrm{Im}\,\Sigma(\omega)=|\Sigma(\omega)|e^{i\Phi(\omega)}.$$

where

$$|\Sigma(\omega)|=(\sqrt{[\mathrm{Re}\,\Sigma(\omega)]^2+[\mathrm{Im}\,\Sigma(\omega)]^2},$$

$$\Phi(\omega)=\arctan\frac{\mathrm{Im}\,\Sigma(\omega)}{\mathrm{Re}\,\Sigma_R(\omega)}. \tag{B1}$$

Thus, for the real part of the mass operator we have

$$P(\omega)=\Sigma_R(\omega)=|\Sigma(\omega)|\cos\Phi. \tag{B2}$$

If the frequency ω is close to the frequency of the phonon mode with quasimomentum $\bar{\omega}_{\mathbf{k}}$, the operator P is determined by a formula of the form

$$P(\omega)=P(\tilde{\omega}_{\mathbf{k}})+(\omega^2-\tilde{\omega}_{\mathbf{k}}^2)\frac{\partial P(\omega)}{\partial\omega^2}. \tag{B3}$$

Using Eqs. (B1) and (B2), instead of Eq. (B3), we have

$$P(\omega)=P(\tilde{\omega}_{\mathbf{k}})+\left(\omega^2-\tilde{\omega}_{\mathbf{k}}^2\left[\frac{\partial|\Sigma(\omega)|}{\partial\omega^2}\cos\Phi-\left|\Sigma(\omega)\right|\frac{\partial\Phi}{\partial\omega^2}\sin\Phi\right]=P(\tilde{\omega}_{\mathbf{k}})+(\omega^2-\tilde{\omega}_{\mathbf{k}}^2)\times\left[\frac{\partial|\Sigma(\omega)|}{\partial\omega^2}\cos\Phi-\frac{\partial\Phi}{\partial\omega^2}\mathrm{Im}\,\Sigma(\omega)\right]. \tag{B4}$$

Based on Eq. (B4) and the determination of the single-particle Green's function given by Eq. (15), it is possible to directly obtain an expression for factor Q in the form of Eq. (21).

*)E-mail: ZHERNOV@KURM.POLUN.KIAE.SU

[1] B. L. Altshuler, A. G. Aronov, A. G. Khmelnitskii, and A. I. Larkin, *Quantum Theory of Solids*, ed. I. M. Lifshits (Mir, Moscow, 1992).
[2] B. Kramer and A. MacKinnon, Rep. Prog. Phys. **56**, 1479 (1993).
[3] G. V. Olemskiĭ, Usp. Fiz. Nauk **165**, 152 (1996).
[4] V. L. Kuz'min and V. P. Romanov, Usp. Fiz. Nauk **166**, 248 (1996).
[5] M. P. van Albada, J. F. de Boer, and A. Lagendijk, Phys. Rev. Lett. **64**, 3132 (1991).
[6] S. Kawato, T. Hattori, T. Takemori, and H. Nakatsuka, Phys. Rev. B **49**, 90 (1994).
[7] B. A. van Tiggelen, A. Lagendijk, M. P. van Albada, and A. Tip, Phys. Rev. B **45**, 12 233 (1992).
[8] B. A. van Tiggelen, A. Lagendijk, and A. Tip, Phys. Rev. Lett. **71**, 1284 (1993).
[9] Qian-Jin Chu and Zhao-Oing Zhag, Phys. Rev. B **38**, 4906 (1988); Phys. Rev. B **72**, 234 (1992).
[10] P. Allen and J. L. Feldman, Phys. Rev. B **48**, 12 581 (1993).

[11] A. P. Zhernov, E. I. Salamatov, and E. P. Chulkin, Phys. Status Solidi B **165**, 355 (1991).
[12] A. P. Zhernov, E. I. Salamatov, and E. P. Chulkin, Phys. Status Solidi B **168**, 81 (1991).
[13] Ping Sheng, Minyao Zhou, and Zhao-Qing Zhang, Phys. Rev. Lett. **72**, 234 (1994).
[14] E. Akkermans and R. Maynard, Phys. Rev. B **32**, 7860 (1985).
[15] A. P. Zhernov and E. P. Chulkin, Fiz. Tverd. Tela **36**, 2302 (1994) [Phys. Solid State **36**, 1254 (1994)].
[16] A. P. Zhernov and E. P. Chulkin, Phys. Status Solidi B **193**, 67 (1996).
[17] A. P. Zhernov and E. P. Chulkin, Zh. Éksp. Teor. Fiz. **109**, 602 (1996) [JETP **82**, 321 (1996)].
[18] M. A. Ivanov and Yu. V. Skripnik, Fiz. Tverd. Tela (Leningrad) **32**, 2965 (1990) [Sov. Phys. Solid State **32**, 1722 (1990)].
[19] M. A. Ivanov and Yu. V. Skripnik, Ukr. Fiz. Zh. **35**, 1856 (1990).
[20] M. A. Ivanov, A. M. Kosevich, E. S. Syrkin, Yu. V. Skripnik, I. A. Gospodarev, and S. B. Feodos'ev, Fiz. Nizk. Temp. **19**, 434 (1993) [Low Temp. Phys. **19**, 305 (1993)].
[21] A. M. Kosevich, *Theory of the Crystal Lattice* (Visha Shkola, Khar'kov, 1988).
[22] H. Bettger, in *Principles of the Theory of Lattice Dynamics* (Akademie-Verlag, Berlin, 1983).
[23] R. Berman, *Thermal Conduction in Solids* (Clarendon Press, Oxford, 1976; Mir, Moscow, 1979).
[24] J. R. Taylor, *Scattering Theory: The Quantum Theory of Nonrelativistic Collisions* (Wiley, New York, 1972; Mir, Moscow, 1979).
[25] M. A. Ivanov, Fiz. Tverd. Tela (Leningrad) **12**, 1285 (1970) [Sov. Phys. Solid State **12**, 1009 (1970)].
[26] M. A. Ivanov and Yu. G. Pogorelov, Fiz. Nizk. Temp. **5**, 1979 (1979) [*sic*].
[27] A. P. Zhernov and T. A. Memedov, Fiz. Met. Metalloved. **62**, 457 (1986).
[28] M. A. Ivanov, Fiz. Tverd. Tela **12**, 1895 (1970) [Sov. Phys. Solid State **12**, 1508 (1971)].
[29] A. P. Zhernov and T. N. Kudagina, Fiz. Tverd. Tela (Leningrad) **17**, 1459 (1975) [Sov. Phys. Solid State **17**, 941 (1975)].
[30] P. B. Allen, X. Du, L. Mihaly, and L. Forro, Phys. Rev. B **49**, 9073 (1994).
[31] V. B. Efimov and L. P. Mezhov-Deglin, JETP Lett. **62**, 952 (1995).
[32] A. N. Taldenkov, A. V. Inyushkin, and S. Yu. Shabanov, Sverkhprovodnost' (KIAE) **7**, 1502 (1994).
[33] A. A. Bush, I. N. Goncharuk, and Yu. É. Kitaev, Zh. Éksp. Teor. Fiz. **102**, 1587 (1992) [Sov. Phys. JETP **75**, 857 (1992)].
[34] Yu. Kitaev, M. F. Limonov, and A. G. Panfilov, Phys. Rev. B **49**, 9933 (1994).
[35] Yu. É. Kitaev, I. F. Limonov, and A. P. Mirgorodskiĭ, Fiz. Tverd. Tela **36**, 865 (1994) [Phys. Solid State **36**, 475 (1994)].
[36] P. P. Parshin, M. G. Zemlyanov, and A. V. Irodova, Fiz. Tverd. Tela **32**, 2134 (1996) [*sic*].
[37] A. P. Zhernov and G. R. Augst, Fiz. Tverd. Tela (Leningrad) **9**, 2196 (1967) [Sov. Phys. Solid State **9**, 1724 (1968)].

Translated by W. J. Manthey

Some aspects of the nonlocal electrodynamics of infinite Josephson junctions

Z. D. Genchev*) and V. I. Vas'kivs'kiĭ

Institute of Electronics, Bulgarian Academy of Sciences, 1784 Sofia, Bulgaria
(Submitted 29 March 1997)
Zh. Éksp. Teor. Fiz. **113**, 955–966 (March 1998)

We derive the basic equations of nonlocal Josephson electrodynamics, valid for any ratios of the characteristic scale of phase-difference variations and junction thickness. The spectrum of generalized Swihart waves is obtained. We also study the effect of finite surface resistance of the superconducting electrodes on the dynamics of vortex structures characteristic of nonlocal electrodynamics. © *1998 American Institute of Physics.* [S1063-7761(98)01303-1]

1. INTRODUCTION

Josephson junctions with a high value of the critical current density cannot be described by the conventional sine–Gordon equation when $\lambda_j(d) \leqslant \lambda$, where

$$\lambda_j(d) = \sqrt{\frac{\Phi_0}{2\pi\mu_0 j_c(2\lambda+2d)}} \tag{1}$$

is the Josephson length, λ is the London magnetic-field penetration depth, and $2d$ is the thickness of the intermediate nonsuperconducting layer, with $\Phi_0 = h/2|e|$ the magnetic flux quantum, and j_c the homogeneous critical current density. Josephson vortices ($\lambda \ll \lambda_j$), which correspond to small values of j_c,

$$j_c \ll j_l \approx \frac{\Phi_0}{4\pi\mu_0\lambda^3}, \quad d \ll \lambda \tag{2}$$

(here we use the terminology introduced by Gurevich[1]), have been studied for a long time.[2] For Abrikosov–Josephson vortices,[1] the characteristic spatial scale of variation of the phase difference φ is much smaller than the London depth λ, and the opposite condition is met:

$$j_c \gg j_l. \tag{3}$$

Several researchers[3–8] found that in this case the magnetostatics and electrodynamics of a Josephson junction become spatially nonlocal. In Refs. 3–8 the effect of the normal dc conductivity of London superconductors was ignored and no detailed analysis of the role of the finite thickness of the normal layer in the junction was carried out. Allowing for these two factors, in Sec. 2 we derive an integro-differential equation (Eq. (24)) that contains both temporal and spatial nonlocalities. Using this equation, in Sec. 3 we analyze the spectral properties of a Josephson junction in the linear approximation. To calculate the maximum normal-layer thickness below which nonlocality effects manifest themselves, in the Appendix we estimate the critical current density for an idealized model of the normal layer in the compound $YBa_2Cu_3O_7$. Finally, in Sec. 4 we study the dynamics of a 4π kink in a thin ($d \to +0$) tunneling junction with current, where we allow for two dissipation channels, quasiparticle tunneling and the surface resistance of the electrodes.

2. EQUATIONS OF NONLOCAL JOSEPHSON ELECTRODYNAMICS

In this section we derive the basic equations of Josephson electrodynamics, which hold for a tunneling junction of arbitrary thickness. We take the simplest geometry: two superconducting half-spaces ($|x| > d$) separated by a nonsuperconducting layer ($-d \leqslant x \leqslant d$). The system is assumed to be homogeneous in y, or $\partial/\partial y = 0$. We introduce the following Fourier representation of an arbitrary function:

$$f(x,z,t) = \int_{-\infty}^{\infty} d\omega \int_{-\infty}^{\infty} dk\ e^{-i\omega t + ikz} f(\omega,k,x)$$

$$= \int_{-\infty}^{\infty} d\omega\ f(\omega,x,z) e^{-i\omega t}. \tag{4}$$

We assume that a magnetic field directed along the y axis satisfies the boundary condition $H(x=d,z,t) = H(x=-d,z,t)$. The solution of Maxwell's equations in the superconductors,

$$-\frac{\partial H(\omega,x,z)}{\partial z} = \sigma_s(\omega) E_x(\omega,x,z), \tag{5a}$$

$$\frac{\partial H(\omega,x,z)}{\partial x} = \sigma_s(\omega) E_z(\omega,x,z), \tag{5b}$$

$$\frac{\partial E_x(\omega,x,z)}{\partial z} - \frac{\partial E_z(\omega,x,z)}{\partial x} = i\mu_0\omega H(\omega,x,z), \tag{5c}$$

where

$$\sigma_s(\omega) = \sigma_{dc} + \frac{i}{\omega\mu_0\lambda^2}, \quad \sigma_{dc} = \lim_{\omega\to 0} \mathrm{Re}\,[\sigma_s(\omega)], \tag{5d}$$

with $\sigma_s(\omega)$ the complex conductivity and μ_0 the permeability of free space, can be written

$$H(\omega,x,z) = \int_{-\infty}^{\infty} dk\ H(\omega,k) e^{ikz} \times \begin{cases} \exp[-V_s(\omega,k)(x-d)], & x > d, \\ \exp[V_s(\omega,k)(x+d)], & x < -d, \end{cases} \tag{6}$$

$$E_x(\omega,x,z)=\int_{-\infty}^{\infty}dk\,\frac{-ik}{\sigma_s(\omega)}H(\omega,k)e^{ikz}\times\begin{cases}\exp[-V_s(\omega,k)(x-d)], & x>d,\\ \exp[V_s(\omega,k)(x+d)], & x<-d,\end{cases}\tag{7}$$

$$E_z(\omega,x,z)=\int_{-\infty}^{\infty}dk\,\tilde{\eta}(\omega,k)H(\omega,k)e^{ikz}\times\begin{cases}\exp[-V_s(\omega,k)(x-d)], & x>d,\\ \exp[V_s(\omega,k)(x+d)], & x<-d,\end{cases}\tag{8}$$

where

$$V_s(\omega,k)=\sqrt{k^2-i\omega\mu_0\sigma_s(\omega)},\tag{9}$$

Re $V_s>0$, and $\tilde{\eta}(\omega,k)=R_s-iX_s$ is the surface impedance,

$$\tilde{\eta}(\omega,k)=\frac{V_s(\omega,k)}{\sigma_s(\omega)},\tag{10}$$

which, as Eqs. (5d), (9), and (10) imply, has the low-frequency approximation

$$\tilde{\eta}(\omega,k)=-i\omega\mu_0\lambda\,\eta(\omega,k),$$
$$\eta(\omega,k)=\sqrt{1+(k\lambda)^2}\left[1+i\,\frac{\omega}{2}\mu_0\lambda^2\sigma_{dc}\frac{1+2(k\lambda)^2}{1+(k\lambda)^2}\right].\tag{11}$$

The function $H(\omega,k)$ can be found by solving the appropriate set of Maxwell's equations for the tunneling layer ($|x|<d$):

$$\nabla\times\mathbf{E}=-\mu_0\frac{\partial\mathbf{H}}{\partial t},\tag{12}$$

$$\nabla\times\mathbf{H}=\mathbf{i}[j_c\sin\varphi(z,t)+j_{QP}(V)]+\sigma_n\mathbf{E}+\varepsilon_0\varepsilon_r\frac{\partial\mathbf{E}}{\partial t},\tag{13}$$

where $\mathbf{i}$ is the unit vector along the x axis, ε_0 is the permittivity of free space, and $\varepsilon_r>0$ is the dielectric constant (relative permittivity). The first term in square brackets in (13) is the Josephson current density, which depends on the phase difference $\varphi(z,t)$ of the macroscopic wave function of paired particles in the two superconducting regions, while $j_{QP}(V)$ is the tunneling-quasiparticle current density, which depends only on the applied potential difference[2]

$$V\equiv V(z,t)=-\int_{-d}^{d}E_x(x,z,t)dx=-\frac{\Phi_0}{2\pi}\frac{\partial\varphi}{\partial t}.\tag{14}$$

This is the well-known Josephson relation. We now show that all components of the electric and magnetic fields can be expressed in terms of the phase-difference function $\varphi(z,t)$. We start by introducing the operation of averaging over x for an arbitrary function $f(x,z,t)$:

$$f(x,z,t)=\langle f(z,t)\rangle+\tilde{f}(x,z,t),$$
$$\langle f(z,t)\rangle=\frac{1}{2d}\int_{-d}^{d}f(x,z,t)dx.\tag{15}$$

Then Eqs. (12) and (13) reduce to

$$-\frac{\partial\langle H\rangle}{\partial z}=j_c\sin\varphi+j_{QP}(V)+\frac{\Phi_0\varepsilon_0\varepsilon_r}{4\pi d}\frac{\partial^2\varphi}{\partial t^2}+\frac{\Phi_0\sigma_n}{4\pi d}\frac{\partial\varphi}{\partial t},\tag{16a}$$

$$-\frac{\partial\tilde{H}}{\partial z}=\sigma_n\tilde{E}_x+\varepsilon_0\varepsilon_r\frac{\partial\tilde{E}_x}{\partial t},\tag{16b}$$

$$-\mu_0\frac{\partial\tilde{H}}{\partial t}=\frac{\partial\tilde{E}_x}{\partial z}+\frac{\tilde{E}_z(x=d)}{d}-\frac{\partial\tilde{E}_z}{\partial x},\tag{16c}$$

$$\frac{\partial\tilde{H}}{\partial x}=\sigma_n\tilde{E}_z+\varepsilon_0\varepsilon_r\frac{\partial\tilde{E}_z}{\partial t},\tag{16d}$$

$$-d\mu_0\frac{\partial\langle H\rangle}{\partial t}=\frac{\Phi_0}{4\pi}\frac{\partial^2\varphi}{\partial z\partial t}-\tilde{E}_z(x=d,z,t).\tag{16e}$$

In deriving Eqs. (16) we employed the fact that $\langle E_z\rangle=0$ and the boundary conditions $\tilde{H}(x=d)=\tilde{H}(x=-d)$ and $\tilde{E}_z(x=d)=-\tilde{E}_z(x=-d)$, which correspond to the description of the electromagnetic field outside the tunneling barrier given in Eqs. (6)–(8). Note that these assumptions do not hold when the materials of the superconductors differ.

The solution of Eqs. (16b)–(16e) for $\tilde{H}$, $\tilde{E}_x$, and $\tilde{E}_z$ can be written

$$\tilde{H}(x,z,t)=\int_{-\infty}^{\infty}\!\!\int d\omega\,dk e^{-i\omega t+ikz}k\,\frac{\sigma_n-i\varepsilon_0\varepsilon_r\omega}{V_n}\times\left[\cosh(V_nx)-\frac{\sinh(V_nd)}{V_nd}\right]A(\omega,k),$$
$$\tilde{E}_x(x,z,t)=\int_{-\infty}^{\infty}\!\!\int d\omega\,dk e^{-i\omega t+ikz}\frac{-ik^2}{V_n}\times\left[\cosh(V_nx)-\frac{\sinh(V_nd)}{V_nd}\right]A(\omega,k),$$
$$\tilde{E}_z(x,z,t)=\int_{-\infty}^{\infty}\!\!\int d\omega\,dk e^{-i\omega t+ikz}k\sinh(V_nx)A(\omega,k),\tag{17}$$

where

$$V_n\equiv V_n(\omega,k)=\sqrt{k^2-i\omega\mu_0\sigma_n-\frac{\varepsilon_r}{c^2}\omega^2},\quad \text{Re } V_n>0,\tag{18}$$

with σ_n the conductivity of the normal layer. The continuity of the tangential components of the electromagnetic field at the boundaries $x=\pm d$ implies that

$$A(\omega,k)=\frac{-i\omega\mu_0\lambda\,\eta}{k\sinh(V_nd)}H(\omega,k)\tag{19}$$

and

$$H(\omega,k)=\frac{-ik\Phi_0\varphi(\omega,k)}{4\pi\mu_0}\left\{d+\lambda\,\eta\left[1+\frac{k_0^2\varepsilon_n(\omega)}{V_n^2}\times[1-V_nd\coth(V_nd)]\right]\right\}^{-1},\tag{20}$$

where $k_0=\omega/c$, $\varepsilon_n(\omega)=\varepsilon_r+i\sigma_n/\omega\varepsilon_0$, and

$$\varphi(\omega,k)=\int_{-\infty}^{\infty}\int \frac{dtdz}{(2\pi)^2}e^{i\omega t-ikz}\varphi(z,t), \tag{21}$$

and we also find that

$$\langle H(z,t)\rangle=-\frac{\Phi_0}{2\pi\mu_0}\int_{-\infty}^{\infty}\int dt'dz'Q(z-z',t-t')\frac{\partial\varphi(z',t')}{\partial z'} \tag{22}$$

and

$$Q(z,t)=\int_{-\infty}^{\infty}\int \frac{d\omega dk}{8\pi^2 d}e^{-i\omega t+ikz}\frac{d+\lambda\,\eta(\omega,k)\dfrac{k_0^2\varepsilon_n(\omega)}{V_n^2(\omega,k)}[1-dV_n(\omega,k)\coth(dV_n(\omega,k))]}{d+\lambda n(\omega,k)\left\{1+\dfrac{k_0^2\varepsilon_n(\omega)}{V_n^2(\omega,k)}[1-dV_n(\omega,k)\coth(dV_n(\omega,k))]\right\}}. \tag{23}$$

The relationship between the magnetic field and the phase difference generally proves to be nonlocal in both space and time. From Eq. (16a) we obtain a generalized sine–Gordon equation for the phase difference in the form of the following integro-differential equation:

$$\sin\varphi+\frac{\alpha}{\omega_j}\frac{\partial\varphi}{\partial t}+\frac{1}{\omega_j^2}\frac{\partial^2\varphi}{\partial t^2}=2\lambda_j^2\lambda\frac{\partial}{\partial z}\int_{-\infty}^{\infty}\int dt'dz'Q(z-z',t-t')\frac{\partial\varphi(z',t')}{\partial z'}, \tag{24}$$

where $\lambda_j^2=\Phi_0/4\pi\mu_0 j_c\lambda$ is the square of the Josephson length (1) corresponding to the limit $d\to+0$, $\omega_j^2=4\pi d j_c/\Phi_0\varepsilon_0\varepsilon_r$, $\alpha=(\sigma_{QP}+\sigma_n)/\omega_j\varepsilon_0\varepsilon_r$, and in (16a) we introduced a linearization of the quasiparticle tunneling current density, $j_{QP}(V)\approx\sigma_{QP}\langle E_x\rangle$.

Formally this equation differs from those used in Refs. 1, 3, and 8 only by the presence of time integrals and a more complicated expression for the kernel.

3. SPECTRAL PROPERTIES OF A JOSEPHSON JUNCTION

We consider the limit of the linear theory of Swihart waves,[2,9] corresponding to small perturbations of the phase difference, which makes it possible to use the approximation $\sin\varphi\approx\varphi$. In this limit we seek the solution of the linearized variant of Eq. (24) in the form $\varphi\propto\exp(-i\omega t+ikz)$. By introducing the notation

$$\delta=\frac{d}{\lambda},\quad q=k\lambda_j,\quad \Omega=\frac{\omega}{\omega_j},\quad \varepsilon=\frac{\lambda_j}{\lambda} \tag{25}$$

and ignoring damping we obtain a dispersion relation for Swihart waves:

$$(\Omega^2-1)\left\{\delta+\sqrt{1+\left(\frac{q}{\varepsilon}\right)^2}\left[1+\frac{\delta^3\Omega^2}{\varepsilon^2}G(V_d)\right]\right\}=q^2\left\{1+\frac{\delta^2\Omega^2}{\varepsilon^2}\sqrt{1+\left(\frac{q}{\varepsilon}\right)^2}G(V_d)\right\}, \tag{26}$$

$$G(V_d)=\begin{cases}\dfrac{1}{V_d^2}(1-V_d\coth V_d), & \\ & \text{if } q^2>\Omega^2\delta,\\ V_d^2=\dfrac{\delta^2}{\varepsilon^2}(q^2-\Omega^2\delta) & \\ -\dfrac{1}{3} & \text{if } q^2=\Omega^2\delta,\\ -\dfrac{1}{V_d^2}(1-V_d\cot V_d), & \\ & \text{if } q^2<\Omega^2\delta.\\ V_d^2=\dfrac{\delta^2}{\varepsilon^2}(\Omega^2\delta-q^2) & \end{cases} \tag{27}$$

The value of the wave vector at the point of inflection of the dispersion curve can be found from the equation

$$C_1^2x^3+(C_1^2-2C_1C_2)x^2+(C_2^2-2C_1C_2)x+C_2^2(1-\delta^2)=0, \tag{28}$$

where

$$x=\left(\frac{q}{\varepsilon}\right)^2=(k\lambda)^2,\quad C_1=1+\frac{\delta^3}{3\varepsilon^2},\quad C_2=\frac{\delta}{\varepsilon^2}.$$

The value of the frequency Ω at $k=0$ coincides with that obtained earlier by Silin.[9] But the asymptotic behavior in the short-wave limit ($q^2\gg\Omega^2\delta$ and $q\gg\varepsilon$) yields $\Omega^2\to\varepsilon^2/\delta$ as $q\to+\infty$, which leads to the following asymptotic value of the frequency:

$$\overline{\omega}=\frac{c}{\lambda\sqrt{\varepsilon_r}}. \tag{29}$$

Numerical calculations of the $\Omega(q)$ dependence for certain values of ε and δ are depicted in Figs. 1 and 2. The

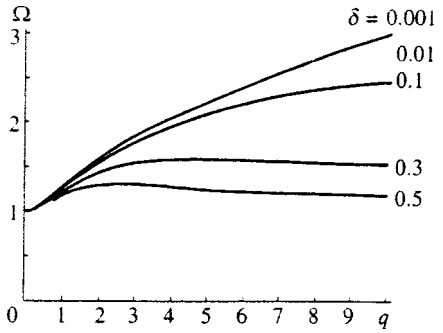

FIG. 1. The $\Omega(q)$ dependence at $\varepsilon=0.8$ for various values of δ.

feasibility of achieving high critical current densities ($\varepsilon \leqslant 1$) by varying the normal-layer thickness in the interval $0 \leqslant \delta \leqslant \delta_m$ is demonstrated in the Appendix.

Allowing for these features of the spectral properties of the Josephson junction, we obtain the following prerequisite of the low-frequency approximation:

$$\max(\omega_j, \bar{\omega})\mu_0\lambda^2\sigma_{dc} \ll 1. \tag{30}$$

The expression (23) for the kernel can be simplified considerably only for thin junctions with $|V_n|d \ll 1$. This is true if

$$kd \ll 1, \quad \frac{\omega}{\bar{\omega}}\frac{d}{\lambda} \ll 1. \tag{31}$$

Hence for this approximation to be valid for the entire spectrum, d must be much less than λ. Then, if conditions (30) and (31) are met and we allow for the fact that $\lim_{x\to 0}[(1 - x\coth x)/x^2] = -1/3$, we arrive at the following expression for the kernel:

$$Q(\omega,k) = \frac{1}{2}\frac{1 - i\omega\tau(k)}{d + \lambda\sqrt{1+(k\lambda)^2}}, \tag{32}$$

where

$$\tau(k) = \mu_0\lambda^3\left\{\frac{\sigma_{dc}}{2}\frac{1+2(k\lambda)^2}{1+(k\lambda)^2} + \frac{\sigma_n d}{3\lambda}\sqrt{1+(k\lambda)^2}\right\}\frac{\sqrt{1+(k\lambda)^2}}{d+\lambda\sqrt{1+(k\lambda)^2}}.$$

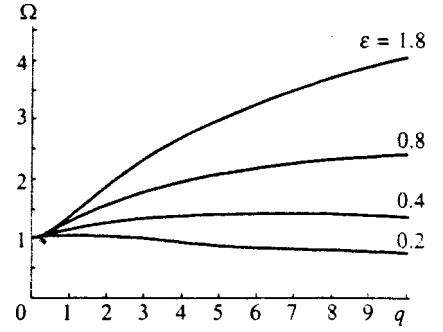

FIG. 2. The $\Omega(q)$ dependence at $\delta=0.1$ for various values of ε.

The real part of the right-hand side of Eq. (32) coincides with an expression used earlier in studies of nonlocal effects in Josephson electrodynamics.[1,8]

4. TRAVELING 4π KINK IN A TUNNELING JUNCTION WITH CURRENT

In this section we discuss some of the corollaries of the general theory for a thin ($d \to +0$) tunneling junction with current. On the right-hand side of Eq. (24) we introduce a dimensionless homogeneous transport-current density j_{dc}/j_c (overlap geometry[2]). In the case of a thin junction, Eqs. (23), (24) and (32) yield

$$\frac{\partial^2\varphi}{\partial t^2} + \alpha\frac{\partial\varphi}{\partial t} + \sin\varphi - \gamma = \int_{-\infty}^{\infty} dz'\left\{\int_{-\infty}^{\infty}\frac{dk}{2\pi V(k)}e^{ik(z-z')} \times\left(1+\beta\frac{\varepsilon^2+2k^2}{\varepsilon^2+k^2}\frac{\partial}{\partial t}\right)\frac{\partial^2\varphi(x',t)}{\partial z'^2}\right\}. \tag{33}$$

Here and in what follows we use dimensionless variables with the substitutions

$$\omega_j t \to t, \quad \frac{z}{\lambda_j} \to z, \quad \frac{\omega}{\omega_j} \to \omega, \quad \lambda_j k \to k$$

and introduce the following notation:

$$V(k) = \sqrt{1+\left(\frac{k}{\varepsilon}\right)^2}, \quad \beta = \frac{1}{2}\omega_j\mu_0\lambda^2\sigma_{dc}, \quad \gamma = \frac{j_{dc}}{j_c}.$$

The parameter β allows for the finite surface resistance of the superconducting electrodes, and ε is defined in (25). Using the integral representation of the Bessel function of the second kind,

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{dke^{ikz}}{\sqrt{1+(k/\varepsilon)^2}} = \frac{\varepsilon}{\pi}K_0(\varepsilon|z|), \quad \varepsilon > 0, \tag{34}$$

we calculate the integrals in (33) and finally obtain

$$\frac{\partial^2\varphi}{\partial t^2} + \alpha\frac{\partial\varphi}{\partial t} + \sin\varphi - \gamma = \frac{\varepsilon}{\pi}\int_{-\infty}^{\infty}dz'\left\{\left(K_0(\varepsilon|z-z'|) + \beta[2K_0(\varepsilon|z-z'|) + \varepsilon|z-z'|K_0'(\varepsilon|z-z'|)]\frac{\partial}{\partial t}\right)\frac{\partial^2\varphi(z',t)}{\partial z'^2}\right\}. \tag{35}$$

Here a prime denotes the derivative of a function with respect to its argument, so that

$$\left(1+\varepsilon\frac{\partial}{\partial\varepsilon}\right)\varepsilon K_0(\varepsilon|z|) = \varepsilon[2K_0(\varepsilon|z|) + \varepsilon|z|K_0'(\varepsilon|z|)]. \tag{36}$$

To analyze the dynamics of traveling vortices (whose velocity $v = c\bar{\nu} = \lambda_j\omega_j\nu$ has yet to be found) in a Josephson junction, we limit our discussion to solutions of the form

$$\varphi(z,t) = \phi(\zeta = z - \nu t). \tag{37}$$

In view of the Josephson relation (14), the traveling wave (37) carries information about the voltage–current-density characteristic γ:

$$V(\zeta)=\frac{\Phi_0\omega_j}{2\pi}\nu\Phi'(\zeta). \tag{38}$$

The boundary conditions for a 4π kink are

$$\phi(+\infty)-\phi(-\infty)=4\pi, \quad \phi^{(1,2)}(\pm\infty)=0, \tag{39}$$

and the dimensionless velocity ν of a vortex is an eigenvalue of the equation

$$\nu^2\phi''(\zeta)-\alpha\nu\phi'(\zeta)+\sin\,\phi(\zeta)-\gamma = \frac{\varepsilon}{\pi}\int_{-\infty}^{\infty}du\left\{\left(K_0(\varepsilon|\zeta-u|)-\beta\nu[2K_0(\varepsilon|\zeta-u|) +\varepsilon|\zeta-u|K_0'(\varepsilon|\zeta-u|)]\frac{d}{du}\right)\phi''(u)\right\}, \tag{40}$$

provided that the boundary conditions (39) are satisfied. Note that if $\beta=0$, Eq. (40) coincides with Eq. (11) of Ref. 1. The case $\varepsilon\to\infty$ corresponds to an ordinary Josephson vortex. If we limit ourselves to junctions with high current densities j_c ($\varepsilon\ll 1$), we are dealing with what is known as small-scale Abrikosov–Josephson vortices.[1,10] Then for the function $K_0(u)$ we can use its asymptotic expression for a small argument, $K_0(u)=\ln(2/u)$.

Using the second condition in (39) and the fact that

$$\int_{-\infty}^{\infty}\ln(|\zeta-u|)\phi^{(2,3)}(u)du=-\fint_{-\infty}^{\infty}\frac{du\,\phi^{(1,2)}(u)}{u-\zeta} \tag{41}$$

(here $\fint$ denotes the Cauchy principal value of the integral), we can transform (40) into

$$\nu^2\phi''(\zeta)-\alpha\nu\phi'(\zeta)+\sin\,\phi(\zeta)-\gamma = \frac{\varepsilon}{\pi}\fint_{-\infty}^{\infty}\frac{du}{u-\zeta}\left(\frac{d\phi}{du}-2\beta\nu\frac{d^2\phi}{du^2}\right). \tag{42}$$

This equation was used to study a traveling 2π kink under conditions of high dissipation ($\alpha\gg 1$) in Ref. 1 for $\beta=0$ and in Ref. 11 for $\beta\neq 0$. Here, bearing in mind the boundary conditions (39), we introduce the ansatz

$$\phi(\zeta)=\theta+4\arctan\frac{\zeta}{s}, \tag{43}$$

where s is the size of the core of the moving vortex, and

$$\gamma=\sin\,\theta\leqslant 1. \tag{44}$$

We also note that at $s=\varepsilon$ and $\nu^2=\varepsilon^2$, Eq. (43) yields the solution for a zero-current nondissipative 4π kink ($\alpha=\beta=\gamma=0$).[12] To estimate the dimensionless velocity ν of a vortex for a junction with current ($\gamma\neq 0$) and with dissipation ($\alpha\neq 0$ and $\beta\neq 0$), it is sufficient to derive two equations from (42) for the unknown variable parameters s and ν. Multiplying both sides of (42) by $\phi'(\zeta)$ and integrating the products with respect to ζ, we obtain a relation that expresses the balance between friction and the Lorentz force:

$$\alpha\nu\int_{-\infty}^{\infty}\phi'^2(\zeta)d\zeta+64\varepsilon s\beta\nu\int_{-\infty}^{\infty}\frac{u^2du}{(u^2+s^2)^3}=-4\pi\gamma. \tag{45}$$

To obtain the second integral relation, we multiply (42) by $\phi''(\zeta)$ and again integrate with respect to ζ. This yields

$$\nu^2\int_{-\infty}^{\infty}\phi''^2d\zeta+\int_{-\infty}^{\infty}\phi''\,\sin\,\phi d\zeta = 16\varepsilon s\int_{-\infty}^{\infty}\frac{s^2-u^2}{(s^2+u^2)^3}du. \tag{46}$$

Employing the identity

$$\sin[\theta+4\arctan\,a]=\frac{a^4-6a^2+1}{(1+a^2)^2}\sin\,\theta+\frac{4a(1-a^2)}{(1+a^2)^2}\cos\,\theta \tag{47}$$

and calculating the integrals in (45) and (46), we obtain the required two equations for ν and s:

$$2\nu\left(\alpha+\frac{\varepsilon}{s}\beta\right)=-s\gamma, \quad \nu^2=\varepsilon s. \tag{48}$$

Eliminating s, we find the dimensionless vortex velocity

$$\nu=A+B-\frac{2\alpha\varepsilon}{3\gamma}, \tag{49}$$

where

$$A=\left(-\frac{q}{2}+\sqrt{Q}\right)^{1/3}, \quad B=\left(-\frac{q}{2}-\sqrt{Q}\right)^{1/3},$$

$$\frac{q}{2}=\left[\left(\frac{2\alpha}{3\gamma}\right)^3+\frac{\beta}{\gamma}\right]\varepsilon^3, \quad Q=\left(1+\frac{16\alpha^3}{27\gamma^2\beta}\right)\frac{\beta^2}{\gamma^2}\varepsilon^6.$$

5. CONCLUSION

The basic equations of nonlocal electrodynamics have been derived for a single Josephson junction of arbitrary thickness positioned between two similar semi-infinite superconductors with finite normal dc conductivity. We have studied the dispersion relation for this junction numerically in the linear approximation. We have also analyzed the conditions under which the various theories of nonlocal Josephson electrodynamics can be used. The theory of a thin junction that allows for finite surface resistance of the superconducting electrodes has been developed. Finally, we have estimated the velocity and size of the core of a 4π kink in such a structure.

We would like to express our gratitude to Yu. M. Aliev for discussions of the results, and to the Bulgarian National Science Foundation.

APPENDIX: ESTIMATES OF THE DEPENDENCE OF JOSEPHSON CRITICAL CURRENT DENSITY ON NORMAL-LAYER THICKNESS

We wish to estimate the maximum normal-layer thickness (or the dimensionless parameter $\delta_m=d_m/\lambda$) below

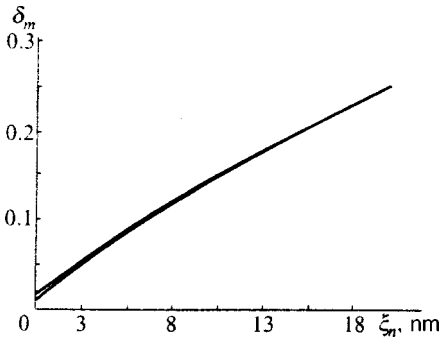

FIG. 3. The parameter δ_m as a function of the coherence length of the normal layer ξ_n.

which nonlocality effects begin to show up. To be more specific, we introduce the definition (see Eqs. ((1) and (25))

$$\varepsilon(\delta_m)=1, \tag{A1}$$

assuming that ordinary local Josephson electrodynamics operates when $\varepsilon \gg 1$, while for junctions with high critical current densities we have $\varepsilon<1$. A simple estimate for the current density in the tunneling layer $|x|<d$,

$$j=-\frac{ie\hbar}{m_s}[\Psi^*(x)\Psi'(x)-\Psi(x)\Psi'^*(x)], \tag{A2}$$

can be made by integrating the Schrödinger equation (see Eq. (A7) with $b_n=0$)

$$\xi_n^2\Psi''(x)=\Psi(x) \tag{A3a}$$

with boundary conditions

$$\Psi(\pm d)=\sqrt{n_s}\,\exp(i\theta_{1,2}). \tag{A3b}$$

The result of these simple calculations is well known (Eq. (8.13) in Ref. 13), and has the form ($\lambda=\sqrt{m_s/4e^2\mu_0 n_s}$)

$$j=j_c\,\sin(\theta_2-\theta_1), \tag{A4a}$$

$$j_c(d)=\frac{\Phi_0}{2\pi\mu_0\lambda^2\xi_n\,\sinh(2d/\xi_n)}. \tag{A4b}$$

Various experimental estimates[14,15] for the intergrain boundary in a $YBa_2Cu_3O_7$ film at $T=4.2$ K yield a coherence length $\xi_n=1-10$ nm for a normal layer, while for the London magnetic-field penetration depth we have $\lambda=140$ nm. If these data are taken into account, Eq. (A1) becomes

$$\xi_n\,\sinh\frac{2d_m}{\xi_n}=2(\lambda+d_m). \tag{A5}$$

In Fig. 3 the parameter $\delta_m=d_m/\lambda$ is plotted as function of the coherence length ξ_n in the normal layer. This numerical example and Fig. 3 show that at $\varepsilon\sim 1$ we must allow for the finite thickness of the normal layer in the junction, with $0\neq\delta\leqslant\delta_m$ and $\delta_m\sim 0.1$.

We believe that for junctions with extremely high critical current densities ($\varepsilon\ll 1$) it is advisable to assess the current–phase relation via modified the Ginzburg–Landau equations ($a_n>0$ and $b_n\geqslant 0$)

$$\frac{\hbar^2}{2m_s}\frac{d^2\Psi}{dx^2}+|a_s|\Psi-b_s|\Psi|^2\Psi=0, \quad |x|>d, \tag{A6}$$

$$\frac{\hbar^2}{2m_n}\frac{d^2\Psi}{dx^2}-a_n\Psi-b_n|\Psi|^2\Psi=0, \quad |x|<d, \tag{A7}$$

and using the following estimates for the YBCO crystal[16]:

$$m_s=8m_e, \quad |a_s(T)|=1.2\times 10^{-21}\left(\frac{T_c-T}{T_c}\right)\text{J},$$

$$b_s=4\times 10^{-49}\ \text{Jm}^3. \tag{A8}$$

Indeed, as shown by Sols and Ferrer[17] for the special case $m_s=m_n$ and $b_s=b_n$, the divergence of the critical current density (A4b) as $d\to+0$ is removed, and in the limit of very thin junctions the Josephson current density becomes the depairing current density:

$$\lim_{d\to 0} j_c(d)=\frac{\Phi_0}{3\sqrt{3}\pi\mu_0\lambda^2\xi_s}, \tag{A9}$$

where $\xi_s=\hbar/\sqrt{2m_s|a_s|}$.

*)E-mail: zgenchev@phys.acad.bg

[1] A. Gurevich, Phys. Rev. B **48**, 12 857 (1993).
[2] A. Barone and G. Paterno, *Physics and Applications of the Josephson Effect*, Wiley, New York (1982).
[3] Yu. M. Aliev, K. N. Ovchinnikov, V. P. Silin, and S. A. Uryupin, Zh. Éksp. Teor. Fiz. **107**, 972 (1995) [JETP **80**, 551 (1995)].
[4] A. F. Volkov, Physica C **192**, 306 (1991).
[5] R. G. Mints and I. B. Shapiro, Phys. Rev. B **49**, 6188 (1994).
[6] G. L. Alfimov and A. F. Popkov, Phys. Rev. B **52**, 4503 (1995).
[7] A. Gurevich and L. D. Cooley, Phys. Rev. B **50**, 13 563 (1994).
[8] Yu. M. Aliev and V. P. Silin, Zh. Éksp. Teor. Fiz. **104**, 2526 (1993) [JETP **77**, 142 (1993)].
[9] V. P. Silin, JETP Lett. **58**, 701 (1993).
[10] V. P. Silin, Zh. Éksp. Teor. Fiz. **110**, 741 (1996) [JETP **83**, 401 (1996)].
[11] Z. D. Genchev, Supercond. Sci. Technol. **10**, 543 (1997).
[12] Yu. M. Aliev and V. P. Silin, Phys. Lett. A **177**, 259 (1993).
[13] T. P. Orlando and K. A. Delin, *Foundations of Applied Superconductivity*, Addison-Wesley, New York (1991).
[14] E. Polturak, G. Koren, D. Cohen, E. Aharoni, and G. Deutscher, Phys. Rev. Lett. , 3038 (1991).
[15] R. Gross, *Interfaces in Superconducting Systems*, S. L. Shinde and D. Rudman (eds.), Springer-Verlag, New York (1996), Chap. 6.
[16] L. N. Bulaevskiĭ, V. L. Ginzburg, and A. A. Sobyanin, Zh. Éksp. Teor. Fiz. **94**, No. 7, 355 (1988) [Sov. Phys. JETP **67**, 1499 (1988)].
[17] F. Sols and J. Ferrer, Phys. Rev. B **49**, 15 913 (1994).

Translated by Eugene Yankovsky

Kinetics of magnetic dipoles and unified theory of NMR spectra in condensed matter

B. N. Provotorov,*) T. P. Kulagina,†) and G. E. Karnaukh‡)

Institute of Chemical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 5 June 1997)
Zh. Éksp. Teor. Fiz. **113**, 967–980 (March 1998)

A new approach to the analysis of magnetic dipole motion in external magnetic field and fields generated by neighboring magnetic dipoles is suggested, and original general kinetic equations for the dipole density are derived. Special cases of these general equations are the Bloch, Redfield, and Provotorov equations, which are widely used in NMR theory. A comparison between NMR spectra calculated with the new theory and published experimental data also shows good agreement in regions to which the equations listed above do not apply. © *1998 American Institute of Physics.* [S1063-7761(98)01403-6]

1. INTRODUCTION

NMR techniques have been extensively used in studies of the molecular motion and structure of various materials.[1–3] It is well known from NMR theory that in order to extract information from NMR spectra about the structure and motion of molecules in a material, one must first derive interpretable kinetic equations for directly observable macroscopic densities of magnetic dipoles from the equation for the density operator $\rho(t)$, which describes the microscopic motion of all ($N\sim10^{23}$) molecules in a sample and is unsolvable in real cases. Such kinetic equations were obtained about forty years ago for the three cases[1,4,5] of most practical interest, based on approximate solutions of the equation for the density operator for a system of $N\sim10^{23}$ microscopic dipoles μ_i coupled through dipole–dipole interactions. An important point is that the derivation of these equations[1,4,5] began by defining, on the basis of simple physical considerations, the main component of the density operator $\rho_0(t)$, which determines the dipole density $M(t)$, and the operator $\rho_1(t)$

$$\rho_1(t)=\rho(t)-\rho_0(t)\ll\rho_0(t) \tag{1}$$

characterizing the time derivatives of the dipole density and inverse spin temperature $\beta_s(t)$. This approach was used in the 1950s[1] to obtain Bloch's kinetic equations for the macroscopic dipole density $M(t)$, which permitted adequate description of NMR in liquids and gases when

$$\omega_d\tau_c\ll1, \tag{2}$$

where $\omega_d=d^2n/\hbar$, d is the dipole moment, n is the dipole concentration, and τ_c is the time in which the local dipole field acting on a specific dipole changes due to thermal motion.

The main component of the density operator $\rho_0(t)$ in liquids for a resonant magnetic field ω_1 of arbitrary amplitude was expressed, using the smallness $\alpha(t)=\mu H_0/kT\ll1$, in the form

$$\rho_0(t)\sim1+\alpha(t)\hat{S}, \tag{3}$$

where $\hat{S}$ is the operator of the total dimensionless nuclear spin of the paramagnetic material, H_0 is the magnitude of the dc magnetic field, and T is the temperature of the substance.

In a solid material and weak resonant field,

$$\omega_{\rm loc}\tau_c\gg1, \tag{4}$$

$$\omega_1\sim\omega_{\rm loc}, \tag{5}$$

where $\omega_{\rm loc}$ is the local field frequency, the main component of the density operator was expressed, using simple physical considerations, in the form[4]

$$\rho_0(t)\sim1+\beta_s(t)(\Delta\hat{S}_z+\omega_1\hat{S}_x+\hat{H}_d^z), \tag{6}$$

where Δ is the frequency offset of the resonant magnetic field, $\beta_s(t)$ is the inverse spin temperature in the material in frequency units, and $\hat{H}_d^z$ is the secular part of the dipole–dipole interaction.

Finally, in the case of a solid material and resonant field $\omega_1\ll\omega_{\rm loc}$, the main component of the density operator was expressed as[5]

$$\rho_0(t)\sim1+\beta_s(t)\Delta\hat{S}_z+\beta_d(t)\hat{H}_d^z, \tag{7}$$

where $\beta_d(t)$ is the inverse temperature of the dipole–dipole reservoir (with dimensions of time).

Further,[1,4,5] $\rho_0(t)$ and the equation for the density operator

$$\frac{d\rho}{dt}=-i[\Delta\hat{S}_z+\omega_1\hat{S}_x+\hat{H}_d^z,\rho] \tag{8}$$

were used in calculating $\rho_1(t)$, which determines the rate of change of dipole densities, reciprocal spin temperatures, and the actual form of the kinetic equations. The resulting equations[1,4,5] for $M(t)$ and $\beta(t)$ allowed the researchers to explain many effects observed in NMR experiments. These successes were the first to demonstrate that it is possible to obtain fairly simple kinetic equations for the dipole density observed in NMR, although the motion of $N\sim10^{23}$ dipoles described by Eq. (8) is extremely complicated. The constraints (2), (4), and (5) imposed on τ_c and ω_1, however,

considerably reduced the applicability of NMR to studies of molecular motion and the structure of materials.

The aim of the reported work was to eliminate these constraints and obtain general kinetic equations for magnetic dipole densities and spin temperatures that could be applied to analysis of NMR spectra at all ω_1 and τ_c. The most difficult problem in deriving these kinetic equations is taking into account the contribution of $\hat{H}_d^z$ for the rate of change of dipole density, which will be discussed in the next section.

2. CONTRIBUTION OF DIPOLE–DIPOLE INTERACTIONS TO KINETIC EQUATIONS FOR DIPOLE DENSITY

An important point in determining the contribution of dipole–dipole interactions to the kinetic equations for the dipole density is that the motion of all magnetic and optical dipoles is controlled by an elementary dynamical process in the form of dipole precession in external fields and local fields generated by neighboring dipoles. This fact was derived from the analysis of the great amount of data on decay of polarizations due to the dipole dephasing obtained in experiments on magnetic resonance[1,3] and resonant optical effects.[6] This process was studied in NMR experiments as early as the 1950s,[1–3] in which the decay of free nuclear polarization was measured. Interestingly enough, following the advent of the laser in the 1960s, it became clear that similar processes govern the motion of optical dipoles. Later, in the 1970s, the key role of the classical elementary dynamical process in dipole kinetics was confirmed by molecular-dynamics calculations of free polarization decay in CaF_2 crystals detected in NMR experiments.[7] These calculations clearly demonstrated that the elementary dynamical process provides an excellent description not only of the dipole dephasing, but also of concomitant oscillations detected in experiments on the decay of free polarization in CaF_2 crystals.

After molecular-dynamics calculations,[7] the only problem in obtaining general kinetic equations was to develop an analytic approach to calculation of the contribution of the dipole–dipole interaction to the time derivative of the dipole density and derivation of the kinetic equations. Such an approach was suggested earlier in Ref. 6, where the contribution of the dynamical process to the rate of change of optical dipole density was determined for the first time not on the basis of approximate calculations, but by using a simple physical reasoning. An important point in determination of this contribution is selection of variables to characterize the dipole density. The most convenient variables for calculations based on this approach are polarizations of layers previously suggested in Ref. 8 to describe dipole precession, or partial dipole densities $\sigma^{\beta}(h,t)$, where $\beta=x,y,z$, which are in the longitudinal local dipole field h at time t.

The major problem in deriving such equations is calculation of the contribution of dipole precession in local magnetic fields to the rate of change of the dipole density determined by the Hamiltonian of the dipole–dipole interaction

$$\hat{H}_d^z=\sum_{i>k} b_{ik}(3\hat{\sigma}_i^z\hat{\sigma}_k^z-(\boldsymbol{\sigma}_i\boldsymbol{\sigma}_k))=3\hat{H}_{zz}+\hat{H}_{is}, \tag{9}$$

where $\hat{\sigma}_i^{\beta}$ are spin operators proportional to dipole moments d_i, and b_{ik} are predetermined coefficients of the dipole–dipole interaction.[1]

The contribution of $\hat{H}_{zz}$ was calculated exactly[8]:

$$\frac{\partial F(h,t)}{\partial t}=ihF(h,t),$$

$$F(h,t)=\sigma^x(h,t)+i\sigma^y(h,t), \tag{10}$$

where

$$h=\frac{\mathrm{Tr}\left[\sum_k b_{ik}\sigma_k^z\hat{P}(h)\right]}{\mathrm{Tr}\,\hat{P}(h)},$$

$\hat{P}(h)$ is the projector on all dipole configurations that generate field h at the site of the ith dipole.

The most difficult part of the calculation is taking into account the contribution of $\hat{H}_{is}$ to the rate of change of the partial dipole density, which was first done[8] by intuitively truncating formally accurate but extremely cumbersome expressions derived using the memory function technique. Only recently[6] was it proved that the contribution of $\hat{H}_{is}$ to the kinetic equations can be easily obtained using five obvious physical properties typical of this interaction: 1) it should be proportional to the local field h and 2) components $\sigma^{x,y}$ of the layer polarization; 3) the isotropic component of this interaction, $\hat{H}_{zz}$, contributes to the rate of change of layer polarization an exactly calculable contribution proportional to $h\sigma^{x,y}$; 4) since Hamiltonian $\hat{H}_{is}$ describes exchange of polarizations between spins, its contribution should vanish after multiplication by $g(h)$, which describes the distribution of local longitudinal fields in dense spin systems, and after integration over all layers, and 5) in the case of equal polarizations in all layers.

Given these simple physical considerations, one can easily determine the form of the contribution of $\hat{H}_d^z$ to the rate of change of layer polarizations measured in free polarization decay experiments, and a simple kinetic equation can be obtained:

$$\frac{dF(h,t)}{dt}=\frac{3}{2}iF(h,t)-i\left(\frac{3}{2}-\alpha\right)\times[hF(h,t)-hF_0(t)-F_1(t)], \tag{11}$$

where

$$F(h,t)=\sigma^x(h,t)+i\sigma^y(h,t), \tag{12}$$

$$F_0(t)=\int dh\; g(h)F(h,t), \tag{13}$$

$$F_1(t)=\int dh\; g(h)hF(h,t). \tag{14}$$

In the next section, we will add to Eq. (11) contributions of spectral diffusion, spin–lattice relaxation, and interaction with ac resonant field of an arbitrary amplitude to the rate of change of the layer polarization; thus we will obtain kinetic equations for description of experimental data.

3. GENERAL KINETIC EQUATIONS FOR LAYER POLARIZATION OF SOLIDS IN NMR EXPERIMENTS

In deriving kinetic equations, it is convenient to introduce the following notation[8]:

$$\sigma_0^\beta = \int \sigma(h,t) g(h)\, dh, \tag{15}$$

$$\sigma_1^\beta = \int \sigma(h,t) h g(h)\, dh. \tag{16}$$

In deriving general kinetic equations with due account of the contributions of dipole–dipole interaction to the rate of change of layer polarizations, we add to equations like (11) the contributions of spectral diffusion and spin–lattice relaxation to the rate of change of dipole density in the form of simple expressions:

$$(\sigma_0^\alpha - \sigma^\alpha)/\tau_\perp, \quad \alpha = x,y, \tag{17}$$

$$\sigma^\alpha / T_\perp, \quad \alpha = x,y. \tag{18}$$

In an ac resonant field of arbitrary amplitude ω_1 aligned with the x-axis, the equation should include terms like

$$\omega_1 \sigma^\alpha, \quad \alpha = y,z. \tag{19}$$

In order to take into account the spin–lattice relaxation of both the Zeeman (with relaxation time $T_{\|Z}$) and dipole–dipole (with time $T_{\|d}$) reservoirs, we add to the equation terms like

$$(\sigma_{eq}^z - \sigma^z)/T_{\|Z},$$

$$h\sigma_1^z / \langle h^2 \rangle T_{\|d}, \tag{20}$$

where σ_{eq}^z is the equilibrium dipole density, and the term

$$(\sigma_0^z + h\sigma_1^z/\langle h^2 \rangle - \sigma^z)/\tau_\| \tag{21}$$

describes the equilibration of the spin system of the solid material.

Thus, by taking into account the precession of dipoles in local magnetic fields (11) and all the processes described above by Eqs. (17)–(21), we obtain the following kinetic equations for the rate of change of dipole density:

$$\frac{d\sigma^x}{dt} = -\left(\Delta + \frac{3h}{2}\right)\sigma^y - \left(\frac{3}{2} - \alpha\right)(h\sigma_0^y + \sigma_1^y - h\sigma^y) + \frac{\sigma_0^x - \sigma^x}{\tau_\perp} - \frac{\sigma^x}{T_\perp}, \tag{22}$$

$$\frac{d\sigma^y}{dt} = \left(\Delta + \frac{3h}{2}\right)\sigma^x + \left(\frac{3}{2} - \alpha\right)(h\sigma_0^x + \sigma_1^x - h\sigma^x) - \omega_1 \sigma^z + \frac{\sigma_0^y - \sigma^y}{\tau_\perp} - \frac{\sigma^y}{T_\perp}, \tag{23}$$

$$\frac{d\sigma^z}{dt} = \omega_1 \sigma^y + \frac{1}{\tau_\|}\left(\sigma_0^z + \frac{h\sigma_1^z}{\langle h^2 \rangle} - \sigma^z\right) + \frac{\sigma_{eq}^z - \sigma^z}{T_{\|Z}} - \frac{h\sigma_1^z}{\langle h^2 \rangle T_{\|d}}, \tag{24}$$

where Δ is the offset from resonance, ω_1 is the resonant field amplitude, $1/\tau_\perp$ is the rate of change of longitudinal local magnetic fields due to spin exchange and thermal motion of atoms, $1/\tau_\|$ is the equilibration rate in the spin system of the solid, $1/T_{\perp(\|)}$ is the rate of transverse (longitudinal) relaxation due to thermal motion, resulting in absorption of photon of energy $\hbar\omega_0$ corresponding to the Larmor frequency ω_0, and $T_{\|Z}$ and $T_{\|d}$ are the times of longitudinal spin–lattice relaxation in the Zeeman and dipole–dipole reservoirs. The parameter $3/2 - \alpha$ is the measure of the nonaveraged part of the isotropic dipole–dipole interaction $\hat{H}_{is}$, which depends on the material structure and its alignment with respect to the dc magnetic field.

In what follows, we will show that the kinetic equations derived by this technique are equivalent to the Redfield, Bloch, and Provotorov equations in the usual domains of applicability of these equations.

4. DERIVATION OF THE REDFIELD EQUATION FROM THE GENERAL KINETIC EQUATION

An important point in the derivation of kinetic equations in a rotating coordinate system is the equation of energy conservation in the spin system:

$$\Delta\sigma_0^z + \omega_1 \sigma_0^x + \frac{3}{2}\sigma_1^z = \text{const}. \tag{25}$$

Let us derive Eq. (25) from the general equations given in the previous section. To this end, we average Eqs. (22) and (24) over the distribution function $g(h)$ and then Eq. (24) over the function $hg(h)$:

$$\frac{d\sigma_0^x}{dt} = -\Delta\sigma_0^y - \frac{3}{2}\sigma_1^y - \frac{\sigma_0^x}{T_\perp}, \tag{26}$$

$$\frac{d\sigma_0^z}{dt} = \omega_1 \sigma_0^y + \frac{\sigma_{eq}^z - \sigma_0^z}{T_{\|Z}}, \tag{27}$$

$$\frac{d\sigma_1^z}{dt} = \omega_1 \sigma_1^y - \sigma_1^z\left(\frac{1}{T_{\|Z}} + \frac{1}{T_{\|d}}\right). \tag{28}$$

Multiplying Eq. (26) by ω_1, Eq. (27) by Δ, and Eq. (28) by 3/2, and summing all these equations, we obtain a fundamental property of the kinetic equations, describing energy exchange between the spin system and lattice:

$$\frac{d(\omega_1\sigma_0^x + \Delta\sigma_0^z + (3/2)\sigma_1^z)}{dt} = -\frac{\omega_1\sigma_0^x}{T_\perp} + \Delta\frac{\sigma_{eq}^z - \sigma_0^z}{T_{\|Z}} - \frac{3}{2}\sigma_1^z\left(\frac{1}{T_{\|Z}} + \frac{1}{T_{\|d}}\right). \tag{29}$$

In the range of times $t \sim \tau_{\|,\perp} \ll T_{\|Z,\|d,\perp}$ the spin system can be considered to be isolated from the lattice. In this case, the spin-system energy in Eq. (29) is an integral of the motion defined by Eq. (25).

We now show from Eqs. (22)–(24) the Redfield equation[4] using energy conservation (25). To this end, we derive from Eqs. (22)–(24) the quasi-equilibrium conditions discussed in Ref. 4. Note that in this case an averaged orientation for all dipoles is established; as a result,

$$\sigma^{\beta}=\sigma_0^{\beta}+\frac{h}{\langle h^2\rangle}\sigma_1^{\beta},\quad \beta=x,y. \tag{30}$$

By substituting expression (30) into Eqs. (22)–(24) and averaging the latter over h and $hg(h)$, we obtain the following system of kinetic equations, in which energy exchange between the spins and lattice can be neglected during spin–spin interaction:

$$\begin{aligned}
&\frac{d\sigma_0^x}{dt}=-\Delta\sigma_0^y-\frac{3}{2}\sigma_1^y,\quad \frac{d\sigma_0^z}{dt}=\omega_1\sigma_0^y,\\
&\frac{d\sigma_0^y}{dt}=\Delta\sigma_0^x+\frac{3}{2}\sigma_1^x-\omega_1\sigma_0^z,\\
&\frac{d\sigma_1^x}{dt}=-\Delta\sigma_1^y-\frac{3}{2}\langle h^2\rangle\sigma_0^y-\frac{\sigma_1^x}{\tau_\perp},\\
&\frac{d\sigma_1^y}{dt}=\Delta\sigma_1^x+\frac{3}{2}\langle h^2\rangle\sigma_0^x-\omega_1\sigma_1^z-\frac{\sigma_1^y}{\tau_\perp},\quad \frac{d\sigma_1^z}{dt}=\omega_1\sigma_1^y.
\end{aligned} \tag{31}$$

In quasi-equilibrium, the kinetic equation system (31) has the sole solution

$$\begin{aligned}
&\sigma_0^y=\sigma_1^y=\sigma_1^x=0,\quad \sigma_0^x=\beta\omega_1,\\
&\sigma_0^z=\beta\Delta,\quad \sigma_1^z=\frac{3}{2}\beta\langle h^2\rangle,
\end{aligned} \tag{32}$$

where β is the reciprocal spin temperature.[4]

From Eqs. (29) and (32) we can derive a kinetic equation for the reciprocal spin temperature, which is Redfield's classical equation[4]:

$$\frac{d\beta_s}{dt}=\frac{\beta_0-\beta_s}{T_{1\rho}}, \tag{33}$$

where

$$\begin{aligned}
&\beta_0=\beta_i T_{1\rho}/T_{\|Z},\quad \beta_i=\beta_{eq}\frac{\Delta^2}{\omega_1^2+\Delta^2+\omega_{\rm loc}^2},\\
&\frac{1}{T_{1\rho}}=\frac{1}{\omega_1^2+\Delta^2+\omega_{\rm loc}^2}\left(\frac{\omega_1^2}{T_\perp}+\frac{\Delta^2}{T_{\|Z}}+\frac{\omega_{\rm loc}^2}{T_c}\right),\\
&\frac{1}{T_c}=\frac{1}{T_{\|Z}}+\frac{1}{T_{\|d}},\quad \omega_{\rm loc}^2=\frac{9}{4}\langle h^2\rangle.
\end{aligned}$$

We now proceed to the derivation of the Bloch equation from Eqs. (22)–(24).

5. DERIVATION OF THE BLOCH EQUATION FROM THE GENERAL KINETIC EQUATION

To derive the Bloch equation, we should note that it applies to the case of fast thermal motion of molecules in the material studied, i.e., $\omega_{\rm loc}\tau_\perp\ll 1$. Under these conditions, the transverse polarization of layers $F(h,t)$ defined by Eq. (10) can be approximated by an expression similar to Eq. (30):

$$F(h,t)=F_0(t)+\frac{h}{\langle h^2\rangle}F_1(t). \tag{34}$$

In order to derive the Bloch equation from Eqs. (22)–(24), it suffices to consider just the contributions of the dipole–dipole interaction to the rate of change of layer polarizations:

$$\left(\frac{dF}{dt}\right)_{dd}=i\alpha hF+(1.5-\alpha)(hF_0+F_1)+\frac{F_0-F}{\tau_\perp}. \tag{35}$$

After averaging Eqs. (34) and (35) over $g(h)$ and $hg(h)$, we obtain

$$\left(\frac{dF_0}{dt}\right)_{dd}=\frac{3}{2}iF_1,\quad \left(\frac{dF_1}{dt}\right)_{dd}=\frac{3}{2}i\langle h^2\rangle F_0-\frac{1}{\tau_\perp}F_1, \tag{36}$$

where the subscript dd indicates that Eq. (36) applies to the part of $F(h,t)$ that depends only on the dipole–dipole interaction of spins and spectral diffusion. Analysis of the exact solution of Eq. (36) allows one to separate out the parameter $9\langle h^2\rangle\tau_\perp^2$, which controls the spin system kinetics. It has been shown that for

$$9\langle h^2\rangle\tau_\perp^2\ll 1 \tag{37}$$

the kinetic equations for $F_0(t)$ and $F_1(t)$ are decoupled:

$$\left(\frac{dF_0}{dt}\right)_{dd}=-\frac{9}{4}\langle h^2\rangle\tau_\perp F_0(t),\quad \left(\frac{dF_1}{dt}\right)_{dd}=-\frac{1}{\tau_\perp}F_1(t). \tag{38}$$

By substituting expressions (35) and (38) into the original system (22)–(24), we obtain the following set of kinetic equations:

$$\begin{aligned}
&\frac{d\sigma_0^x}{dt}=-\Delta\sigma_0^y-\left(\frac{1}{T_\perp}+\frac{9}{4}\langle h^2\rangle\tau_\perp\right)\sigma_0^x,\\
&\frac{d\sigma_0^y}{dt}=\Delta\sigma_0^x-\left(\frac{1}{T_\perp}+\frac{9}{4}\langle h^2\rangle\tau_\perp\right)\sigma_0^y-\omega_1\sigma_0^z,\\
&\frac{d\sigma_0}{dt}=\omega_1\sigma_0^y+\frac{\sigma_{eq}^z-\sigma_0^z}{T_{\|Z}}.
\end{aligned} \tag{39}$$

Thus, when Eq. (37) holds, we have obtained the set of equations (39), which is equivalent to the Bloch equations[1] if we write $1/T_\perp'=1/T_\perp+9\tau_\perp\langle h^2\rangle/4$.

6. DERIVATION OF THE BASIC EQUATION FROM THE GENERAL KINETIC EQUATIONS

To derive the basic equation, note that the local field at the site of a specific spin is largely determined by the nearest neighbors, i.e., we can consider a finite number (10^3–10^4) of local field h values. Then we can rewrite the general equations (22)–(24) in matrix form by introducing vector and matrix operators. To this end, we order the spatial basis according to local field values h. Then each function $a(h)$ is replaced by the vector **a** whose components are $a(h_j)$. For convenience, we introduce the following notation: **I** is a vector all of whose elements are 1, $\hat{K}$ is a matrix all of whose elements are 1, and $\hat{D}(\mathbf{a})$ is the diagonal matrix whose diagonal elements are the components of vector **a**. Let us express the kinetic matrix operators in the explicit form:

$$\hat{U}=i\alpha\hat{D}(\mathbf{h})+i\left(\frac{3}{2}-\alpha\right)[\hat{D}(\mathbf{h})\hat{K}\hat{D}(\mathbf{g})+\hat{K}\hat{D}(\mathbf{hg})]$$

$$+\frac{1}{\tau_\perp}\hat{K}\hat{D}(\mathbf{g})-\left(\frac{1}{\tau_\perp}+\frac{1}{T_\perp}\right)\hat{E}, \tag{40}$$

$$\hat{A}=\frac{1}{\tau_\parallel}\left[\hat{K}\hat{D}(\mathbf{g})+\frac{1}{\langle h^2\rangle}\,\hat{D}(\mathbf{h})\hat{K}\hat{D}(\mathbf{hg})-\hat{E}\right], \tag{41}$$

$$\hat{B}=\frac{1}{T_{\parallel d}\langle h^2\rangle}\,\hat{D}(\mathbf{h})\hat{K}\hat{D}(\mathbf{h}\cdot\mathbf{g}), \tag{42}$$

after which Eqs. (22)–(24) take the compact form

$$\frac{d\mathbf{F}}{dt}=(\hat{U}+i\Delta\hat{E})\mathbf{F}-i\omega_1\boldsymbol{\sigma}^z,$$

$$\frac{d\boldsymbol{\sigma}^z}{dt}=\omega_1\ \mathrm{Im}(\mathbf{F})+\left(\hat{A}-\frac{\hat{E}}{T_{\parallel Z}}-\hat{B}\right)\boldsymbol{\sigma}^z+\sigma^z_{eq}\frac{\mathbf{I}}{T_{\parallel Z}}. \tag{43}$$

The first line of Eq. (43) defines the vector **F**, and substituting it into the second line, we obtain a closed equation for the longitudinal polarization vector $\boldsymbol{\sigma}^z$:

$$\begin{aligned}\frac{d\boldsymbol{\sigma}^z}{dt}=&-\omega_1^2\ \mathrm{Re}\left\{\int\ \exp[(\hat{U}+i\Delta\hat{E})(t-t')]\boldsymbol{\sigma}^z dt'\right\}\\&+\omega_1\ \mathrm{Im}\{\exp[(\hat{U}+i\Delta\hat{E})t]\}\mathbf{F}(h,0)\\&+\left(\hat{A}-\hat{E}/T_{\parallel Z}-\hat{B}\right)\boldsymbol{\sigma}^z+\sigma^z_{eq}\frac{\mathbf{I}}{T_{\parallel Z}},\end{aligned} \tag{44}$$

where $\mathbf{F}(h,0)$ is the vector of the initial transverse polarization of the layer. Note that the matrix $\hat{U}$ describes free polarization decay, so it determines $G(t)$, i.e., the signal due to this decay:

$$\begin{aligned}G(t)=(\exp[\hat{U}t]\mathbf{I},\ \mathbf{g})&=\sum_n\ (\hat{U}^n\mathbf{I},\ \mathbf{g})t^n/n!\\&=\sum_n\ M_n t^n/n!\end{aligned} \tag{45}$$

$$d^nG/dt^n=(\exp[\hat{U}t]\hat{U}^n\mathbf{I},\ \mathbf{g}). \tag{46}$$

This means that the initial transverse polarization can be neglected at times larger than the free polarization decay time, i.e., in the region where the theory of Ref. 5 applies.

Note that under the conditions of the theory of Ref. 5, when $\omega_1\ll\omega_{\mathrm{loc}}$, $\omega_1^2\tau_\parallel\tau_\perp\ll 1$ at times of order $\tau_\parallel$ ($\tau_\parallel<10^{-5}$ s; Ref. 9), which is much shorter than the free polarization decay time, quasi-equilibrium is established with the longitudinal magnetization vector:

$$\boldsymbol{\sigma}^z(t)=\Delta\beta_Z\mathbf{I}+\frac{9}{4}\beta_d\mathbf{h}, \tag{47}$$

where the typical time of change in β_Z and β_d is much longer than the free polarization decay time.

Substituting Eq. (47) into Eq. (44) and taking scalar products of the resulting equation and **g** and **hg**, we obtain a system of linear kinetic equations for the inverse temperatures:

$$\frac{d\beta_Z}{dt}=a_{11}\beta_Z+a_{12}\beta_d+\frac{\beta_{eq}-\beta_Z}{T_{\parallel Z}},$$

$$\frac{d\beta_d}{dt}=a_{21}\beta_Z+a_{22}\beta_d-\left(\frac{1}{T_{\parallel Z}}+\frac{1}{T_{\parallel d}}\right)\beta_d, \tag{48}$$

where $\sigma^z_{eq}=\beta_{eq}\Delta$ and the coefficients a_{ij} are linear combinations of the integrals

$$Y_n=\int_0^\infty\ \exp(i\Delta t)\frac{d^nG(t)}{dt^n}dt. \tag{49}$$

In deriving Eq. (49) we have used the relationships

$$\mathbf{h}=-\frac{2}{3}i\left(\hat{U}+\frac{\hat{E}}{T_\perp}\right)\mathbf{I},\quad \mathbf{hg}=\frac{2}{3}i\left(\hat{U}^*+\frac{\hat{E}}{T_\perp}\right)\mathbf{g}, \tag{50}$$

where the matrix U^* is the Hermitian conjugate of $\hat{U}$, and the fact that the inverse temperatures β_z and β_d are essentially constant over the free polarization decay time.

To calculate integrals in Eq. (49), we have derived the generating function

$$J(x)=\sum_{n=0}^{n-1}\ Y_n x^n=\frac{Y_0-x\Sigma_{n=0}^{n-1}M_n x^n}{1+i\Delta x}. \tag{51}$$

From Eq. (51) we obtain

$$Y_n=(-i\Delta)^n\pi[f(\Delta)+if'(\Delta)]-\sum_{s=0}^{n-1}\ M_s(-i\Delta)^{n-s-1}, \tag{52}$$

where $f(\Delta)$ describes the absorption line profile and $f'(\Delta)$ is the dispersion signal. Using Eq. (52), we obtain the following coefficients for equation system (48):

$$\begin{aligned}a_{11}&=-\pi\omega_1^2 f(\Delta),\\a_{21}&=\frac{4}{9}\frac{\pi\omega_1^2}{\langle h^2\rangle}\left(\Delta^2 f(\Delta)-\frac{\Delta f'(\Delta)}{T_\perp}\right),\\a_{12}&=\pi\omega_1^2\left(f(\Delta)-\frac{f'(\Delta)}{\Delta T_\perp}\right),\\a_{22}&=\frac{4}{9}\frac{\pi\omega_1^2}{\langle h^2\rangle}\left[-\Delta^2 f(\Delta)-\frac{1}{\pi T_\perp}+\frac{2\Delta f'(\Delta)}{T_\perp}+\frac{f(\Delta)}{T_\perp^2}\right].\end{aligned} \tag{53}$$

Substituting Eq. (53) into Eq. (48), neglecting terms proportional to $1/T_\perp$, which might be important under other circumstances, and taking into account that $1/T_{\parallel Z}\ll 1/T_{\parallel d}$, we obtain the well-known equations[5]

$$\begin{aligned}\frac{d\beta_Z}{dt}&=\pi\omega_1^2 f(\Delta)(\beta_d-\beta_Z)+\frac{\beta_{eq}-\beta_Z}{T_{\parallel Z}},\\\frac{d\beta_d}{dt}&=\frac{4}{9}\frac{\pi\omega_1^2\Delta^2}{\langle h^2\rangle}f(\Delta)(\beta_Z-\beta_d)-\frac{\beta_d}{T_{\parallel d}}.\end{aligned} \tag{54}$$

7. COMPARISON OF THEORY AND EXPERIMENTAL DATA

It is noteworthy that Eqs. (22)–(24) are linear in both $\sigma^\beta(h,t)$ and h, which allows us to solve them both analytically and numerically. We have obtained analytic expressions for the Fourier transforms of all projections of layer polarizations $\sigma^\beta_{0,1}(t)$, where the latter are determined by conditions (15) and (16). To compare our calculations to experi-

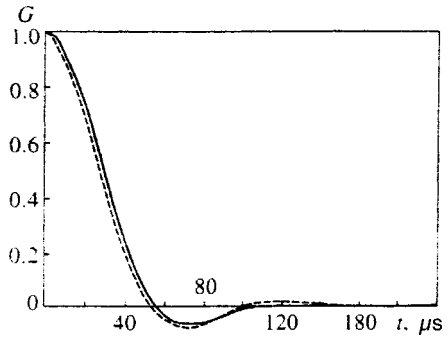

FIG. 1. Decay of free polarization in CaF_2 [111]: the solid line represents calculations and the dashed line experimental data.[10]

mental data, we have obtained a solution[11] of (22)–(24) for line shape $f(\Delta)$, which is the Fourier transform of the ratio $\sigma_0^y(t)/\sigma_{eq}^z$, and σ_{eq}^z is the equilibrium polarization constant, and we have written a computer program to calculate NMR spectra in less than a minute. All calculations have been performed for the case when $g(h)$ is a Gaussian distribution function of local fields.

Figure 1 shows measured[10] and calculated signals of free polarization decay in solid CaF_2 with the [111] orientation. The analysis of these spectra based on the theory described above has indicated that the oscillating component of decaying free polarization depends on the nonaveraged part of the isotropic dipole–dipole exchange interaction, which is characterized by the parameter $3/2-\alpha$ (oscillations vanish at $\alpha=3/2$). It describes collective coherent oscillations of dipoles, in which the polarization of individual dipoles simultaneously passes through zero.

Our calculations have revealed the effect of spectral diffusion on the oscillation amplitude, and yielded estimates of spectral diffusion times: $\tau_\perp=4\cdot10^{-4}$ s and $\tau_\parallel<10^{-5}$ s. In our calculation we have used the following values of parameters in Eqs. (22)–(24): $M_2=8.08\cdot10^8$ s^{-2}, $T_{\parallel d}=20$ s, $T_{\parallel Z}=480$ s, $T_\perp=2$ s, $\omega_1=10^{-6}$ s^{-1}, $\alpha=1.25$, crystal orientation [111].

This theory has been also used to analyze saturation experiments.[10] Figure 2 shows theoretical and experimental

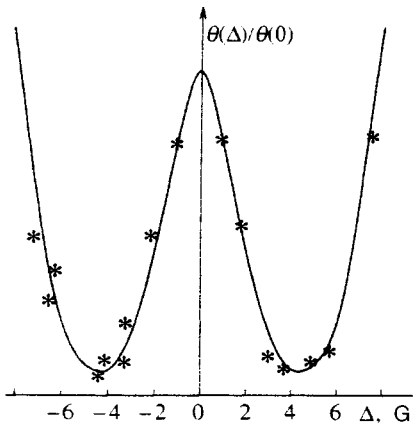

FIG. 2. NMR saturation time vs. offset in a CaF_2 [111] single crystal: the solid line shows calculations with Eq. (55) using the parameters of Fig. 1; asterisks are experimental data.[10]

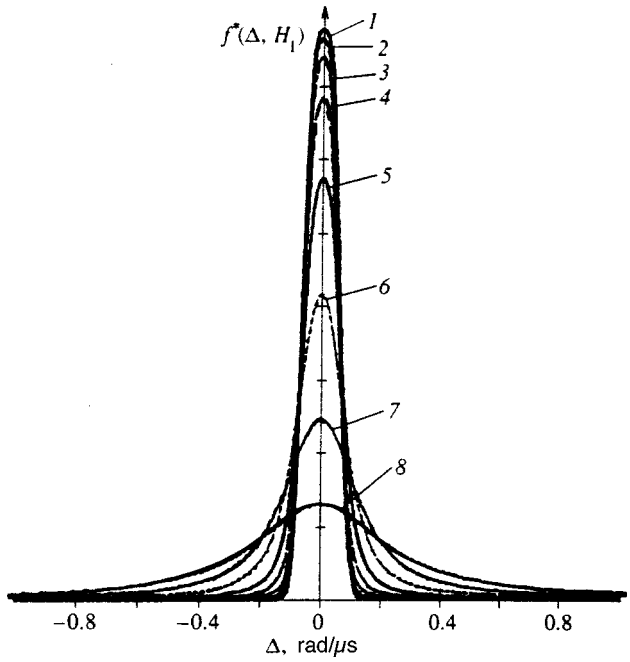

FIG. 3. Dependence of the saturated absorption line in CaF_2 [111] (normalized to unit area) on the ac resonant magnetic field amplitude H_1: (*1*) $H_1=0$; (*2*) 0.08 G; (*3*) 0.16 G; (*4*) 0.32 G; (*5*) 0.64 G; (*6*) 1.28 G; (*7*) 2.56 G; (*8*) 5.12 G. Calculations were performed with parameters of Fig. 1.

data on the saturation time θ as a function of frequency offset Δ. In Ref. 10 the data were fitted to calculations by the theory given in Ref. 5:

$$\frac{1}{\theta(\Delta)}=f(\Delta)\left(1+\frac{\Delta^2}{M_2}\right), \tag{55}$$

where $f(\Delta)$ is the absorption line profile taken directly from experiment, and M_2 is the second moment of the absorption line. In this work the function $f(\Delta)$ was calculated for the first time using Eqs. (22)–(24) with parameters corresponding to the data in Fig. 1 at $\omega_1=0.27$ G.

For comparison with the data of saturation experiments at $\omega_{\rm loc}\tau\sim1$,[4] we used Eqs. (22)–(24) and the parameters of Fig. 1 to calculate the absorption line profiles in CaF_2 with [111] orientation at various values of the saturation field H_1. The resulting curves are given in Fig. 3, and they demonstrate how the line profile changes from Gaussian to Lorentzian as H_1 increases. We have obtained the dependence of the absorption line amplitude at the resonant frequency ω_0 on ω_1 (Fig. 4). The latter graph shows that the calculations for a CaF_2 single crystal are in qualitative agreement with experimental data for Al and Cu powder,[4] which is consistent with the Redfield equations (33) derived from the generalized equation system.

8. CONCLUSIONS

In conclusion, we emphasize that NMR spectra have been calculated for the first time using consistent kinetic equations for the magnetic dipole density, based on the equation for the density matrix. Note also that such equations were first derived to describe free polarization decay[8] by intuitively simplifying formally rigorous expressions for the

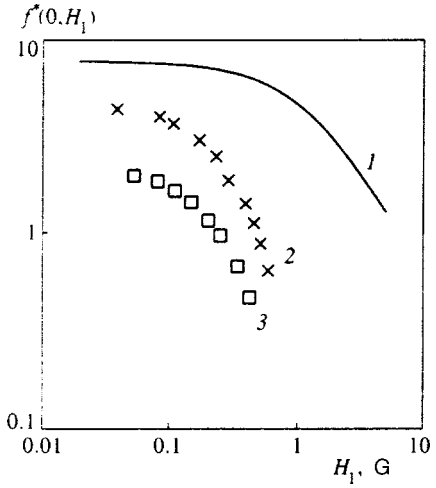

FIG. 4. Amplitude of absorption line with normalized integral as a function of the ac resonant field amplitude H_1. *1*) calculation for CaF_2 with parameters of Fig. 1; *2*, *3*) measurements of Cu and Al powder, respectively.[4]

coefficients obtained using the memory function technique, but the latter could not be calculated in practice. Similar kinetic equations were later derived by analyzing a simple dynamical process in the form of dipole precession in local dipole and external magnetic fields, based on five physical considerations.[6] The latter can be considered a justification of the kinetic equations for partial dipole densities that have been successfully used in analysis of NMR spectra in condensed media.[11,12] The theory advanced in this paper allows one to calculate in several minutes NMR spectra measured in solids using the computer program mentioned above, to derive from experimental data the characteristic spin relaxation times ($T_\perp$ and $T_\parallel$) and spectral diffusion times ($\tau_\parallel$ and $\tau_\perp$) of Zeeman and dipole–dipole reservoirs ($T_{\parallel Z}$ and $T_{\parallel d}$), the second moment M_2 of the absorption line and all higher-order moments M_i, and the parameter α, which is defined as the ratio between the second and fourth moments.[8]

Note also that the same five physical considerations allowed us to derive kinetic equations for partial dipole densities for the case of ESR[13] and interpret spectra measured in a solution of chemical radicals over a wide range of dipole concentrations.

We are grateful to V. A. Atsarkin for useful discussions of this work.

This work was supported by the Russian Fund for Fundamental Research.

*)E-mail: prov@icp.ac.ru
†)E-mail: tan@icp.ac.ru
‡)E-mail: karnaukh@icp.ac.ru

[1] A. Abraham, *The Principles of Nuclear Magnetism*, Oxford University Press (1961).
[2] Ch. Slichter, *Principles of Magnetic Resonance*, Springer-Verlag, Heidelberg (1978).
[3] A. Abraham and M. Goldman, *Nuclear Magnetism. Order and Disorder*, Clarendon Press, Oxford (1982).
[4] A. G. Redfield, Phys. Rev. **98**, 1787 (1955).
[5] B. N. Provotorov, Zh. Éksp. Teor. Fiz. **41**, 1582 (1961) [Sov. Phys. JETP **14**, 1126 (1961)].
[6] A. V. Ivanova and B. N. Provotorov, Zh. Éksp. Teor. Fiz. **107**, 473 (1995) [JETP **80**, 254 (1995)].
[7] S. J. Jensen and O. Platz, Phys. Rev. B **7**, 31 (1971).
[8] G. E. Karnaukh, A. A. Lundin, B. N. Provotorov, and K. T. Summanen, Zh. Éksp. Teor. Fiz. **91**, 2229 (1986) [Sov. Phys. JETP **64**, 1324 (1986)].
[9] G. E. Karnaukh, T. P. Kulagina, and B. N. Provotorov, in Abstracts of the XXVII Congress AMPERE, Kazan (1994), p. 40.
[10] A. E. Mefed and M. I. Rodak, Zh. Éksp. Teor. Fiz. **59**, 404 (1970) [Sov. Phys. JETP **32**, 220 (1970)].
[11] G. E. Karnaukh, T. P. Kulagina, and B. N. Provotorov, Bull. Magn. Reson. **17**, 122 (1995).
[12] T. P. Kulagina and G. E. Karnaukh, in Abstracts of the 3rd Intern. Conf. on Magn. Res., Wurzburg (1995), p. 100.
[13] E. L. Kurach and B. N. Provotorov, Khim. Fiz. **10**, 784 (1991).

Translation provided by the Russian Editorial office

The Hall effect in $La_{0.67}Ba_{0.33}MnO_3$

N. G. Bebenin,*) R. I. Zaĭnullina, V. V. Mashkautsan, A. M. Burkhanov, and V. V. Ustinov

Institute of Metal Physics, Ural Branch of the Russian Academy of Sciences, 620219 Ekaterinburg, Russia

V. V. Vasil'ev and B. V. Slobodin

Institute of Solid-State Chemistry, Ural Branch of the Russian Academy of Sciences, 620219 Ekaterinburg, Russia

(Submitted 11 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 981–987 (March 1998)

The Hall effect in polycrystalline barium-substituted lanthanum manganite $La_{0.67}Ba_{0.33}MnO_3$ has been investigated in the temperature interval $298<T<355$ K. It is found that the anomalous Hall coefficient in this material is two orders of magnitude greater than the normal coefficient. At $T_0=333$ K the normal Hall coefficient changes sign, which indicates a change in the type of conductivity. The temperature dependence of the normal Hall coefficient, electrical conductivity, and magnetoresistance is explained on the basis of the concept of motion of the mobility edge attendant as the temperature changes. © *1998 American Institute of Physics.* [S1063-7761(98)01503-0]

1. INTRODUCTION

The production of multilayer metallic films possessing a large magnetoresistance at room temperature and having definite potential for practical applications[1] has stimulated the search for other materials with similar properties. A few years ago ''colossal'' magnetoresistance (CMR) was detected in lanthanum manganite films (see the reviews in Refs. 2 and 3). Subsequent growth of interest in these half-forgotten compounds led to an upsurge of information about their magnetic and electrical properties. The effect of various factors on the magnitude of the CMR has been described in detail.[2,3] Unfortunately, an understanding of the CMR mechanism in $La_{1-x}M_xMnO_3$ (M=Ca, Sr, Ba) has not yet been achieved. Historically, the first model explaining the interrelationship between the electrical and magnetic effects in these compounds was Zener's double-exchange model;[4] however, this model fails to give a satisfactory explanation of a number of experimental facts.[5–7] Efforts to give a more adequate description of the magnetic and transport properties of lanthanum manganites have been based on very diverse suppositions. Thus, Refs. 5–7 stress the importance of allowing for lattice deformation and conclude that the current carriers in these compounds are small-radius polarons. References 2, 8, and 9 emphasize the similarity of the phenomena in $LaMnO_3$-based materials and magnetic semiconductors. A number of authors assume that the state of these materials is two-phase (ferro- and antiferromagnetic), and one of the possible mechanisms of CMR is modification of such a state by a magnetic field.[2,10]

The difference in points of view on the nature of CMR is due in part to the insufficiency of the experimental data. Usually, the results of studies of magnetic and resistive properties are analyzed. Measurements of the Hall effect could yield substantial information, but there are precious few indications of them in the literature. We are aware of only one work[11] that investigates the Hall effect in $La_{2/3}Ca_{1/3}MnO_3$ films. The results reported there are, however, difficult to understand: first, far from the Curie temperature $T_c=232$ K the hole concentration is almost an order of magnitude less than could be expected from the doping level, and second, the mobility has a maximum while the concentration has a minimum at the Curie point. It is possible that these strange results arose because the anomalous Hall effect was not separated out.

The present work is an experimental study of the Hall effect in $La_{0.67}Ba_{0.33}MnO_3$ in the region of the magnetoresistance extremum. The normal (R_0) and anomalous (spontaneous) (R_s) Hall coefficients are determined. Analysis of the temperature dependence of R_0 and R_s suggests that the transport properties of lanthanum manganites in the vicinity of the Curie point are associated with a change in the position of the mobility edge relative to the Fermi level.

2. SAMPLES AND EXPERIMENTAL TECHNIQUE

Measurements were performed in polycrystalline samples with nominal composition $La_{0.67}Ba_{0.33}MnO_3$. The constituent powders, synthesized by co-precipitation from solution,[12] were cold-pressed and then annealed in a stream of oxygen at 1200 °C for 18 h. X-ray analysis showed that the polycrystal obtained in this way is single-phase and has a lattice that can be assumed to be cubic with period $a=3.9$ Å. The Mn^{3+} and Mn^{4+} content was determined by potentiometric titration. It was found that the Mn^{4+} ions constitute 40% of the total number of manganese ions, whence the Mn^{4+} concentration is equal to 6.7×10^{21} cm^{-3}.

Samples for measurements of the Hall effect and the magnetoresistance were cut in the shape of rectangular wafers with dimensions $10\times3\times0.9$ mm^3. Magnetization mea-

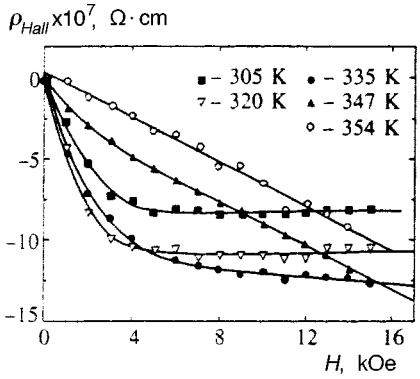

FIG. 1. Dependence of the Hall resistance on the magnetic field at different temperatures.

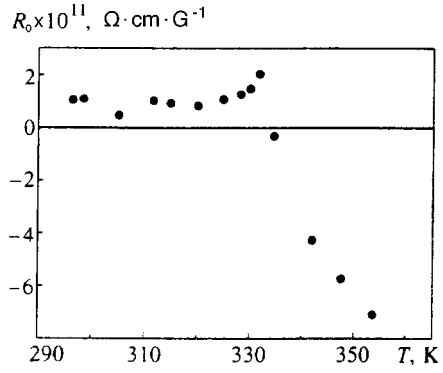

FIG. 2. Temperature dependence of the normal Hall coefficient.

surements were performed on smaller wafers with the same relative dimensions. Magnetization curves were recorded on a vibration magnetometer. Measurements of the Hall effect were performed using the four-point potentiometric method at constant current $I=50$ mA in fields up to 15 kOe for two mutually opposed directions of the magnetic field vector and two mutually opposed directions of the current.

The presence of a large magnetoresistance effect imposes rigid constraints on the quality of the contacts and the temperature stability in the measurement of the Hall voltage isotherms. Indium contacts were fashioned by ultrasound soldering. The temperature stability was 0.05 K or better.

3. EXPERIMENTAL RESULTS

The form of the magnetization curves of the $La_{0.67}Ba_{0.33}MnO_3$ sample is typical of ferromagnets. Saturation was reached in fields $H<5$ kOe. The Curie temperature, determined by the method of thermodynamic coefficients, was found to be equal to $T_c=349$ K.

The resistivity at room temperature was $\rho(T=293\ \mathrm{K})=3.5\times10^{-2}\ \Omega\cdot\mathrm{cm}$. The temperature dependence of the resistance is typical for polycrystalline samples of lanthanum manganite. In the $\rho(T)$ curve at $H=0$ a resistance peak is observed at $T_R=339$ K. Imposing a magnetic field shifts the peak toward higher temperatures. Note that the peaks in the $\rho(T)$ curves for different H correspond to the same magnetization value $M\approx140$ G.

The magnetoresistance, defined by $r=(\rho(H)-\rho(0))/\rho(0)$, is negative. The temperature dependence $r(T)$ has a minimum at $T=T_{MR}(H)$. For example, in a $H=5$ kOe field $T_{MR}=334$ K and $r(T_{MR})=-4.7\%$.

The H dependence of the Hall resistance ρ_{Hall} at different temperatures is plotted in Fig. 1. In the indicated temperature interval it satisfies $\rho_{\mathrm{Hall}}<0$. With growth of the temperature the dependence of ρ_{Hall} on H changes from a nonlinear dependence with abrupt growth in small fields to a linear dependence near T_c.

As is well known, in ferromagnets[13]

$$\rho_{\mathrm{Hall}}=R_0B+R_sM, \tag{1}$$

where B is the magnetic field induction in the sample; in our case $B\approx H$. In the investigated temperature region the magnetization depends sensitively on the field intensity. In the ferromagnetic region we can set $M(H)=M_s+\chi H$, where M_s is the spontaneous magnetization and χ is the susceptibility of the paramagnetic process. Having determined M_s and χ from the magnetization curves, we can readily find the Hall coefficients.[14] Since $M(H)$ and $\rho_{\mathrm{Hall}}(H)$ are known, R_0 and R_s can be determined in a different way by constructing the dependence of ρ_{Hall}/H on M/H. In this way it is also possible to analyze the results in the paramagnetic region.

To find the Hall coefficients, we used the results of measurements of $M(H)$ and $\rho_{\mathrm{Hall}}(H)$ in the region $5\leqslant H\leqslant15$ kOe. The two methods for finding R_0 and R_s give identical results. Figures 2 and 3 plot the temperature dependence of the Hall coefficients. The normal Hall coefficient for $T\leqslant325$ K is positive, essentially independent of temperature, and equal to $R_0=(9\pm2)\times10^{-12}\ \Omega\cdot\mathrm{cm}\cdot\mathrm{G}^{-1}$. As the temperature is increased past 325 K, R_0 at first grows rapidly, reaches a sharp maximum, and then abruptly falls to zero and changes sign at $T_0\approx333$ K. With further increase of the temperature, the normal Hall coefficient remains negative and grows in magnitude.

The anomalous Hall coefficient is negative in the ferromagnetic temperature region, with $|R_s|$ exceeding $|R_0|$ by two orders of magnitude. At $T=T_0$ the anomalous Hall coefficient has a wide minimum, and at $T=T_c$ changes sign.

4. DISCUSSION

In Refs. 15 and 16 it was shown that in the ferromagnetic region the resistance of polycrystals of lanthanum manganite is determined mainly by intercrystal boundaries, and

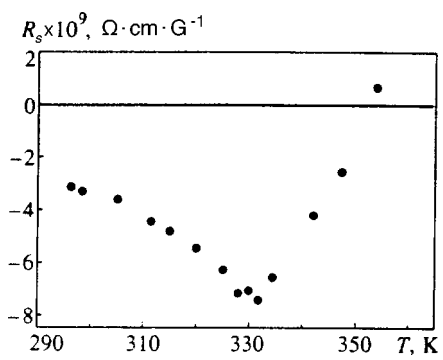

FIG. 3. Temperature dependence of the anomalous Hall coefficient.

the resistance of polycrystalline samples can exceed the resistance of single-crystal samples by two to three orders of magnitude. Under these conditions the Hall effect is governed mainly by crystallites.[14]

Calculating the hole concentration n_p in the temperature region $T \leqslant 325$ K with the help of the formula $R_0 = (ecn_p)^{-1}$, we find that $n_p = (7 \pm 2) \times 10^{21}$ cm^{-3}, which essentially coincides with the Mn^{4+} ion concentration. Assuming that the resistance of the crystallites is one to two orders less than that of the investigated polycrystal, we obtain an estimate for the hole mobility on the order of 1 cm$^2 \cdot$V$^{-1} \cdot$s^{-1}. These two estimates allow us to conclude that for $T < T_0$ the conductivity of $La_{0.67}Ba_{0.33}MnO_3$ is determined by band charge carriers undergoing strong scattering. This conclusion and also the large magnitude of the anomalous Hall coefficient are in agreement with the result that the valence band of $LaMnO_3$, as band calculations show (see, e.g., Ref. 17), is formed mainly by manganese d states and has width $W \sim 1$ eV.

The change in the nature of the temperature dependence and sign of the normal Hall coefficient is evidence of a change in the type of conductivity at $T = T_0$. Optical measurements also indicate different types of conductivity in the ferro- and paramagnetic regions.[18] In our opinion, the change in the type of conductivity is due to movement of the mobility edge in the vicinity of the Curie temperature. Since the valence band of $LaMnO_3$ is narrow, the carrier states are very sensitive to the presence of disorder. Disorder due to the presence of impurity atoms, vacancies, etc., can be assumed to be temperature-independent, whereas magnetic disorder due to fluctuations of the magnetization depends on T. For the case of a narrow-band ferromagnetic semiconductor the position of the mobility edge (taking only magnetic disorder into account) was calculated in Ref. 19 according to the $s-d$ model. It was shown that as the temperature is increased from the ferromagnetic region to the paramagnetic region, magnetic fluctuations cause the mobility edge to shift by an amount comparable with the width of the band. Although the results obtained in this work cannot be applied directly to an analysis of the conduction mechanisms in lanthanum manganites, even in combination with familiar general ideas of the physics of disordered systems,[20–22] they allow a qualitative explanation of the above experimental facts.

At $T = 0$ disorder is due to the presence of a dopant impurity and lattice defects. The mobility edge E_c in this case is located inside the band (see Fig. 4) so that $E_c(T=0) < E_{\text{top}}$, where E_{top} is the energy corresponding to the top of the valence band (for simplicity we ignore band deformation and formation of density-of-states tails). With growth of the temperature the mobility edge shifts toward the center of the band due to the appearance of magnetic disorder, and, according to Ref. 19, a large part of the shift takes place already in the ferromagnetic temperature region. The largest value of the difference $E_{\text{top}} - E_c$ is reached in the far paramagnetic region. If nonmagnetic disorder is absent, then in the case of a narrow-band material $E_c(T=0) - E_c(T \gg T_c) = \alpha W$ since W is the smallest energy parameter; in Ref. 19 the value $\alpha = 1/4$ was obtained. A magnetic field suppresses magnetic disorder and, consequently, shifts the mobility edge away from the center toward the top of the valence band.

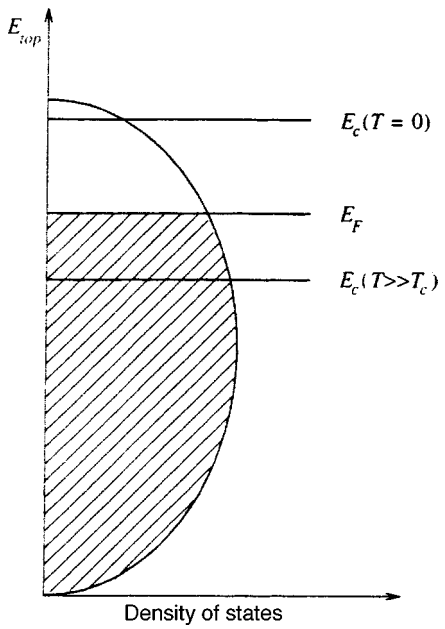

FIG. 4. Diagram of the location of the Fermi level and the mobility edge.

The Fermi level E_F in our case is found inside the valence band. Utilizing the elliptic model of the density of states, it easy to obtain an estimate for the Fermi energy: $E_{\text{top}} - E_F \approx 0.4W$, which is similar to the difference $E_c(T=0) - E_c(T \gg T_c)$. Consequently, movement of the mobility edge can lead to a large change in the number of carriers in the delocalized states.

The conductivity for $T < 325$ K has a metallic character, whence it follows that the Fermi level at these temperatures is located far from the mobility edge. As T_c is approached, because of the shift of E_c the difference $E_c - E_f$ decreases, which leads to a decrease in the number of carriers in the delocalized states and, consequently, to a growth of the normal Hall coefficient. At some temperature T_{MS} the mobility edge crosses the Fermi level, i.e., a metal–semiconductor transition occurs. If the temperature exceeds T_{MS}, then the Fermi level lies in the region of localized states and the conductivity does not have a bandlike character. It is well known[21] that if the conductivity is due to activation of current carriers to the mobility edge, then the normal Hall coefficient is negative as a rule, regardless of whether the carriers are electrons or holes. Consequently, the change of sign of R_0 from positive at $T < T_0$ to negative at $T > T_0$ observed in our samples can be explained by a transition at $T = T_0$ from bandlike conductivity to conductivity due to activation of holes to the mobility edge; thus $T_{MS} = T_0$.

The change in the type of conductivity is also reflected in the temperature dependence of the anomalous Hall coefficient, as is indicated by the presence of a minimum at $T = T_0$ in the $R_s(T)$ curve.

Near T_{MS} growth of the temperature leads to a rapid growth of the resistance. On the other hand, at values of T noticeably above T_{MS}, the difference $E_F - E_c$ varies weakly; therefore the resistance is of an activation nature with activation energy equal to this difference, and decreases with

growth of T. Consequently, the resistance peak should be found at some temperature $T_R > T_{MS}$, as is observed.

A magnetic field shifts the mobility edge toward the top of the band. In a weak magnetic field the change in the resistance caused by this shift should be greatest at the metal–semiconductor transition; therefore the extremum in the magnetoresistance should be found at $T = T_{MS}$. The investigated samples have $T_{MR} = 334$ K for $H = 5$ kOe, i.e., T_{MR} indeed essentially coincides with $T_{MS} = 333$ K.

5. CONCLUSION

The study reported here of the Hall effect in polycrystalline samples of $La_{0.67}Ba_{0.33}MnO_3$ shows that the anomalous Hall effect in this material greatly exceeds the normal effect. In the ferromagnetic temperature region far from the Curie point the normal Hall coefficient R_0 is positive, its value is essentially independent of temperature, and the hole concentration n_p, determined with the help of R_0, is near the concentration of Mn^{4+} ions. At $T = 333$ K the normal Hall coefficient changes sign; as the temperature is increased further R_0 remains negative but its magnitude increases. Such behavior of $R_0(T)$ indicates that at $T = 333$ K a change in the type of conductivity takes place.

The observed temperature dependence of the normal Hall coefficient, the electrical resistance, and the magnetoresistance can be explained in terms of movement of the mobility edge as the temperature varies, changing the number of current carriers in the delocalized states. Such a mechanism for the change in the type of conductivity in the vicinity of T_c is apparently a general feature of all heavily doped materials based on $LaMnO_3$. It should be noted in passing that at low doping levels the situation may be more complicated.

ACKNOWLEDGMENTS

This work was carried out with the support of the Russian Fund for Fundamental Research (Project No. 97-02-16008).

*)E-mail: elph@ifm.e-burg.su

[1] R. L. White, IEEE Trans. Magn. **30**, 346 (1994).
[2] R. von Helmolt, J. Wecker, K. Samwer, and K. Barner, J. Magn. Magn. Mater. **151**, 411 (1995).
[3] É. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996)
[4] C. Zener, Phys. Rev. **82**, 403 (1951).
[5] A. J. Millis, P. B. Littlewood, and B. I. Shraiman, Phys. Rev. Lett. **74**, 5144 (1995).
[6] A. J. Millis, R. Mueller, and B. I. Shraiman, Phys. Rev. B **54**, 5389 (1996).
[7] A. J. Millis, R. Mueller, and B. I. Shraiman, Phys. Rev. B **54**, 5405 (1996).
[8] R. von Helmolt, J. Wecker, B. Holzapfel, L. Schulz, and K. Samwer *et al.*, Phys. Rev. Lett. **71**, 2331 (1993).
[9] A. Nossov, J. Pierre, V. Vassiliev, and V. Ustinov, Solid State Commun. **101**, 361 (1997).
[10] L. I. Koroleva, R. V. Demin, A. M. Balbashov, JETP Lett. **65**, 474 (1997).
[11] J. E. Nuñez-Regueiro, D. Gupta, and A. M. Kadin, J. Appl. Phys. **79**, 5179 (1996).
[12] V. G. Vasil'ev, A. A. Ivakin, and A. A. Fotiev, Zh. Neorg. Khim. **39**, 3 (1994).
[13] S. V. Vonsovskiĭ, *Magnetism*, Vols. 1 and 2 (Wiley, New York, 1974).
[14] E. V. Kuchis, *Galvanomagnetic Effects and Methods of Their Study* [in Russian] (Radio i Svyaz', Moscow, 1990).
[15] H. Y. Hwang, S.-W. Cheong, N. P. Ong, and B. Batlogg, Phys. Rev. Lett. **77**, 2041 (1996).
[16] A. Gupta, G. Q. Gong, Gang Xiao, P. R. Duncombe, P. Lecoeur, P. Trouilloud, Y. Y. Wang, V. P. Dravid, and J. Z. Sun, Phys. Rev. B **54**, R15 629 (1996).
[17] S. Satpathy, Z. S. Popović, and F. R. Vukajlović, J. Appl. Phys. **79**, 4555 (1996).
[18] N. N. Loshkareva, Yu. P. Sukhorukov, A. P. Nosov *et al.*, Fiz. Tverd. Tela **39**, 1616 (1997) [Phys. Solid State **39**, 1440 (1997)].
[19] E. M. Kogan and M. I. Auslender, Phys. Status Solidi B **147**, 613 (1988).
[20] J. M. Ziman, *Models of Disorder: The Theoretical Physics of Homogeneously Disordered Systems* (Cambridge University Press, Cambridge, 1979).
[21] N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (Clarendon Press, Oxford, 1971).
[22] N. F. Mott, Adv. Phys. **34**, 329 (1985).

Translated by Paul F. Schippnick

Spectrum of plasma oscillations in structures with a periodically inhomogeneous two-dimensional electron plasma

O. R. Matov, O. F. Meshkov, and V. V. Popov*)

Saratov Branch of the Radio Engineering and Electronics Institute, Russian Academy of Sciences, 410019 Saratov, Russia
(Submitted 16 May 1997)
Zh. Éksp. Teor. Fiz. **113**, 988–999 (March 1998)

We develop a rigorous electrodynamic theory of plasma oscillations in a periodically inhomogeneous two-dimensional electron system with a rectangular profile of the spatial modulation of the equilibrium electron concentration. We calculate the frequencies and radiative damping of the main plasma-oscillation types with a zero reduced wave vector. We show that the frequency splitting and the radiative damping of the oscillations are nonmonotonic functions of the modulation percentage and the ratio of the widths of the bands of two-dimensional plasma with low and high electron concentrations. The results of calculations are compared with experimental data and with the results of a perturbation theory developed in earlier work of other researchers. We discuss the physical mechanism for the emergence of radiative damping of plasma oscillations. © *1998 American Institute of Physics.* [S1063-7761(98)01603-5]

1. INTRODUCTION

Plasma oscillations in periodically inhomogeneous two-dimensional ($2D$) electron systems have been experimentally investigated by many researchers[1–7] using submillimeter Fourier transform spectroscopy. The experimenters used metal–insulator–semiconductor (MIS) structures based on p-Si (Refs. 1–4) and GaAs/AlGaAs heterostructures (Refs. 5–7). Periodic spatial modulation of the density of a $2D$ electron plasma, N_s, in the inversion channel of the MIS structure or in the GaAs/AlGaAs heterojunction was created by the field effect produced by applying a bias voltage (positive in the MIS structure and negative in the GaAs/AlGaAs heterojunction) to a periodic gate electrode, which is a solid semitransparent (for electromagnetic waves) periodically corrugated conducting NiGr layer[1–6] or a metallic (Al) grid.[7] The typical modulation period L is about (0.5–1) μm. Plasma oscillations manifest themselves in the form of resonances in the transmission spectrum of electromagnetic radiation passing through the structure.

The main results of these experimental studies can be summarized as follows. As the percentage of spatial modulation of the density of the $2D$ electron plasma increases with the average surface electron number density $\overline{N_s}$ remaining constant, the plasma oscillation frequencies decrease. Note that in the case of a homogeneous $2D$ layer ($\overline{N_s}=N_s$), the plasma frequency ω_p is proportional to $\sqrt{\overline{N_s}}$ (Refs. 8 and 9):

$$\omega_p=\sqrt{N_s e^2 k/2m^*\varepsilon_0\overline{\varepsilon}}, \tag{1}$$

where k is the plasmon wave number, e and m^* are the electron charge and effective mass, ε_0 is the permittivity of free space, and $\overline{\varepsilon}$ is the effective dielectric constant, which depends on the geometry of the structure. The resonances observed in the experiments described in Refs. 1–7 correspond to excitation of plasma oscillations with wave numbers $k=k_n=2n\pi/L$ $(n=1,2,3,\ldots)$.

In a periodically inhomogeneous $2D$ electron system, aside from a decrease in plasma oscillation frequencies, there is a splitting of plasma resonances due to the emergence of forbidden frequency minibands in the continuous plasmon spectrum (1). Here the splitting of resonances manifests itself most vividly in structures with asymmetric profiles of the electron density distribution.[3]

As the modulation percentage increases, the continuity of the $2D$ electron system is violated eventually and a periodic system of isolated quasi-one-dimensional electron channels emerges.[5,6] As the average surface electron density decreases, the frequency of plasma oscillations of electrons localized in these channels increases.

Krasheninnikov and Chaplik[10] were the first to undertake a theoretical study of plasma oscillations in a periodically inhomogeneous $2D$ electron plasma in an approximation in which the spatial modulation of the equilibrium electron density was assumed weak. According to Ref. 10, the spectrum of plasma oscillations at $k_n=2n\pi/L$ $(n=1,2,3,\ldots)$ is described by

$$\omega_n^{\pm}=\omega_p^{(0)}\left(1\pm|N_{2n}|/N_0\right)^{1/2}, \tag{2}$$

where ω_n^+ and ω_n^- are the upper and lower edges of the $2n$th forbidden frequency band (odd-numbered forbidden bands appear at $k_n=(2n-1)\pi/L$); the N_{2n} are the Fourier expansion coefficients for the periodic electron density distribution in the $2D$ system, with $N_{2n}\ll N_0$, $N_0=\overline{N_s}$; and $\omega_p^{(0)}$ is given by Eq. (1) with N_s replaced by N_0. Equation (2) shows that the perturbation theory approach does not predict a decrease

in the average frequency (the ''softening'' effect) of plasma oscillations, $(\omega_n^+ + \omega_n^-)/2$ as the modulation percentage grows. Moreover, the splitting of plasma resonances, $\Delta\omega_1 = \omega_1^+ - \omega_1^-$, observed in the experiments described in Refs. 1–3 exceeds (by a factor greater than two in some experiments) the value calculated by (2).

Eliasson *et al.*[11] were the first to theoretically substantiate the ''softening'' effect of plasma oscillations that appears as the modulation percentage grows. Further theoretical investigations of Cataudella and Ramaglia[12] and Meshkov[13] suggested that the frequencies $\omega_n^{\pm}$ drop to zero when the modulation amplitude reaches unity (which corresponds to continuity of the $2D$ system being violated). The rigorous quantum mechanical calculations done by Wulf *et al.*[14] support the conclusion, obtained by the classical hydrodynamic approach, that in a continuous $2D$ electron system the plasma oscillation frequency decreases as the percentage of spatial modulation grows.

The increase in the frequency of plasma oscillations of electrons localized in a system of isolated quasi-one-dimensional channels with the decrease in the average surface electron density was explained by Shikin *et al.*,[15] while Schaich *et al.*[16] built an approximate theory of the passage of electromagnetic waves through a structure containing periodically inhomogeneous $2D$ electron plasma for arbitrary modulation percentages.

In all of the above papers, plasma oscillations are considered in the electrostatic approximation. However, as noted earlier, plasma oscillations observed in experiments involving periodically inhomogeneous $2D$ electron systems have wave numbers $k_n = 2n\pi/L$ $(n = 1,2,3,\ldots)$, which corresponds to the center of the first Brillouin zone $(k = 0)$ in the reduced-zone scheme for a periodic structure. Hence the plasma oscillations prove to be related to homogeneous (in the plane of the $2D$ system) fields of transverse electromagnetic waves. Because of this relation, absorption (or emission) of electromagnetic waves by plasma oscillations in periodic structures is possible. A meaningful description of oscillations in a $2D$ plasma requires a rigorous electrodynamic approach. The characteristic of the relation between plasma oscillations and electromagnetic radiation is the radiative damping of such oscillations.[17] It is understood that the electrostatic approximation cannot specify the magnitude of the radiative damping of one or another plasma mode.

In Ref. 18–21, a rigorous electrodynamic theory of plasma oscillations in a homogeneous $2D$ electron plasma was developed. The theory takes into account the relation of such oscillations to electromagnetic radiation in structures with a periodic metallic lattice, including the case where there is an external magnetic field perpendicular to the plane of the $2D$ electron system.[22] The method based on the scattering-matrix technique[18] was used by Wilkinson *et al.*[7] in their studies of plasma oscillations in a periodically inhomogeneous $2D$ electron system. Their calculations made it possible to explain quantitatively the data gathered in the experiment of the same researchers,[7] who employed submillimeter Fourier transform spectroscopy to study $2D$ plasma oscillations in GaAs/AlGaAs structures with a gate electrode in the form of a periodic metallic lattice. In particular, they corroborated theoretically the reduction of the $2D$ plasma oscillation frequency with an increase in the percentage of spatial modulation of the electron density.

However, the results obtained in Ref. 7 cannot be used directly to explain the results of the experiments described in Refs. 1–3, 5, and 6, where the electron density in the $2D$ system was modulated by employing a semitransparent periodically corrugated solid gate electrode. Moreover, the calculations in Ref. 7 were done for a fixed ratio of the widths of the bands of two-dimensional plasma with low and high electron concentrations, which makes the description of many features of the plasmon spectrum in a periodic $2D$ system impossible.

As noted earlier, the splitting of plasma resonances observed in experiments has so far been described in the perturbation-theory setting in the electrostatic approximation.

To carry out a theoretical analysis of the features of plasma oscillations in periodically inhomogeneous $2D$ electron systems in connection with the above experiments and to establish the limits of the perturbation-theory approach, the rigorous electrodynamic theory developed by the present authors in their theoretical study of $2D$ plasma oscillations in a semiconductor heterostructure with a lateral periodic metallic lattice has been advanced still further.[18,19]

In Sec. 2 we describe the theoretical model and give the main relations. Section 3 is devoted to examining the results of calculations and comparing them with the existing experimental and theoretical data. In Sec. 4 we summarize the conclusions that follow from our results.

2. MODEL AND BASIC RELATIONS

Suppose that a $2D$ electron plasma occupies the space at the surface of a substrate (the $y = 0$ plane) with a dielectric constant ε_1. The equilibrium surface electron concentration in the plane of the $2D$ system is a periodic function of position x, $N_s(x) = N_s(x+L)$, of the form

$$N_s(x) = \begin{cases} N_A & \text{if } 0 < x < w, \\ N_B & \text{if } w < x < L. \end{cases} \tag{3}$$

We assume that above the $2D$ plasma there is a layer of an insulator of thickness d, and above the insulator (for $y > d$) there is a medium, with the dielectric constants ε_2 and ε_3, respectively.

In accordance with the conditions of the experiments described in Refs. 1–3, 5, and 6, we assume that the desired electric and magnetic fields are uniform along the bands of $2D$ plasma (along the z axis), and limit ourselves to TM polarization, which means that the nonzero components of the electric field are E_x and E_y and the nonzero component of the magnetic field is H_z. For TE polarization the component E_x is zero, so that fields with TE polarization cannot interact with longitudinal plasma oscillations in the direction x observed in the experiments.[1–3,5,6]

We write the desired fields and currents in the periodic structure in the form of expansions in spatial Fourier transforms. For instance, for the z-component of the magnetic field we have

$$H_z(x,y)=\sum_{m=-\infty}^{\infty} H_{zm}(y)\exp(-i\beta_m x), \tag{4}$$

where

$$H_{zm}(y)=\frac{1}{L}\int_0^L H_z(x,y)\exp(i\beta_m x)dx$$

are the amplitudes of the spatial Fourier transforms, and

$$\beta_m=k+2m\pi/L,$$

with k the wave number reduced to the center of the first Brillouin zone. The amplitudes of the Fourier components of the magnetic field in the three different media can be written

$$H_{zm}^{(1)}(y)=A_m\exp(\alpha_m^{(1)}y), \quad y\leqslant 0, \tag{5}$$

$$H_{zm}^{(2)}(y)=B_m\sinh(\alpha_m^{(2)}y)+C_m\cosh(\alpha_m^{(2)}y), \quad 0\leqslant y\leqslant d, \tag{6}$$

$$H_{zm}^{(3)}(y)=D_m\exp(-\alpha_m^{(3)}y), \quad y\geqslant d. \tag{7}$$

Here $(\alpha_m^{(j)})^2=\beta_m^2-k_0^2\varepsilon_j$ $(j=1,2,3)$, $k_0=\tilde{\omega}\sqrt{\varepsilon_0\mu_0}$, μ_0 is the permeability of free space, and A_m, B_m, C_m, and D_m are constants. Generally, the dielectric constants ε_j are complex-valued, $\varepsilon_j=\varepsilon_j'+i\varepsilon_j''$, where the ε_j'' are determined by the dielectric losses in the corresponding media.

Writing the time dependence in exponential form, $\exp(i\tilde{\omega}t)$, Maxwell's equations yield

$$E_{xm}^{(j)}=-\frac{i\tilde{\omega}\mu_0}{k_0^2\varepsilon_j}\frac{\partial H_{zm}^{(j)}}{\partial y}. \tag{8}$$

Let us write the boundary conditions at the surfaces $y=0$ and $y=d$ as follows:

$$E_{xm}^{(1)}=E_{xm}^{(2)}, \quad H_{zm}^{(2)}-H_{zm}^{(1)}=I_{xm} \tag{9}$$

at $y=0$, and

$$E_{xm}^{(2)}=E_{xm}^{(3)}, \quad H_{zm}^{(3)}-H_{zm}^{(2)}=0 \tag{10}$$

at $y=d$, where the I_{xm} are the amplitudes of the spatial Fourier transforms of the current density in the $2D$ electron plasma.

Using Eqs. (5)–(10), we can derive the following relation linking the electric field and the surface current density in the $y=0$ plane:

$$E_{xm}(0)=G_m I_{xm}. \tag{11}$$

The surface impedance G_m is given by

$$G_m=iZ_0\frac{\chi_m^{(3)}/\chi_m^{(2)}+\coth(\alpha_m^{(2)}d)}{\chi_m^{(2)}+\chi_m^{(1)}\chi_m^{(3)}/\chi_m^{(2)}+(\chi_m^{(1)}+\chi_m^{(3)})\coth(\alpha_m^{(2)}d)}, \tag{12}$$

where $\chi_m^{(j)}=\varepsilon_j k_0/\alpha_m^{(j)}$, and $Z_0\simeq 377\ \Omega$ is the impedance of free space.

On the other hand, for the current density in the $2D$ electron system we have

$$I_x(x)=\sigma_A E_x(x,0) \quad \text{if } 0<x<w, \tag{13}$$

$$I_x(x)=\sigma_B E_x(x,0) \quad \text{if } w<x<L. \tag{14}$$

In the local approximation (the Drude model), the surface conductivities σ_A and σ_B can be written

$$\sigma_{A,B}=\frac{e^2 N_{A,B}}{m^*}\frac{\tau}{1+i\tilde{\omega}\tau}, \tag{15}$$

where τ is the phenomenological relaxation time of the electron momentum in the $2D$ plasma.

If we now go to the Fourier representation in (13) and (14) and allow for (11), we obtain

$$\sum_{m=-\infty}^{\infty}(1-G_m\sigma_A)I_{xm}\exp(-i\beta_m x)=0, \quad 0<x<w, \tag{16}$$

$$\sum_{m=-\infty}^{\infty}(1-G_m\sigma_B)I_{xm}\exp(-i\beta_m x)=0, \quad w<x<L. \tag{17}$$

Bearing in mind that

$$I_{xm}=\frac{1}{L}\int_0^L I_x(x)\exp(i\beta_m x)dx,$$

we can write (16) and (17) as integral equations for the surface current density:

$$\sum_{m=-\infty}^{\infty}(1-G_m\sigma_A)\exp(-i\beta_m x)\int_0^L I_x(\zeta)\exp(i\beta_m\zeta)d\zeta=0, \quad 0<x<w, \tag{18}$$

$$\sum_{m=-\infty}^{\infty}(1-G_m\sigma_B)\exp(-i\beta_m x)\int_0^L I_x(\zeta)\exp(i\beta_m\zeta)d\zeta=0, \quad w<x<L. \tag{19}$$

Next we write the distribution of the current density over one period of the structure as

$$I_x(x)=I_{xA}(x)+I_{xB}(x),$$

where $I_{xA}(x)=0$ if $w<x<L$, and $I_{xB}(x)=0$ if $0<x<w$.

The functions $I_{xA}(x)$ and $I_{xB}(x)$ can be approximated in the intervals $0<x<w$ and $w<x<L$ by the expansions

$$I_{xA}(x)=\exp(-ikx)\sum_{n=0}^{N}p_n P_n(x'), \tag{20}$$

$$I_{xB}(x)=\exp(-ikx)\sum_{n=0}^{N}q_n P_n(x''), \tag{21}$$

where $P_n(x')$ and $P_n(x'')$ are Legendre polynomials of the first kind of degree n defined, respectively, in the intervals $0<x<w$ and $w<x<L$, $x'=2x/w-1$, $x''=2(x-w)/s-1$, $s=L-w$, and p_n and q_n are unknown constant coefficients. Note that the functions $I_{xA,B}(x)\exp(ikx)$ are the complex-valued amplitudes of the current densities $I_{xA,B}(x)$ in the $2D$ electron system; they are periodic functions of position x with period L.

We write the current density in Eqs. (18) and (19) in the form of the sum of the expansions (20) and (21) and employ the Galerkin procedure[23] for solving the integral equations (18) and (19), where for the orthogonal basis functions we

take the Legendre polynomials $P_n(x')$ and $P_n(x'')$ in the corresponding intervals. This procedure makes it possible to pass from the system of integral equations (18) and (19) for the function $I_x(x)$ to a system of $2(N+1)$ homogeneous algebraic equations for the coefficients p_n and q_n:

$$\sum_{n=0}^{N} A_{kn}p_n+\sum_{n=0}^{N} B_{kn}q_n=0, \quad k=0,1,2,\ldots,N,$$

$$\sum_{n=0}^{N} C_{kn}p_n+\sum_{n=0}^{N} D_{kn}q_n=0, \quad k=N+1,\ldots,2N+1, \quad (22)$$

where

$$A_{kn}=wi^n\sum_{m=-\infty}^{\infty} j_k\left(\frac{m\pi w}{L}\right)j_n\left(\frac{m\pi w}{L}\right)(1-G_m\sigma_A),$$

$$B_{kn}=si^n\sum_{m=-\infty}^{\infty} (-1)^m j_k\left(\frac{m\pi w}{L}\right)j_n\left(\frac{m\pi s}{L}\right)(1-G_m\sigma_A),$$

$$C_{kn}=wi^n\sum_{m=-\infty}^{\infty} (-1)^m j_k\left(\frac{m\pi s}{L}\right)j_n\left(\frac{m\pi w}{L}\right)(1-G_m\sigma_B),$$

$$D_{kn}=si^n\sum_{m=-\infty}^{\infty} j_k\left(\frac{m\pi s}{L}\right)j_n\left(\frac{m\pi s}{L}\right)(1-G_m\sigma_B). \quad (23)$$

The spherical Bessel functions of the first kind of degree n, $j_n(m\pi w/L)$ and $j_n(m\pi s/L)$, emerge in Eqs. (23) as a result of calculations of integrals of type[24]

$$\int_{-1}^{1} P_n(\zeta)\exp(i\rho\zeta)d\zeta=2i^n j(\rho),$$

which appear after substituting the expansions (20) and (21) into Eqs. (18) and (19).

The dispersion relation, which connects the frequency and the reduced wave numbers of the natural oscillations of the system, is determined from the condition that the determinant of the system of linear homogeneous equations (22) be zero. For a fixed real value of the wave number k, the roots of the dispersion relation yield the values of the complex-valued frequencies $\tilde{\omega}=\omega+i\gamma$. The real part ω is the frequency of the natural oscillations, and the imaginary part γ is the damping coefficient for these oscillations, the damping being due both to dissipative losses and electromagnetic radiation emitted by the structure (radiative losses). If we neglect dissipative losses, then $\gamma=\gamma_r$, where γ_r is the radiative damping coefficient.

The Galerkin procedure and the series in (23) are convergent. Below we give the results of numerical calculations for $N=4$ and for terms with $|m|\leqslant 30$ retained in (23), which makes it possible to ensure 1% accuracy in the calculated natural frequencies.

3. RESULTS AND DISCUSSION

In this section we discuss calculations of a $2D$ electron system with a periodic distribution profile for the equilibrium electron concentration of the form (3), for the following values of the modulation amplitude:

$$\Delta n_s=\frac{N_A-N_B}{2\overline{N_s}}<1, \quad N_A>N_B>0.$$

In accordance with the experimental situation, we study the plasma oscillations with a zero wave vector k. Such oscillations manifest themselves in the form of plasma resonances in the transmission spectra of the external electromagnetic wave (uniform in the plane of the $2D$ system) passing through the structures in the experiments described in Refs. 1–3, 5, and 6.

The presence of a periodic distribution profile of the equilibrium electron concentration in the $2D$ system splits each plasma oscillation with a wave number $k_n=2n\pi/L$ (which corresponds to $k=0$ in the reduced-zone scheme) into two oscillations with different frequencies $\omega_n^{\pm}$. In the adopted case of a profile of the equilibrium electron concentration (Eq. (3)) that is symmetric with respect to the centers of the intervals $0<x<w$ and $w<x<L$, one oscillation is nonradiative ($\gamma_r=0$) and the other is characterized by nonzero radiative damping. Obviously, only radiative modes can interact directly with an external electromagnetic field and hence can be observed in experiment.

The magnitude of radiative damping rapidly decreases as n grows, so that plasma resonances with $n=1$ are the most degenerate resonances in the transmission spectrum of the electromagnetic wave. In this connection, below we consider only the fundamental (i.e., with the lowest frequencies) plasma oscillations with $n=1$ and drop the label 1 for convenience.

Figure 1 depicts the results of calculating the frequencies and radiative damping coefficients of plasma oscillations as functions of the parameter w/L obtained without allowance for dissipative losses in the structure ($\tau\to\infty$ and $\varepsilon_j''=0$). The values of the other parameters are characteristic of MIS structures based on p-Si that are used in experiments.[1–3] The average surface electron concentration $\overline{N_s}=(N_Aw+N_Bs)/L$ is assumed constant. The distributions of the electric field component $E_x(x)$ and of the oscillations of the surface charge density $\rho(x)$ in the plane of the $2D$ system are depicted in Fig. 2 for radiative and nonradiative oscillations.

Figure 1 shows that the frequencies of the plasma oscillations decrease as the percentage of the modulation of the equilibrium electron concentration grows over essentially the entire range of the parameter w/L, due to the localization of the oscillation field in the region of the $2D$ plasma with the lower electron concentration (Fig. 2). Naturally, at $w/L=0$ and $w/L=1$ the frequencies ω^+ and ω^- coincide and are equal to the frequency of plasma oscillations in a structure with a homogeneous $2D$ electron layer with the electron concentration $N_s=\overline{N_s}$. Theoretically, the frequencies of plasma oscillations tend to zero as $\Delta n_s\to 1$ ($N_B\to 0$). These conclusions agree with earlier theoretical results.[7,11–13]

The frequency splitting $\Delta\omega=\omega^+-\omega^-$ of the radiative and nonradiative oscillations is a nonmonotonic function of the parameter w/L. For small values of w/L the radiative oscillation frequency is higher than the frequency of the nonradiative mode, while for large values of w/L the situation is just the opposite. The value of w/L at which the oscillations

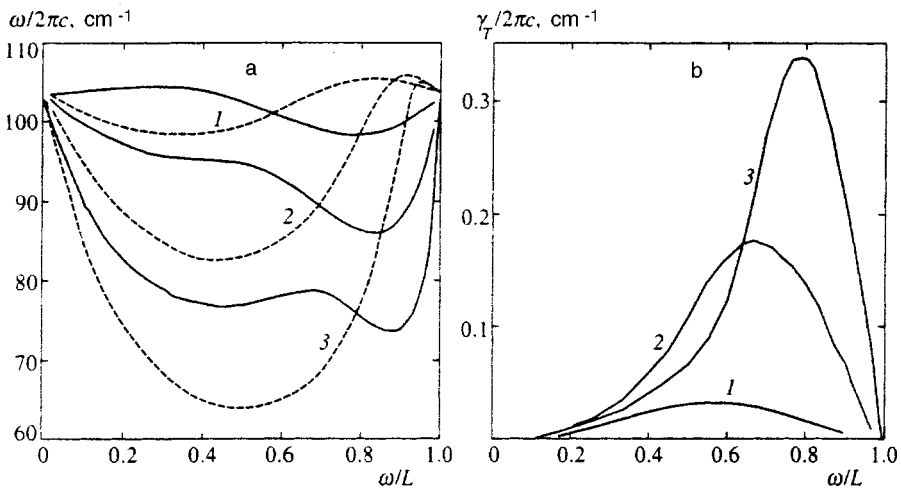

FIG. 1. The w/L-dependence of the frequencies (a) and radiative damping coefficients (b) for radiative (solid curved) and nonradiative (dashed curves) plasma oscillations for various values of the amplitude of the spatial modulation of the electron concentration $\Delta n_s = 0.2$ (curves *1*), 0.5 (curves *2*), and 0.7 (curve *3*). The parameters of the structure are $\varepsilon_1 = 11.45$, $\varepsilon_2 = 3.9$, $\varepsilon_3 = 1$, $\bar{N}_s = 3 \times 10^{12}$ cm^{-2}, $L = 6 \times 10^{-5}$ cm, $d = 2 \times 10^{-6}$ cm, and $m^* = 0.2\ m_e$, with $c = 1/\sqrt{\varepsilon_0 \mu_0}$ the speed of light.

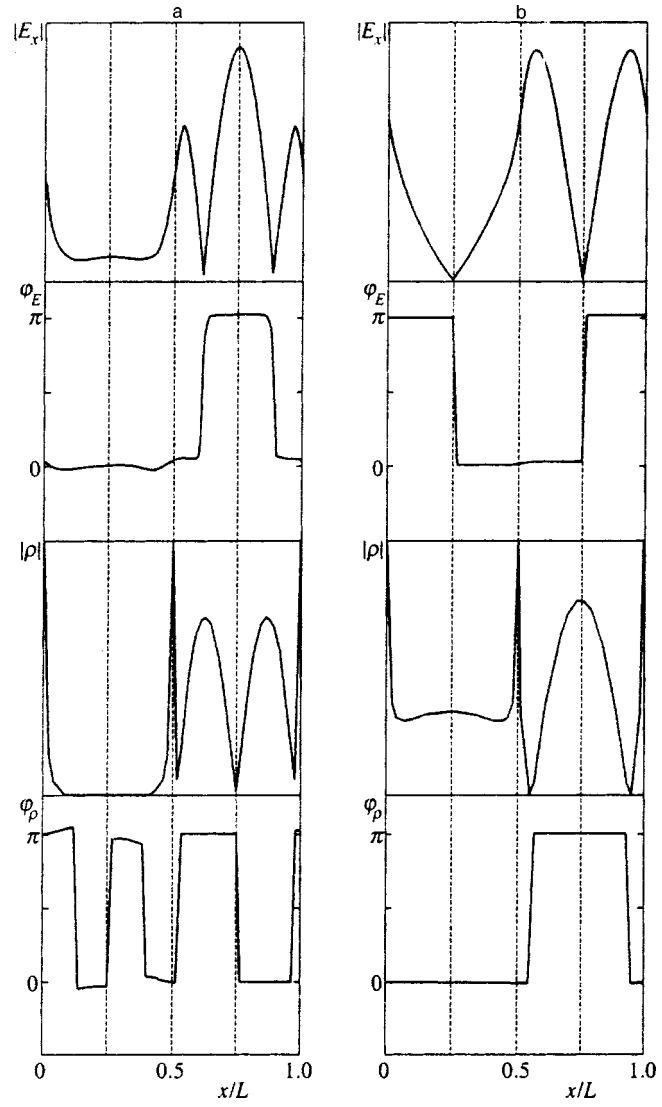

FIG. 2. Distributions of the amplitude and phase of the longitudinal electric field $E_x = |E_x| \exp(i\varphi_E)$ and of the oscillations of the surface charge density $\rho = |\rho| \exp(i\varphi_\rho)$ in the plane of the $2D$ electron system for radiative (a) and nonradiative (b) oscillations at $w/L = 0.5$ and $\Delta n_s = 0.7$. The other parameters of the structure are the same as in Fig. 1.

become degenerate ($w^+ = w^-$) approaches 0.5 as the modulation percentage decreases, which agrees with the perturbation theory formula (2) if

$$\frac{|N_2|}{N_0} = \frac{\sin(2\pi w/L)}{2\pi(N_B/(N_A - N_B) + w/L)}. \tag{24}$$

This corresponds to the given profile (3) of the concentration distribution. Estimates show that perturbation theory techniques provide satisfactory agreement (to within 1%) with the results of rigorous calculations of $\Delta\omega$ only when $\Delta n_s < 0.05$.

The asymmetry of the periodic profile of the electron concentration distribution in a $2D$ system (which was ignored in this work) must lead (at least in principle) to radiative damping of both fundamental plasma oscillations, which makes it possible to observe the splitting of the frequencies of these oscillations in experiments. The splitting of a plasma resonance due to the excitation of plasma oscillations at the frequencies ω^+ and ω^- has been observed in the experiments discussed in Refs. 1–3, which indicates that the symmetry of the profile of the concentration distribution is violated (in the work of Mackens *et al.*[1,2] this violation was uncontrolable). Moreover, the insulating layer (SiO_2) used in the experiments discussed in Refs. 1–3 had a periodically varying thickness, a situation that was also not taken into account by the theory.

These facts impede direct quantitative comparison of the experimental data and the results of calculations. For instance, the estimates done by the perturbation theory formula (2) with the use of (24) yield a value of $\Delta\omega$ that is almost twice as small as the experimentally observed value.[1] Our rigorous calculations yield even a smaller value of $\Delta\omega$ (by approximately 15%) for the percentage of modulation of the equilibrium electron concentration realized in the experiment of Mackens *et al.*:[1] $\Delta n_s \simeq 0.26$. Allowance for dissipative losses by inclusion into the calculations of the values of τ and ε_j'' characteristic of the experiment discussed in Ref. 1 has essentially no effect on the values of the natural frequencies (with an accuracy of better than 1%). We believe that the main reason for the considerable discrepancy between the experimental and theoretical data is the presence in the ex-

perimental structures of an additional inhomogeneity related to the periodic corrugation of the surface of the insulator layer. Note that Heitmann[3] has reported the observation of substantial splitting of the plasma resonance at $w/L=0.5$ and a modulation percentage Δn_s of approximately 0.7, which qualitatively agrees with the results of our calculation depicted in Fig. 1a. Perturbation theory techniques yield in this case $\Delta\omega=0$, so that they cannot be used to explain Heitmann's results[3] qualitatively.

As expected, the maximum value of the radiative damping coefficient increases with the modulation percentage, but in the process the maximum of γ_r shifts to larger values of w/L (Fig. 1b). This yields a nonmonotonic dependence of γ_r on the modulation percentage for any fixed value of w/L.

Figure 2a shows that the distribution of the electric field component E_x in a radiative plasma oscillation is symmetric with respect to the centers of the intervals $0<x<w$ and $w<x<L$. At the same time, the distribution $E_x(x)$ for the nonradiative mode (Fig. 2b) has nodes at the central points of these intervals, i.e., is antisymmetric. The distributions of the nonequilibrium additions to the surface charge density $\rho(x)$ for the radiative and nonradiative oscillations have opposite symmetry parities in relation to the corresponding distributions $E_x(x)$. Thus, in the case of a radiative oscillation in the plane of a $2D$ electron system there emerges a chain of electric dipoles formed by the nonequilibrium charges of opposite signs at the edges of the bands of $2D$ plasma with different equilibrium electron concentrations. The electromagnetic radiation emitted by these dipoles is the cause of radiative damping. In the case of nonradiative oscillations no dipoles are formed (see the distribution $\rho(x)$ in Fig. 2b). If the distribution profile for the equilibrium density of the $2D$ electron plasma is asymmetric, then the symmetry of the distributions $E_x(x)$ and $\rho(x)$ is violated, which leads to radiative damping of both oscillations.

4. CONCLUSION

We have analyzed theoretically the plasma oscillations in a periodically inhomogeneous $2D$ electron plasma. The model employed was that of a continuous $2D$ electron system with a rectangular profile of the spatial modulation of the equilibrium electron concentration with an arbitrary modulation percentage.

Using a rigorous electrodynamic approach, we have calculated the frequencies and radiative damping of two fundamental (lowest in frequency) types of plasma oscillations with a zero reduced wave vector. The splitting of the plasma oscillation frequencies has proved to be a nonmonotonic function of the ratio s/w of the bands of $2D$ plasma with low and high electron concentrations, respectively. At a certain value of s/w inside the interval $0<s/w<1$ degeneracy sets in and the oscillation frequencies are inverted. As the modulation percentage grows, degeneracy sets in at smaller values of the parameter s/w.

The results of our rigorous calculations have been compared with the experimental data taken from the literature on submillimeter Fourier transform spectroscopy of plasma oscillations in a periodically inhomogeneous $2D$ electron plasma and with the results of an earlier approximate theory of a small periodic perturbation of the equilibrium electron concentration. We have found that perturbation theory techniques yield results that are in satisfactory agreement with those of rigorous calculations only at values of the amplitude Δn_s of concentration modulation smaller than 0.05. The most probable reason for the large discrepancy between the experimental and theoretical values of the frequency splitting is the presence of a periodic gate electrode in the experimental structures, with the effect of this electrode ignored by the theory.

Because of the symmetry of the periodic profile of the distribution equilibrium electron concentration considered in this paper, one of the fundamental plasma oscillations is nonradiative. The maximum value of the radiative damping coefficient γ_r of the other fundamental oscillation increases with modulation percentage, while the maximum of γ_r is observed for smaller values of s/w. As a result, there emerges a nonmonotonic dependence of the radiative damping coefficient on the amplitude of the modulation of the equilibrium electron concentration at any fixed value of s/w.

On the basis of an analysis of the distributions of the longitudinal electric field in the radiative and nonradiative plasma oscillations and the distributions of the nonequilibrium surface charge density in a $2D$ electron plasma we have attempted to explain the physical mechanism of emission of electromagnetic waves by the system.

This work was supported by the Russian Fund for Fundamental Research (Project No. 96-02-19211).

*)E-mail: vapr@scnit.saratov.su

[1] U. Mackens, D. Heitmann, L. Prager, J. P. Kotthaus, and W. Beinvogl, Phys. Rev. Lett. **53**, 1485 (1984).
[2] D. Heitmann, J. P. Kotthaus, U. Mackens, and U. Beinvogl, Superlattices Microstruct. **1**, 35 (1985).
[3] D. Heitmann, Surf. Sci. **170**, 332 (1986).
[4] T. Zettler and J. P. Kotthaus, Semicond. Sci. Technol. **3**, 413 (1988).
[5] J. P. Kotthaus, W. Hansen, H. Pohlmann, and M. Wassermeier, Surf. Sci. **196**, 600 (1988).
[6] T. Demel, D. Heitmann, and P. Grambow, in *Proc. NATO ARW ''Spectroscopy of Semiconductor Microstructures''*, Venice (1989), NATO ASI Series B, Physics, Vol. **206**, G. Fasol, A. Fasolino, and P. Lugly (eds.), Plenum Press, New York (1989), p. 75.
[7] R. J. Wilkinson, C. D. Ager, T. Duffield, H. P. Hughes, D. G. Hasko, H. Ahmed, J. E. F. Frost, D. C. Peacock, D. A. Ritchie, G. A. C. Jones, C. R. Whitehouse, and N. Apsley, J. Appl. Phys. **71**, 6049 (1992).
[8] T. N. Theis, Surf. Sci. **98**, 515 (1980).
[9] A. V. Chaplik, Surf. Sci. Rep. **5**, 289 (1985).
[10] M. V. Krasheninnikov and A. V. Chaplik, Fiz. Tekh. Poluprovodn. **15**, 32 (1981) [Sov. Phys. Semicond. **15**, 19 (1982)].
[11] G. Eliasson, P. Hawrylak, Ji-Wei Wu, and J. J. Quinn, Solid State Commun. **60**, 3 (1986).
[12] V. Cataudella and V. M. Ramaglia, Phys. Rev. B **38**, 1838 (1988).
[13] S. V. Meshkov, J. Phys.: Condens. Matter **3**, 1773 (1991).
[14] U. Wulf, E. Zeeb, P. Gies, R. R. Gerhardts, and W. Hanke, Phys. Rev. B **42**, 7637 (1990).
[15] V. B. Shikin, T. Demel', and D. Heitmann, Zh. Éksp. Teor. Fiz. **96**, 1406 (1989) [Sov. Phys. JETP **69**, 797 (1989)].
[16] W. L. Schaich, P. W. Park, and A. H. MacDonald, Phys. Rev. B **46**, 12 643 (1992).
[17] M. V. Krasheninnikov and A. V. Chaplik, Zh. Éksp. Teor. Fiz. **88**, 129 (1985) [Sov. Phys. JETP **61**, 75 (1985)].
[18] C. D. Ager and H. P. Hughes, Phys. Rev. B **44**, 13 452 (1991).
[19] O. R. Matov, O. V. Polishchuk, and V. V. Popov, Pis'ma Zh. Tekh. Fiz.

18, No. 8, 86 (1992) [Sov. Tech. Phys. Lett. **18**, 645 (1992)].
[20] C. D. Ager, R. J. Wilkinson, and H. P. Hughes, J. Appl. Phys. **71**, 1322 (1992).
[21] O. R. Matov, O. V. Polischuk, and V. V. Popov, Int. J. Infrared Millim. Waves **14**, 1455 (1993).
[22] O. R. Matov, O. F. Meshkov, O. V. Polishchuk, and V. V. Popov, Zh. Éksp. Teor. Fiz. **109**, 876 (1996) [JETP **82**, 471 (1996)].
[23] G. Korn and T. Korn, *Mathematical Handbook*, 2nd ed., McGraw-Hill, New York (1968).
[24] V. Ya. Arsenin, *Methods of Mathematical Physics and Special Functions* [in Russian], Nauka, Moscow (1974), p. 346

Translated by Eugene Yankovsky

Absence of saturated ferromagnetism in the two-dimensional Hubbard model with two holes for $U=\infty$

Yu. V. Mikhaĭlova

State Scientific Center ''Scientific-Research Institute 'Heat Device,' '' 129085 Moscow, Russia
(Submitted 9 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 1000–1008 (March 1998)

For a square Hubbard lattice with infinite repulsion energy U the following exact result has been obtained: the ferromagnetic state with maximum spin is not the ground state of the system if the number of holes is equal to two. © *1998 American Institute of Physics.* [S1063-7761(98)01703-X]

1. INTRODUCTION

The Hubbard model, first introduced to explain ferromagnetism, is a simple model describing compounds with strong correlation coupling. The Hubbard Hamiltonian is usually written in the form

$$\hat{H}=-t\sum_{\langle i,j\rangle,\sigma}(\hat{c}^{+}_{i,\sigma}\hat{c}_{j,\sigma}+\hat{c}^{+}_{j,\sigma}\hat{c}_{i,\sigma})+U\sum_{i}(n_{i\uparrow}n_{i\downarrow}).$$

where $c^{+}_{i,\sigma}$ and $c_{i,\sigma}$ are the creation and annihilation operators of the electron at the site i with spin projection σ, $n_{i\sigma}$ is the number of electrons at the site i with spin σ, and the sum over $\langle i,j\rangle$ extends over pairs of nearest neighbors. The presence of only two parameters—the jump to the neighboring site t and the Coulomb repulsion energy U per site—make the model extraordinarily attractive for study.

The two-dimensional Hubbard model with infinite repulsion may be considered as a zeroth approximation for the description of a wide class of compounds with anomalous magnetic and electrical properties, including high-temperature superconductors. Therefore the question of the nature of the ground state of this model is of exceptional importance for understanding the mechanism of high-temperature superconductivity. The available references in the literature contain contradictory statements. All numerical studies of cluster systems (see, e.g., Refs. 1 and 2) give one and the same picture: for systems with fixed total spin of the particles S the ground-state energy $E_0(S)$ is a monotonic function of S. In the case of one hole the ground state of the system corresponds to maximum spin (saturated ferromagnetism) in the following cases: 1) a free boundary, 2) an even number of particles in each direction, and 3) positive jump energy ($t>0$). Here the ground-state energy is a monotonically decreasing function of S. If none of these three conditions is satisfied, the ground state corresponds to the minimum possible spin of the particles, whereby $E_0(S)$ grows with S. If the number of holes is greater than one, the ground state of the system corresponds to minimum spin ($S=0$ or $S=1/2$) and the ground-state energy is a monotonically increasing function of S. In his 1966 paper Nagaoka[3] considered only a simple cubic (square) lattice with periodic boundary conditions and an even number of particles in each direction. He gave a rigorous proof of the maximality of the spin for the ground-state energy in the case of one hole. The remainder of his paper is usually interpreted as follows: at low hole concentrations for a cubic lattice the ground state has maximum lattice spin (the state of saturated ferromagnetism) for all $U<U_{\max}$. According to Ref. 3, the limiting value $U_{\max}$ falls as the hole concentration grows. In fact Nagaoka formulated his result somewhat differently: for a simple cubic lattice with one hole and $U=\infty$ the ground state corresponds to maximum spin; for finite U and n holes the ferromagnetic state with maximum total spin is not the ground state if

$$\alpha n/N<t/U,$$

where N is the number of lattice sites, and α is a numerical parameter of order unity. Nagaoka obtained this result in the gas approximation assuming a small but macroscopic hole concentration n/N. Formally the case of two holes was not considered in Ref. 3.

In the present paper, for a periodic two-dimensional lattice with an even number of sites in each direction, I obtain an upper estimate for the difference

$$\Delta=E_0(S_{\max}-1)-E_0(S_{\max})$$

between the ground-state energy with spin equal to 1 less than its maximum value and the energy of the Nagaoka state. The estimate is obtained variationally. I consider the quantity

$$\tilde{\Delta}=\frac{\langle\Psi(\hat{H}-E_0(S_{\max}))\Psi\rangle}{\langle\Psi,\Psi\rangle}.$$

The explicit form of the test function is given for which $\tilde{\Delta}<0$.

2. MAIN EQUATIONS

Consider a rectangular $N_x\times N_y$ lattice. We denote by $N=N_xN_y$ the number of sites, by $\hat{a}^{+}_{i}$ ($\hat{a}_i$) the creation (annihilation) operator at the ith site of a particle with spin up, and by $\hat{b}^{+}_{i}$ ($\hat{b}_i$) the creation (annihilation) operator at the ith site of a particle with spin down. We assume that the system possesses translational invariance, and we consider states with prescribed quasimomentum $\alpha=(\alpha_x,\alpha_y)$:

$$\alpha_x=\frac{2\pi}{N_x}\,i,\quad i=0,1,...,N_x-1,$$

$$\alpha_y=\frac{2\pi}{N_y}\,i,\quad i=0,1,...,N_y-1.$$

For such states the complete set

$$\Phi_{ij}=\frac{1}{\sqrt{N}}\,(1+\exp(i\alpha_x)\hat{K}_x+\exp(2i\alpha_x)\hat{K}_x^2+\ldots+\exp(iN_{x-1}\alpha_x)\hat{K}_x^{N_x-1})(1+\exp(i\alpha_y)\hat{K}_y+\exp(2i\alpha_y)\hat{K}_y^2+\ldots+\exp(iN_{y-1}\alpha_y)\hat{K}_y^{N_y-1})\hat{a}_i\hat{a}_j\Phi_0, \tag{1}$$

may be specified, where $\hat{K}_x$ ($\hat{K}_y$) is the translational shift operator (by one lattice site) in the x (y) direction, and $\Phi_0=\hat{b}_1^+\hat{a}_2^+\ldots\hat{a}_N^+|\rangle$, where $|\rangle$ is the empty state. The function Φ_{ij} is a translationally invariant state with fixed distances between the flipped spin and each hole (equal to the distances between the first and ith or jth site, respectively). The translational shift operators $\hat{K}_x$ ($\hat{K}_y$) are defined as follows:

$$\hat{K}_x\Phi=\exp(-i\alpha_x)\Phi,\quad \hat{K}_y\Phi=\exp(-i\alpha_y)\Phi.$$

The functions Φ_{ij} satisfy the following normalization conditions:

$$\langle\Phi_{ij},\Phi_{lm}\rangle=\delta_{il}\delta_{jm}-\delta_{im}\delta_{jl}. \tag{2}$$

As the basis functions we take

$$\Omega_{\mathbf{k}_1,\mathbf{k}_2}=\frac{1}{M}\sum_{i,j}\Phi_{ij}\exp(i\mathbf{k}_1\cdot\mathbf{r}_i)\exp(i\mathbf{k}_2\cdot\mathbf{r}_j). \tag{3}$$

where the set $\mathbf{k}=(k_x,k_y)$ coincides with the set of quasimomenta $\boldsymbol{\alpha}=(\alpha_x,\alpha_y)$ with the exception of the case $k_x=k_y=0$. The vector $\mathbf{k}$ may be treated as a hole momentum in a system in which the flipped spin is at rest.

Obviously, the functions $\Omega_{\mathbf{k}_1,\mathbf{k}_2}$ are antisymmetric in the indices $\mathbf{k}_1,\mathbf{k}_2$:

$$\Omega_{\mathbf{k}_1,\mathbf{k}_2}=-\Omega_{\mathbf{k}_2,\mathbf{k}_1}. \tag{4}$$

Thus, the obvious requirement is fulfilled that for a given quasimomentum $\boldsymbol{\alpha}$ we have $(N-1)(N-2)/2$ independent functions $\Omega_{\mathbf{k}_1,\mathbf{k}_2}$:

We note the useful relation

$$\sum_{\mathbf{k}_1}\Omega_{\mathbf{k}_1\mathbf{k}_2}=-\frac{1}{N}\sum_{i,j}\Phi_{ij}\exp(i\mathbf{k}_2\cdot\mathbf{r}_j) \tag{5}$$

and analogously

$$\sum_{\mathbf{k}_2}\Omega_{\mathbf{k}_1\mathbf{k}_2}=-\frac{1}{N}\sum_{i,j}\Phi_{ij}\exp(i\mathbf{k}_1\cdot\mathbf{r}_j), \tag{5a}$$

where the sum over $\mathbf{k}$ is taken over the $N-1$ indicated values.

The functions Ω_{ij} satisfy the following normalization conditions:

$$\langle\Omega_{\mathbf{k}_1,\mathbf{k}_2},\Omega_{\mathbf{p}_1,\mathbf{p}_2}\rangle=(\delta_{\mathbf{k}_1,\mathbf{p}_1}\delta_{\mathbf{k}_2,\mathbf{p}_2}-\delta_{\mathbf{k}_1,\mathbf{p}_2}\delta_{\mathbf{k}_2,\mathbf{p}_1})-\frac{1}{N}\times(\delta_{\mathbf{k}_1,\mathbf{p}_1}+\delta_{\mathbf{k}_2,\mathbf{p}_2}-\delta_{\mathbf{k}_1,\mathbf{p}_2}-\delta_{\mathbf{k}_2,\mathbf{p}_1}). \tag{6}$$

In addition,

$$\langle\Omega_{\mathbf{k}_1,\mathbf{k}_2},\Phi_{\mathbf{r}_1,\mathbf{r}_2}\rangle=\frac{1}{N}\,(\exp(-i\mathbf{k}_1\cdot\mathbf{r}_1-i\mathbf{k}_2\cdot\mathbf{r}_2)-\exp(-i\mathbf{k}_1\cdot\mathbf{r}_2-i\mathbf{k}_2\cdot\mathbf{r}_1)). \tag{7}$$

Thus the functions Φ_{ij} can be expressed in terms of the functions $\Omega_{\mathbf{k}_1,\mathbf{k}_2}$ as follows:

$$\Phi_{ij}=\frac{1}{N}\sum_{\mathbf{k}_1\mathbf{k}_2}\Omega_{\mathbf{k}_1\mathbf{k}_2}(\exp(-i\mathbf{k}_1\cdot\mathbf{r}_i)-1)\times(\exp(-i\mathbf{k}_2\cdot\mathbf{r}_j)-1). \tag{8}$$

The energy spectrum E is found by solving the Schrödinger equation

$$\hat{H}\Psi=E\Psi. \tag{9}$$

where $\hat{H}$ is the Hubbard Hamiltonian for $U=\infty$:

$$\hat{H}=t\sum_{i,j}{}'\,[(\hat{a}_i^+\hat{a}_j+\hat{a}_j^+\hat{a}_i)(1-\hat{b}_i^+\hat{b}_i)(1-\hat{b}_j^+\hat{b}_j)+(\hat{b}_i^+\hat{b}_j+\hat{b}_j^+\hat{b}_i)(1-\hat{a}_i^+\hat{a}_i)(1-\hat{a}_j^+\hat{a}_j)] \tag{10}$$

or

$$\hat{H}=t\sum_{i,j,\sigma}\hat{X}_i^{\sigma 0}\hat{X}_j^{0\sigma}, \tag{10a}$$

where $\hat{X}_i^{\sigma 0}$ ($\hat{X}_i^{0\sigma}$) are the Hubbard operators. The sum in expressions (10) and (10a) extend over nearest neighbors. In what follows we taken the quantity t as our unit of energy, i.e., we set $t=1$ in Eq. (10).

We represent the wave function Ψ as an expansion over the set $\Omega_{\mathbf{k}_1,\mathbf{k}_2}$:

$$\Psi=\sum_{\mathbf{k}_1,\mathbf{k}_2}c_{\mathbf{k}_1,\mathbf{k}_2}\Omega_{\mathbf{k}_1,\mathbf{k}_2}. \tag{11}$$

3. VARIATIONAL ESTIMATES

Using test wave functions in the form (11) makes it easier to estimate the maximum eigenvalue of the Hamiltonian (10).

By virtue of the periodicity of the lattice and the evenness of the number of sites in each direction, under the substitution $\hat{a}_i^+\leftrightarrow\hat{a}_i^+(-1)^i$, $\hat{a}_i\leftrightarrow\hat{a}_i(-1)^i$ and correspondingly $\hat{b}_i^+\leftrightarrow\hat{b}_i^+(-1)^i$, $\hat{b}_i\leftrightarrow\hat{b}_i(-1)^i$, the Hamiltonian changes sign: $\hat{H}\leftrightarrow-\hat{H}$. This means that the energy spectrum E does not depend on the sign of t. Therefore, from the proof that the maximum eigenvalue $E_{\max}$ is greater than some value Λ ($E_{\max}>\Lambda$) it automatically follows that $E_{\min}<-\Lambda$.

Let us consider test functions for which the values $c_{\mathbf{k}_1,\mathbf{k}_2}$ are nonzero only if one of the vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ is equal to $\mathbf{p}_0=(p_x,p_y)=(\pi,\pi)$ and the other is equal to any of the four possible vectors:

$$\mathbf{p}_1=\left(\pi,\pi+\frac{2\pi}{L}\right),\quad \mathbf{p}_2=\left(\pi,\pi-\frac{2\pi}{L}\right),$$
$$\mathbf{p}_3=\left(\pi-\frac{2\pi}{L},\pi\right),\quad \mathbf{p}_4=\left(\pi+\frac{2\pi}{L},\pi\right), \tag{12}$$

where $L=\sqrt{N}$.

The vectors $\mathbf{p}_i$ were chosen such that for all $\mathbf{p}_i$ the energy of the two free quasiparticles is equal to the maximum energy of the states of the $N-2$ states with spin $S=S_{\max}$:

$$\varepsilon_0=\varepsilon_{\mathbf{p}_0,\mathbf{p}_i}=8\left(1+\frac{\cos(2\pi/L)-1}{4}\right)\approx 8\left(1-\frac{\pi^2}{2N}\right). \tag{13}$$

We assume the test wave function to be symmetric with respect to the substitution $y\leftrightarrow -y$ under the condition $c_{\mathbf{p}_4,\mathbf{p}_0}=0$; therefore the nonzero coefficients are $c_{\mathbf{k}_1,\mathbf{k}_2}$

$$c_{\mathbf{p}_1,\mathbf{p}_0}=\gamma,\quad c_{\mathbf{p}_2,\mathbf{p}_0}=\gamma,\quad c_{\mathbf{p}_3,\mathbf{p}_0}=2\delta,\quad c_{\mathbf{p}_4,\mathbf{p}_0}=0,$$
$$c_{\mathbf{p}_0,\mathbf{p}_1}=-\gamma,\quad c_{\mathbf{p}_0,\mathbf{p}_2}=-\gamma,\quad c_{\mathbf{p}_0,\mathbf{p}_3}=-2\delta,\quad c_{\mathbf{p}_0,\mathbf{p}_4}=0. \tag{14}$$

We will treat the coefficients γ and δ as variational parameters.

The value $\tilde{\Delta}=\langle\Psi,(\hat{H}-E_0(S_{\max}))\Psi\rangle/\langle\Psi,\Psi\rangle$, according to Eq. (A16) in the Appendix, is calculated in terms of the quantities $f_i(\mathbf{p})$. Simple calculation gives

$$f_1(\mathbf{p}_1)=c_{\mathbf{p}_1,\mathbf{p}_0}=\gamma,\quad f_1(\mathbf{p}_2)=c_{\mathbf{p}_2,\mathbf{p}_0}=\gamma,$$
$$f_1(\mathbf{p}_3)=c_{\mathbf{p}_3,\mathbf{p}_0}=2\delta,\quad f_1(\mathbf{p}_4)=c_{\mathbf{p}_4,\mathbf{p}_0}=0,$$
$$f_1(\mathbf{p}_0)=-2\gamma-2\delta,\quad f_2(\mathbf{p}_1)=c_{\mathbf{p}_1,\mathbf{p}_0}\cos p_{0x}=-\gamma,$$
$$f_2(\mathbf{p}_2)=-\gamma,\quad f_2(\mathbf{p}_3)=-2\delta,$$
$$f_2(\mathbf{p}_4)=0,\quad f_2(\mathbf{p}_0)=2(\gamma+\delta)\left(1-\delta\,\frac{1-\cos(2\pi/L)}{\gamma+\delta}\right),$$
$$f_3(\mathbf{p}_1)=f_3(\mathbf{p}_2)=f_3(\mathbf{p}_3)=f_3(\mathbf{p}_4)=0,$$
$$f_3(\mathbf{p}_0)=2\delta\sin(2\pi/L),$$
$$f_4(\mathbf{p}_1)=c_{\mathbf{p}_1,\mathbf{p}_0}\cos p_{0y}=-\gamma,\quad f_4(\mathbf{p}_2)=-\gamma,$$
$$f_4(\mathbf{p}_3)=-2\delta,\quad f_4(\mathbf{p}_4)=0,$$
$$f_4(\mathbf{p}_0)=2(\gamma+\delta)\left(1-\gamma\,\frac{1-\gamma\cos(2\pi/L)}{\gamma+\delta}\right),$$
$$f_5(\mathbf{p}_0)=0,\quad f_5(\mathbf{p}_1)=f_5(\mathbf{p}_2)=f_5(\mathbf{p}_3)=f_5(\mathbf{p}_4)=0, \tag{15}$$

Substituting Eqs. (15) into Eq. (A9), we obtain

$$|\Psi|^2=2\sum|c_{\mathbf{k}_1\mathbf{k}_2}|^2-\frac{8}{N}(|\gamma|^2+|\delta|^2+2|\gamma+\delta|^2). \tag{16}$$

Let us calculate the value of $\Delta=X-\varepsilon_0|\Psi|^2$. Employing Eqs. (A9) and (A10), we find

$$\Delta=\frac{8\varepsilon_0}{N}(|\gamma|^2+|\delta|^2+2|\gamma+\delta|^2)+\frac{8}{N}\sum_{\mathbf{k}_1\mathbf{k}_2}(\cos k_{1x}+\cos k_{1y})|f_1(\mathbf{k}_1)|^2+\frac{8}{N}\sum_{\mathbf{k}_1}\{f_1(\mathbf{k}_1)[f_2^*(\mathbf{k}_1)+f_4^*(\mathbf{k}_1)]+f_1^*(\mathbf{k}_1)[f_2(\mathbf{k}_1)+f_4(\mathbf{k}_1)]\}$$
$$-\frac{8}{N}\sum_{\mathbf{k}_1}|f_2(\mathbf{k}_1)|^2\cos(k_{1x}-\alpha_x)+\frac{8}{N}\sum_{\mathbf{k}_1}|f_3(\mathbf{k}_1)|^2\cos(k_{1x}-\alpha_x)+\frac{8}{N}\sum_{\mathbf{k}_1}[f_2^*(\mathbf{k}_1)f_3(\mathbf{k}_1)+f_3^*(\mathbf{k}_1)f_2(\mathbf{k}_1)]\sin(k_{1x}-\alpha_x)$$
$$-\frac{8}{N}\sum_{\mathbf{k}_1}|f_4(\mathbf{k}_1)|^2\cos(k_{1y}-\alpha_y)+\frac{8}{N}\sum_{\mathbf{k}_1}[f_4^*(\mathbf{k}_1)f_5(\mathbf{k}_1)+f_5^*(\mathbf{k}_1)f_4(\mathbf{k}_1)]\sin(k_{1y}-\alpha_y)+\frac{8}{N}\sum_{\mathbf{k}_1}|f_5(\mathbf{k}_1)|^2\cos(k_{1y}-\alpha_y). \tag{17}$$

Calculating the sums over the vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ in expression (17), we find ($\alpha=2\pi/L$)

$$r_1=\frac{8}{N}\sum_{\mathbf{k}_1\mathbf{k}_2}(\cos k_{1x}+\cos k_{1y})|f_1(\mathbf{k}_1)|^2$$
$$=-\frac{32}{N}\left(1-\frac{1-\cos(2\pi/L)}{2}\right)(|\gamma|^2+|\delta|^2)-\frac{64}{N}|\gamma+\delta|^2, \tag{18}$$

$$r_2=\frac{16}{N}\sum_{\mathbf{k}_1}f_1(\mathbf{k}_1)[f_2^*(\mathbf{k}_1)+f_4^*(\mathbf{k}_1)]$$
$$=-\frac{64}{N}(|\gamma|^2+|\delta|^2)-\frac{128}{N}|\gamma+\delta|^2\left(1-\frac{1-\cos(2\pi/L)}{2}\right), \tag{19}$$

$$r_3=-\frac{8}{N}\sum_{\mathbf{k}_1}|f_2(\mathbf{k}_1)|^2[\cos(k_{1x}-\alpha_x)+\cos(k_{1y}-\alpha_y)]$$
$$=\frac{32}{N}(|\gamma|^2+|\delta|^2)+\frac{64}{N}|\gamma+\delta|^2-\frac{8}{N}|\gamma|^2\left[2-\cos\left(\frac{2\pi}{L}-\alpha_y\right)+2-2\cos\alpha_x-\cos\left(\frac{2\pi}{L}+\alpha_x\right)\right]$$
$$-\frac{8}{N}|\delta|^2\left[2-\cos\alpha_y-\cos\left(\frac{2\pi}{L}+\alpha_x\right)\right]+\frac{32}{N}|\gamma+\delta|^2(\cos\alpha_x+\cos\alpha_y-2), \tag{20}$$

$$r_4=\frac{8}{N}\sum_{\mathbf{k}_1}[f_2^*(\mathbf{k}_1)f_3(\mathbf{k}_1)+f_3^*(\mathbf{k}_1)f_2(\mathbf{k}_1)]\times[\sin(k_{1x}-\alpha_x)+\sin(k_{1y}-\alpha_y)]$$
$$=\frac{32}{N}(\gamma+\delta)\delta\left(1-\delta\,\frac{1-\cos(2\pi/L)}{\gamma+\delta}\right)$$

$$\times[\sin\alpha_x+\sin\alpha_y]\sin\frac{2\pi}{L}, \tag{21}$$

$$r_5=\frac{8}{N}\sum_{\mathbf{k}_1}|f_3(\mathbf{k}_1)|^2\cos(k_{1x}-\alpha_x)$$
$$+\frac{8}{N}\sum_{\mathbf{k}_1}|f_5(\mathbf{k}_1)|^2\cos(k_{1y}-\alpha_y)$$
$$=-\frac{16}{N}(\gamma+\delta)\delta\sin^2\frac{2\pi}{L}\cos\alpha_x. \tag{22}$$

Finally, we obtain

$$\Delta=\frac{64}{N}(\gamma+\delta)\delta\alpha_x\frac{2\pi}{L}-\frac{16}{N}|\gamma+\delta|^2(\alpha_x^2+\alpha_y^2)$$
$$-\frac{4|\gamma|^2}{N}\left[\alpha_x^2+\left(\frac{2\pi}{L}-\alpha_x\right)^2+\left(\frac{2\pi}{L}-\alpha_y\right)^2\right]$$
$$-\frac{16|\delta|^2}{N}\left[\alpha_y^2+\alpha^2+\left(\frac{2\pi}{L}+\alpha_x\right)^2+\left(\frac{2\pi}{L}+\alpha_y\right)^2\right]. \tag{23}$$

Note that for $\gamma=-(2/15)\delta$, $\alpha_x=2\alpha$, and $\alpha_y=0$ the quantity Δ is equal to

$$\tilde{\Delta}=\frac{128\pi^2}{15N^2}|\delta|^2. \tag{24}$$

Therefore the maximum energy is greater than ε_0 by the amount

$$\Delta\varepsilon=\frac{128\pi^2|\alpha|^2}{15N^2\langle\Psi\Psi\rangle}\geqslant\frac{0.4\pi^2}{N^2}. \tag{25}$$

This means that there exists an energy level E^+ for a system with two holes such that

$$E^+\geqslant\varepsilon_0+\frac{0.4\pi^2}{N^2}. \tag{26}$$

As a consequence of the symmetry of the energy spectrum relative to change of sign of t, noted above, it follows from the inequality (26) that there also exists an energy level E^- for a system with two holes such that

$$E^-\leqslant-\varepsilon_0-\frac{0.4\pi^2}{N^2}. \tag{27}$$

4. CONCLUSION

It follows from inequality (27) that the ground state of a system with two holes corresponds to total spin S of the particles less than the maximum possible value. In the present paper, as our test functions we have chosen functions with $S=S_{\max}-1$. Even in this case the ground state is lower than the Nagaoka state. For states with lower S this estimate can possibly be strengthened.

The results of this work show that the state of saturated ferromagnetism is not the ground state for a system with two holes.

This conclusion in no way contradicts the results of Nagaoka.[3] In this regard, it must be stressed that Nagaoka proved that the ground state of a system with $U=\infty$ and one hole, which corresponds to saturated ferromagnetism, corresponds to a state with lower S for a larger number of holes and $U<U_0$. Thus, in Ref. 3 he proved the absence of saturated ferromagnetism in the indicated case. In his proof Nagaoka made an assumption about the structure of the wave function that is not necessarily valid if the ground state is degenerate for a system with maximum spin and prescribed value of the projection of the total spin.

Thus, Ref. 3 derives a sufficient ($U<U_0$) but not necessary condition for the absence of saturated ferromagnetism. In his 1966 paper Nagaoka sought a wave function of completely determinate form under the condition of nondegeneracy of the ground state for a system with maximum spin. In the case of two holes the ground state for a system with maximum spin is degenerate; therefore the ground-state wave function cannot coincide with the wave function found by Nagaoka. This applies to the three-dimensional as well as the two-dimensional case. The test function used in the present paper can be obtained as the solution of a secular equation for the zeroth-order wave function in the expansion in the density as the small parameter (in the given case $1/N$).

APPENDIX A

To calculate the quantity

$$\Delta=\frac{\langle\Psi,\hat{H}\Psi\rangle}{\langle\Psi,\Psi\rangle} \tag{A1}$$

we find an expression for the normalization $\langle\Psi,\Psi\rangle$ and $X=\langle\Psi,\hat{H}\Psi\rangle$ in terms of the coefficients $c_{\mathbf{k}_1,\mathbf{k}_2}$ in the wave function expansion (11).

Taking Eqs. (6) and (7) into account, we obtain for the normalization of the wave function $\langle\Psi,\Psi\rangle$

$$\langle\Psi,\Psi\rangle=2\sum_{\mathbf{k}_1\mathbf{k}_2}|c_{\mathbf{k}_1\mathbf{k}_2}|^2-\frac{4}{N}\sum_{\mathbf{k}_1}\left|\sum_{\mathbf{k}_p}c_{\mathbf{k}_1\mathbf{k}_p}\right|^2. \tag{A2}$$

Introducing the notation

$$\varepsilon(\mathbf{k})=-2(\cos k_x+\cos k_y), \tag{A3}$$

$$f_1(\mathbf{k})=\sum_{\mathbf{k}_1}c_{\mathbf{k}\mathbf{k}_1}, \tag{A4}$$

$$f_2(\mathbf{k})=\sum_{\mathbf{p}}c_{\mathbf{k}\mathbf{p}}\cos p_x, \tag{A5}$$

$$f_4(\mathbf{k})=\sum_{\mathbf{p}}c_{\mathbf{k}\mathbf{p}}\cos p_y, \tag{A6}$$

$$f_3(\mathbf{k}_1)=\sum_{\mathbf{k}_2}c_{\mathbf{k}_1\mathbf{k}_2}\sin k_{2x}, \tag{A7}$$

$$f_5(\mathbf{k}_1)=\sum_{\mathbf{k}_2}c_{\mathbf{k}_1\mathbf{k}_2}\sin k_{2y}. \tag{A8}$$

we have

$$\langle\Psi,\Psi\rangle=2\sum_{\mathbf{k}_1\mathbf{k}_2}|c_{\mathbf{k}_1\mathbf{k}_2}|^2-\frac{4}{N}\sum_{\mathbf{k}}|f_1(\mathbf{k}_1)|^2. \tag{A9}$$

Similarly, for $X=\langle\Psi,\hat{H}\Psi\rangle$ we obtain

$$X=2\sum_{\mathbf{k}_1\mathbf{k}_2}\varepsilon_{\mathbf{k}_1\mathbf{k}_2}|c_{\mathbf{k}_1\mathbf{k}_2}|^2+\frac{8}{N}\sum_{\mathbf{k}_1\mathbf{k}_2}(\cos k_{1x}+\cos k_{1y})|f_1(\mathbf{k}_1)|^2+\frac{8}{N}\sum_{\mathbf{k}_1}\{f_1(\mathbf{k}_1)[f_2^*(\mathbf{k}_1)+f_4^*(\mathbf{k}_1)]+f_1^*(\mathbf{k}_1)[f_2(\mathbf{k}_1)+f_4(\mathbf{k}_1)]\}-\frac{8}{N}\sum_{\mathbf{k}_1}|f_2(\mathbf{k}_1)|^2\cos(k_{1x}-\alpha_x)+\frac{8}{N}\sum_{\mathbf{k}_1}[f_2^*(\mathbf{k}_1)f_3(\mathbf{k}_1)+f_3^*(\mathbf{k}_1)f_2(\mathbf{k}_1)]\times\sin(k_{1x}-\alpha_x)+\frac{8}{N}\sum_{\mathbf{k}_1}|f_3(\mathbf{k}_1)|^2\cos(k_{1x}-\alpha_x)-\frac{8}{N}\sum_{\mathbf{k}_1}|f_4(\mathbf{k}_1)|^2\cos(k_{1y}-\alpha_y)+\frac{8}{N}\sum_{\mathbf{k}_1}|f_5(\mathbf{k}_1)|^2\cos(k_{1y}-\alpha_y)+\frac{8}{N}\sum_{\mathbf{k}_1}[f_4^*(\mathbf{k}_1)f_5(\mathbf{k}_1)+f_5(\mathbf{k}_1)f_4(\mathbf{k}_1)]\times\sin(k_{1y}-\alpha_y). \tag{A10}$$

We note the case in which only terms with hole momenta $\mathbf{p}_1$ and $\mathbf{p}_2$ are present in the expansion (11) of the wave function Ψ, i.e.,

$$c_{\mathbf{k}_1\mathbf{k}_2}=\delta_{\mathbf{k}_1\mathbf{p}_1}\delta_{\mathbf{k}_2\mathbf{p}_2}-\delta_{\mathbf{k}_1\mathbf{p}_2}\delta_{\mathbf{k}_2\mathbf{p}_1}. \tag{A11}$$

Here

$$f_1(\mathbf{p}_1)=1,\quad f_1(\mathbf{p}_2)=-1,\quad f_2(\mathbf{p}_1)=\cos p_{2x},$$
$$f_2(\mathbf{p}_2)=-\cos p_{1x},\quad f_4(\mathbf{p}_1)=\cos p_{2y},$$
$$f_4(\mathbf{p}_2)=-\cos p_{1y},\quad f_3(\mathbf{p}_1)=\sin p_{2x},$$
$$f_3(\mathbf{p}_2)=-\sin p_{2x},\quad f_5(\mathbf{p}_1)=\sin p_{2y},$$
$$f_5(\mathbf{p}_2)=-\sin p_{1y} \tag{A12}$$

and, consequently,

$$\langle\Psi,\Psi\rangle=4-\frac{8}{N}, \tag{A13}$$

$$X=-4\varepsilon_{p_1p_2}+\frac{8}{N}(\cos p_{1x}+\cos p_{1y}+\cos p_{2x}+\cos p_{2y})+\frac{16}{N}(\cos p_{1x}+\cos p_{2x}+\cos p_{1y}+\cos p_{2y})-\frac{8}{N}(\cos(2p_{2x}+p_{1x}-\alpha_x)+\cos(2p_{1x}+p_{2x}-\alpha_x))-\frac{8}{N}(\cos(2p_{2y}+p_{1y}-\alpha_y)+\cos(2p_{1y}+p_{2y}-\alpha_y)). \tag{A14}$$

If we have $p_{2x}+p_{1x}=\alpha_x$ and $p_{2y}+p_{1y}=\alpha_y$, which corresponds to the wave function of a state $S=S_{\max}$, then

$$X=4\left(1-\frac{2}{N}\right)\varepsilon_{\mathbf{p}_1\mathbf{p}_2}.$$

Taking Eq. (A13) into account, we have

$$X=-(\varepsilon_{\mathbf{p}_1}+\varepsilon_{\mathbf{p}_2})\langle\Psi,\Psi\rangle,$$

as could be expected, since in the given case we used the exact wave function with eigenvalue of the Hamiltonian $E=-(\varepsilon_{\mathbf{p}_1}+\varepsilon_{\mathbf{p}_2})$.

We also make note of the following convenient fact. Let the functions $c_{\mathbf{k}_1,\mathbf{k}_2}=c(k_{1x},k_{1y},k_{2x},k_{2y})$ be symmetric under to the substitution $x\leftrightarrow y$. Then

$$f_1(k_{1x},k_{1y})=\Lambda f_1(k_{1y},k_{1x}),$$
$$f_2(k_{1x},k_{1y})=\Lambda f_4(k_{1y},k_{1x}),$$
$$f_3(k_{1x},k_{1y})=\Lambda f_5(k_{1y},k_{1x}), \tag{A15}$$

where $\Lambda=\pm1$.

Substituting Eqs. (A15) into Eq. (A10), we obtain

$$X=2\sum_{\mathbf{k}_1\mathbf{k}_2}\varepsilon_{\mathbf{k}_1\mathbf{k}_2}|c_{\mathbf{k}_1\mathbf{k}_2}|^2+\frac{8}{N}\sum_{\mathbf{k}_1\mathbf{k}_2}(\cos k_{1x}+\cos k_{1y})|f_1(\mathbf{k}_1)|^2+\frac{16}{N}\sum_{\mathbf{k}_1}\{f_1(\mathbf{k}_1)f_2^*(\mathbf{k}_1)+f_1^*(\mathbf{k}_1)f_2(\mathbf{k}_1)\}-\frac{8}{N}\sum_{\mathbf{k}_1}|f_2(\mathbf{k}_1)|^2(\cos(k_{1x}-\alpha_x)+\cos(k_{1x}-\alpha_y))+\frac{8}{N}\sum_{\mathbf{k}_1}[f_2^*(\mathbf{k}_1)f_3(\mathbf{k}_1)+f_3^*(\mathbf{k}_1)f_2(\mathbf{k}_1)](\sin(k_{1x}-\alpha_x)+\sin(k_{1x}-\alpha_y))+\frac{8}{N}\sum_{\mathbf{k}_1}|f_3(\mathbf{k}_1)|^2(\cos(k_{1x}-\alpha_x)+\cos(k_{1x}-\alpha_y)). \tag{A16}$$

[1] Y. Takahashi, Progr. Theor. Phys. Jpn. **41**, 228 (1972).
[2] R. O. Zaĭtsev and Yu. V. Mikhaĭlova, Fiz. Nizk. Temp. **17**, 999 (1991) [Sov. J. Low Temp. Phys. **17**, 521 (1991)].
[3] Y. Nagaoka, Phys. Rev. **147**, 392 (1966).

Translated by Paul F. Schippnick

Electron spectrum and wave functions of icosahedral quasicrystals

D. V. Olenev,*) É. I. Isaev, and Yu. Kh. Vekilov

Moscow State Steel and Alloys Institute, 117936 Moscow, Russia
(Submitted 14 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 1009–1025 (March 1998)

Electron spectra and wave functions of icosahedral quasicrystals have been investigated in the tight-binding approximation using the two-fragment structural model (the Amman–MacKay network) with ''central'' decoration. A quasicrystal has been considered as a limiting structure in a set of optimal cubic approximants with increasing lattice constants. The method of level statistics indicates that the energy spectrum of an icosahedral quasicrystal contains a singular (nonsmooth) component. The density of electron states has been calculated for the first four optimal cubic approximants of the icosahedral quasicrystal, and the respective Lebesgue measures of energy spectra of these approximants have been obtained. Unlike the case of a one-dimensional quasiperiodic structure, the energy spectrum of an icosahedral quasicrystal does not contain a hierarchical gap structure typical of the Cantor set of measure zero in a one-dimensional quasicrystal. Localization of wave functions in an icosahedral quasicrystal has been studied, and their ''critical'' behavior has been detected. The effect of disorder due to substitutional impurities on electron properties of icosahedral quasicrystals has been investigated. This disorder makes the electron spectrum ''smoother'' and leads to a tendency to localization of wave functions. © *1998 American Institute of Physics.* [S1063-7761(98)01803-4]

1. INTRODUCTION

Quasicrystals are objects that have noncrystallographic symmetry and coordination long-range order (i.e., the capability of coherent scattering of incident radiation).[1,2] Quasicrystals are metallic alloys, but their physical properties are different from those of crystalline and amorphous metallic phases, although they have much in common with the latter. Like metals, quasicrystals have a specific heat component linear in temperature, which is usually smaller than the calculated value of the free-electron model.[3] At the same time, the electrical resistivity of quasicrystals (unlike that of amorphous phases of similar compositions) is anomalously high (in the stable icosahedral phase of AlPdRe it is as high as $2\ \Omega\cdot$cm at 0.5 K[4]), highly susceptible to their composition, decreases with increasing temperature (its temperature coefficient is negative at all temperatures), and increases as the degree of structural order increases and defects are annealed.[3,5]

Decagonal quasicrystals are characterized by anisotropic conductivity: the conductivity along the packing axis of crystal planes with a finite period (the ten-fold axis) behaves like that of normal metals, and the conductivity in the crystal planes themselves behaves in the manner described above.[6] In addition, the conductivity along the periodic axis of decagonal quasicrystals is an order of magnitude higher than the conductivity in crystal planes.[7] Other specific features are the large magnitude and strong temperature dependence of the Hall coefficient, conductivity in the optical band deviating from the Drude formula, and strong temperature dependence of the thermopower.[3,6] The magnetoresistance in weak magnetic fields is usually negative.[3] Essentially all quasicrystal phases, including alloys containing transitional metals, are weak diamagnetics over a wide temperature range (except for alloys containing Mn).[6] Various models have been suggested to describe these properties, but accurate information on the electron spectra of quasicrystals is required to assess the utility of these interpretations.

Electron spectra of quasicrystal structures have been studied theoretically rather well in the cases of one- and two-dimensional quasicrystals using the two-fragment structural model (one- and two-dimensional Penrose tilings, respectively).[8–11] The energy spectrum of a one-dimensional quasicrystal lattice is a Cantor set of Lebesgue measure zero with a self-similar sequence of gaps; its wave functions demonstrate critical behavior, being neither localized nor delocalized.[8,9]

In the two-dimensional configuration, the energy spectrum has nonzero Lebesgue measure, contains a singular component, and most of its wave functions are critical.[10,11] One consequence of the peculiar electron spectra of one- and two-dimensional quasicrystals is that their resistance is a power-law function of their dimensions.[10] An investigation of electron properties of three-dimensional icosahedral quasicrystals with Danzer's model structure (constructed on the basis of four tetrahedra) revealed large oscillations in the density of electron states throughout the spectral band, and the localization of individual states in specific topological configurations.[12]

The shapes of electron spectra and wave functions in three-dimensional quasicrystals with the structure of the Amman–MacKay network (a three-dimensional analogue of the Penrose tiling) have not been studied in detail since, starting with the publication by Marcus,[13] many researchers

believed that this system had no interesting features in its spectrum and wave functions.[12] In reality, published conclusions are inconsistent. For example, Marcus,[13] who studied the electronic properties of a three-dimensional quasicrystal with the structure of the Amman–MacKay network in the tight-binding approximation, detected no fundamental differences from the crystalline structure, and concluded that its electron spectrum was smooth and electron wave functions were delocalized.[13] Niizeki and Akamatsu,[14] in contrast, suggested the existence of critical wave functions and a peculiar shape of the electron spectrum—a singular continuous shape in the three-dimensional case. Similar conclusions can be derived from the results obtained earlier in Refs. 15 and 16.

First-principles calculations and calculations of the density of electron states in the weak-binding approximation for some crystalline approximants, i.e., three-dimensional crystalline structures whose local atomic structure is similar to that of quasicrystals, indicate the existence of strong oscillations in the density of states with a deep pseudogap at the Fermi level. This allowed the researchers to suggest that the Hume–Rothery criterion could be applied to quasicrystal stability.[3,7,17–25] The presence of a pseudogap, however, cannot account for the physical properties of quasicrystals described above. Their interpretation demands more detailed information about the wave functions and features of the electron spectrum in quasicrystals.

In the present work, we have studied in the tight-binding approximation features of the electron spectrum and localization of wave functions in a three-dimensional icosahedral quasicrystal treated as a structural limit of a sequence of cubic approximants. This publication is a continuation of our earlier studies.[15,16] The paper is organized as follows. Section 2 considers the most widely used two-fragment structural model of quasicrystals, describes the formalism of the projection method applied to the Amman–MacKay network, and considers the structure of periodic approximants of icosahedral quasicrystals. Section 3 is dedicated to the analysis of electron properties of icosahedral quasicrystals in the tight-binding approximation. Subsection 3.1 describes techniques used in this study and approximations. Calculations of the electron spectrum and wave functions in a defect-free icosahedral quasicrystal, along with the effect of disorder due to substitutional impurities in a quasicrystal, are discussed in Subsection 3.2. Section 4 is devoted to the discussion of results, and Sec. 5 contains conclusions.

2. PERIODIC APPROXIMANTS OF ICOSAHEDRAL QUASICRYSTALS

Probably the most common model of the quasicrystal frame is based on the quasiperiodic packing of two elementary structural units (the two-fragment model). This model applied to a one-dimensional quasicrystal generates the Fibonacci sequence of short and long segments. In the two-dimensional case the two-fragment model is the Penrose tiling composed of two rhombi with acute angles of 36° and 72° at their vertices (prolate and oblate rhombi, respectively), and in the three-dimensional case, this is a generalization of the Penrose tiling (alternatively termed the Amman–MacKay network) composed of prolate and oblate rhombohedrons. Real quasicrystalline structures are usually modeled by the Amman–MacKay network decorated in a complex manner, in which atoms are located at positions of different fragments forming the rhombohedron network (rhombohedral vertices, edges, faces, and inner volumes). In such structures some positions can be vacant (structural voids) and chemical disorder due to substitutional impurities is also possible.[25]

There are three basic techniques for construction of the Amman–MacKay network, namely the projection technique, the method of multiple grids, and the method using self-similarity of quasiperiodic structures.[10,11,26,27] The most commonly used method is the projection technique, which is performed as follows. In order to construct the Amman–MacKay network with a primitive decoration (atoms are located only at rhombohedron vertices), an integer hyperlattice is generated in a six-dimensional space, and three-dimensional subspaces, parallel ($\parallel$) and perpendicular ($\perp$), with an irrational orientation with respect to the hyperlattice are constructed. Lattice sites close to the parallel subspace are projected onto it, and this projection is the basis of the quasicrystal structure frame. The subspace orientations are selected so that the unit vectors of the six-dimensional hyperlattice projected onto the parallel subspace are aligned with the five-fold axes of the icosahedron. A site $\mathbf{r}'$ of the six-dimensional hyperlattice is present in the projection onto the parallel subspace if $n(\mathbf{r}'_\perp)=1$, where $n(\mathbf{r}_\perp)$ is a form-function of the hyperlattice unit cell projection on the orthogonal subspace (the form-function is unity within the projection and zero outside it). Here $\mathbf{r}_\perp$ is the perpendicular component of $\mathbf{r}$. The inside volume of the hypercell is defined by the condition

$$\mathbf{r}=\sum_{i=1}^{6} x_i\mathbf{a}_i, \quad 0<x_i<1,$$

where $\mathbf{a}_i$ are the basis vectors of the six-dimensional cubic lattice, so

$$n(\mathbf{r}_\perp)=\begin{cases} 1, & \text{if } \mathbf{r}_\perp=\sum_{i=1}^{6} x_i\mathbf{a}_{i\perp}, \quad 0<x_i<1, \\ 0, & \text{if otherwise.} \end{cases} \tag{1}$$

Following Shaw *et al.*,[28] we use the following set of basis vectors in the perpendicular and parallel subspaces:

$$\mathbf{a}_{1\perp}=\eta(1,0,-\tau), \quad \mathbf{a}_{2\perp}=\eta(1,0,\tau),$$
$$\mathbf{a}_{3\perp}=\eta(-\tau,1,0), \quad \mathbf{a}_{4\perp}=\eta(0,-\tau,1), \tag{2}$$
$$\mathbf{a}_{5\perp}=\eta(0,\tau,1), \quad \mathbf{a}_{6\perp}=\eta(-\tau,-1,0);$$
$$\mathbf{a}_{1\parallel}=\eta(\tau,0,1), \quad \mathbf{a}_{2\parallel}=\eta(\tau,0,-1),$$
$$\mathbf{a}_{3\parallel}=\eta(1,\tau,0), \quad \mathbf{a}_{4\parallel}=\eta(0,1,\tau), \tag{3}$$
$$\mathbf{a}_{5\parallel}=\eta(0,-1,\tau), \quad \mathbf{a}_{6\parallel}=\eta(1,-\tau,0),$$

where $\eta=(1+\tau^2)^{-1/2}$. The polyhedron defining the region of nonzero values of $n(\mathbf{r}_\perp)$ in Eq. (1) is an orthogonal cross section of the ''projection tube'' (a region in the six-dimensional space in which a vector of the six-dimensional

space is projected onto the parallel subspace) shaped like a rhombic triacontahedron. The irrational number τ characterizing the rhombic triacontahedron, $\tau=(1+\sqrt{5})/2$, is called the ''golden mean.''

Let us leave the set of basis vectors (3) unchanged and replace the set (2) with basis vectors $\mathbf{a}_{i\perp}$, $i=1,\ldots,6$ derived from the vectors defined in Eq. (2) by replacing τ with an approximate rational number (the ratio between two successive Fibonacci numbers f_n, given by the expressions $f_0=f_1=1$, $f_{n+1}=f_n+f_{n-1}$, where $n\geqslant 1$ and $\lim_{n\to\infty}[f_{n+1}/f_n]=\tau$), namely, $\bar{\tau}=f_{n+1}/f_n$, where f_{n+1} and f_n are integers. In this case, the orthogonal section of the ''projection tube'' is a distorted rhombic triacontahedron, so the projection procedure generates in the parallel space cubic structures locally isomorphic to the Amman–MacKay network. The smaller the difference between τ and its rational approximation $\bar{\tau}$, the larger the period of the resulting cubic structure (for $\bar{\tau}=f_{n+1}/f_n$ the period is $L=2\eta(f_{n+1}\tau+f_n)$), and the more accurate the approximation to the quasicrystal structure. For this reason, such structures are called periodic approximants to icosahedral quasicrystals. If ratios between successive Fibonacci numbers are treated as approximations to $\bar{\tau}$, the approximants are optimal in the sense discussed by Shaw *et al.*[28] Let us label the approximants by f_{n+1}/f_n, i.e., the notation ''f_{n+1}/f_n-approximant'' means that τ is replaced with f_{n+1}/f_n in constructing a distorted ''projection tube.'' Note that the f_{n+1}/f_n approximants have a bcc structure for $n=1,4,\ldots,3k+1,\ldots$ and a simple cubic structure for other n. Thus, the icosahedral quasicrystal can be considered a structural limit of the sequence of optimal cubic approximants with increasing period.

In order to obtain the Amman–MacKay network decorated in a complex manner, a six-dimensional cubic hyperlattice and a set of ''projection tubes'' are used (a separate ''tube'' for each position in the six-dimensional lattice). ''Projection tubes'' corresponding to different positions in the decorated Amman–MacKay network become more narrow with the increase in the structural fragment dimensionality. The polyhedrons representing orthogonal sections of ''projection tubes'' for sites on edges, two-dimensional and three-dimensional faces of the six-dimensional hyperlattice are the rhombic icosahedron, rhombic dodecahedron, and rhombohedron (prolate and oblate), respectively.[25]

3. ELECTRONIC PROPERTIES OF THE TWO-FRAGMENT MODEL OF ICOSAHEDRAL QUASICRYSTAL

Presently there is no universal technique for investigating electronic properties of quasiperiodic structures. Since quasicrystals have no long-range translational order, traditional techniques for calculating electron bands in solids based on Bloch's theorem do not apply. Methods that do not require translational invariance of the system, such as the recursion technique and the like, are poorly suited to the calculation of energy spectra and electron wave functions in quasicrystals.[12,13] The structures studied in the present work are therefore optimal cubic approximants to icosahedral quasicrystals. Thus, the icosahedral quasicrystal is considered to be a structural limit of the sequence of such approximants.

3.1. Model, approximations, and calculation techniques

Electronic properties of a two-fragment structural model of icosahedral quasicrystals have been studied in the tight-binding approximation. In order to minimize the number of adjustable parameters of the model, we used a Hamiltonian with constant hopping integrals between nearest neighbors. As follows from the results for one- and two-dimensional quasicrystals, such a Hamiltonian describes basic features of a quasicrystal and allows one to analyze qualitatively the effect of quasiperiodic properties on electron spectra of quasicrystals in one and two dimensions.[8–11,29] In this paper, we consider the ''central'' decoration of approximants with atoms of one type, namely, atoms with one s-orbital per atom are located at rhombohedral centers. In this case, the Hamiltonian can be expressed as

$$H=\sum_j |j\rangle\epsilon_j\langle j|+\sum_{j,j\neq i}|i\rangle t_{ij}\langle j|.$$

If atoms of only one type are present, the diagonal elements ϵ_j can be equated to zero. In this case the Schrödinger equation in the tight-binding approximation can be written in the form

$$\sum_j t_{ij}\psi_j=E\psi_i, \tag{4}$$

where the transfer integrals are equated to a nonzero constant ($t_{ij}=-1$) only in the case of nearest neighbors, i.e., atoms of rhombohedra with a common face (the inclusion of other neighbors does not significantly alter the results but complicates the calculations). Since an increase in the approximant order by one (the order of the 1/1-approximant is one, that of the 2/1-approximant is two, etc.) leads to a decrease in the Brillouin zone volume by a considerable factor (about τ^{-3}), the singular point of the icosahedral quasicrystal in the thermodynamic limit is $\mathbf{k}=0$. We therefore numerically diagonalized the Hamiltonian matrices (4) for various periodic approximants with periodic boundary conditions, and calculated the distribution of energy levels at $\mathbf{k}=0$.

In investigating the smoothness of the energy spectrum, we used the method of level statistics (LS-method).[10,11,29] It is based on two key equations. The first yields the relative number of gaps between levels with widths $\Delta E\leqslant BN^{\beta}$:

$$D(\beta)=\frac{1}{N-1}\sum_{j=1}^{N-1}\theta\left(\beta-\log_N\left[\frac{\epsilon_{j+1}-\epsilon_j}{B}\right]\right), \tag{5}$$

where N is the number of atoms in the approximant basis, $B=\epsilon_N-\epsilon_1$ is the total band width, θ is the Heaviside theta-function. The second equation yields the band fraction occupied by gaps between levels with widths $\Delta E\leqslant BN^{\beta}$:

$$F(\beta)=\frac{1}{B}\sum_{j=1}^{N-1}(\epsilon_{j+1}-\epsilon_j)\theta\left(\beta-\log_N\left[\frac{\epsilon_{j+1}-\epsilon_j}{B}\right]\right). \tag{6}$$

If $B=\text{const}$, these functions should satisfy the following conditions in the thermodynamic limit, irrespective of the smoothness of the spectrum[10,11,29]: $D(\beta)=1$ for $\beta>-1$ and $F(\beta)=0$ for $\beta<-1$. In crystalline and amorphous systems (with smooth spectra) the curves of $D(\beta)$ and $F(\beta)$ jump from zero to one at $\beta=-1$ in the thermodynamic limit.[10,11,29] Therefore an electron spectrum is considered unsmooth or singular if the gaps between levels as functions of the system dimension are not described by the function $1/N$ in the thermodynamic limit.

Localization properties of wave functions in the icosahedral quasicrystal have been studied using the statistics of $2p$-norms of wave functions,[10,11] which are defined as

$$\|\psi\|_{2p}\equiv\frac{\Sigma_n|\psi_n|^{2p}}{(\Sigma_n|\psi_n|^2)^p}, \tag{7}$$

where ψ_n are amplitudes of the electron wave function (expansion coefficients of the wave function in the tight-binding basis).

By considering the icosahedral quasicrystal as a structural limit of a sequence of optimal cubic approximants with increasing period, one can try to calculate thermodynamic limits for curves describing distributions of $2p$-norms of eigenvectors of the Hamiltonian in Eq. (4). In this paper, the statistical analysis of the distributions of $2p$-norms of wave functions in the icosahedral quasicrystal is based on the function $I_{2p}(\gamma)$,[10,11] which yields the relative number of states with $2p$-norms $\|\psi\|_{2p}\leqslant N^{\gamma}$, i.e.,

$$I_{2p}(\gamma)\equiv\frac{1}{N}\sum_{n=1}^{N}\theta(\gamma-\log_N\|\psi^{(n)}\|_{2p}). \tag{8}$$

The wave functions were classified in accordance with their normalization integrals.[10] Wave functions were considered to be delocalized if

$$\int_{|\mathbf{r}|<R}|\psi(\mathbf{r})|^2d\mathbf{r}\sim R^d,$$

where d is the space dimensionality. They were assumed to be localized when their finite norms existed:

$$\int_{\infty}|\psi(\mathbf{r})|^2d\mathbf{r}.$$

Wave functions which could not be normalized in an infinite space but were not delocalized were defined as ''critical.''

In order to analyze the effect of chemical disorder (disorder due to substitutional impurities[25]) on the electron spectrum and localization of wave functions in an icosahedral quasicrystal, atoms of a different element were randomly introduced into the basis of optimal periodic approximants (a two-component system was analyzed). In this case, the Hamiltonian can be expressed as

$$H=\sum_j|j\rangle\epsilon_j\langle j|+\sum_{j,j\neq i}|i\rangle t_{ij}\langle j|.$$

The diagonal elements ϵ_j can assume two possible values, namely ϵ^{A} and ϵ^{B}, depending on whether atom A or B is at the jth site. The difference between the two types of atoms can be characterized by the energy parameter δ:

$$\delta=\epsilon^{\mathrm{A}}-\epsilon^{\mathrm{B}}.$$

The transfer integrals for nearest neighbors are t^{AA}, t^{AB}, or t^{BB}, depending on which atoms occupy sites i and j. The difference between the transfer integrals for the different types of atoms can be characterized by the energy parameter δ_1:

$$\delta_1=t^{\mathrm{AA}}-t^{\mathrm{BB}}.$$

The transfer integral t^{AB} is described in the additive limit,[30]

$$t^{\mathrm{AB}}=(t^{\mathrm{AA}}+t^{\mathrm{BB}})/2.$$

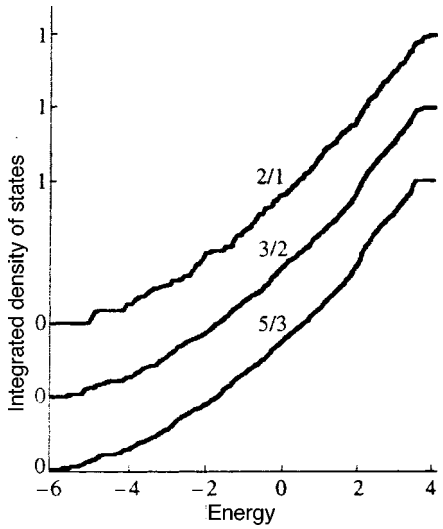

FIG. 1. Integrated density of states for periodic approximants 2/1, 3/2, and 5/3 of an icosahedral quasicrystal in arbitrary units.

3.2. Numerical calculations

In our work, we investigated four optimal cubic approximants to the icosahedral quasicrystal: 1/1, 2/1, 3/2, and 5/3. The unit cells of these approximants contain 32, 136, 576, and 2440 atoms, respectively. Figure 1 shows the integrated density of states calculated by taking energy levels at $\mathbf{k}=0$ for cubic approximants 2/1, 3/2, and 5/3 to the icosahedral quasicrystal (the curve for the 1/1-approximant is not given because its Brillouin zone is much larger than those of approximants of higher orders, so calculations based on energy levels at one $\mathbf{k}$-point of the Brillouin zone are not representative).

One can see in Fig. 1 that the integrated density of states converges rather rapidly and show no hierarchical structure of gaps, which is typical of the Cantor set in the spectrum of a one-dimensional quasicrystal. Figures 2 and 3 show $D(\beta)$ and $F(\beta)$ calculated for the three icosahedral approximants 2/1, 3/2, and 5/3. It is clear in Fig. 2 that $D(\beta)$ for these three approximants almost converges to the thermodynamic limit in the region $\beta<-1$. This means that a thermodynamically large number of gaps between levels are narrower than in traditional systems as $N\to\infty$ (in the thermodynamic limit the band gap between levels in crystalline and amorphous

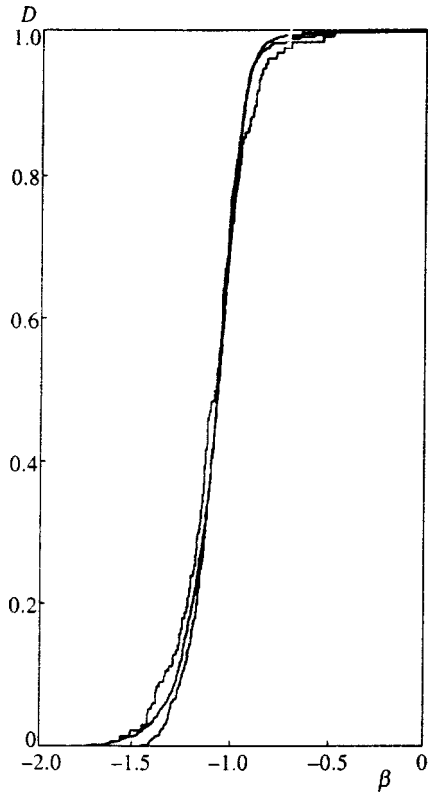

FIG. 2. Fraction $D(\beta)$ of gaps with widths $\Delta E \leqslant BN^{\beta}$ for periodic approximants 2/1, 3/2, and 5/3 of an icosahedral quasicrystal. Curves with smaller steps correspond to higher approximant orders. The total band width was equated to the limiting value $B_{\infty}=9.69$.

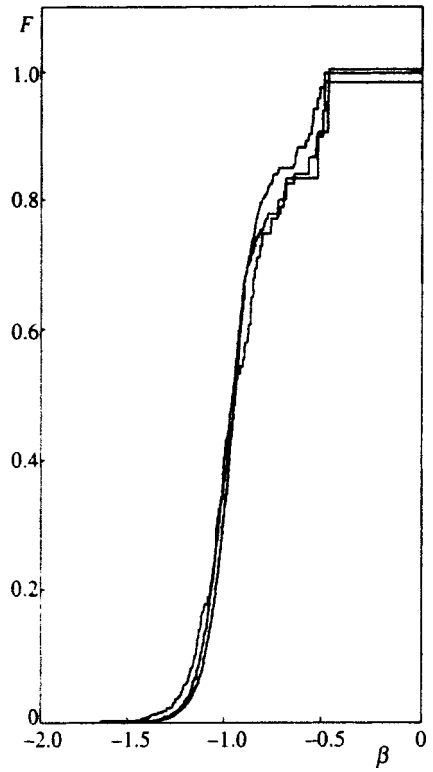

FIG. 3. Fraction $F(\beta)$ of the band filled with gaps between levels with widths $\Delta E \leqslant BN^{\beta}$ for periodic approximants 2/1, 3/2, and 5/3 of an icosahedral quasicrystal. Curves with smaller steps correspond to higher approximant orders. The total band width in the calculations was equated to the limiting value $B_{\infty}=9.69$.

systems as a function of the system size is given by[10] $\Delta E \propto 1/N$). Figure 3 illustrates the convergence of $F(\beta)$ for the approximants under consideration in the region $\beta > -1$, which indicates that a finite fraction of the energy band is occupied by gaps between levels wider than those in traditional systems. Thus, our results indicate that for $N \to \infty$ the spectrum of the icosahedral quasicrystal contains a singular component.

As in the two-dimensional case,[10] it is impossible to determine beyond doubt whether the energy spectrum of the icosahedral quasicrystal contains a regular component since the answer to this question requires investigation of $D(\beta)$ and $F(\beta)$ near $\beta=-1$ for approximants of higher orders.

To determine whether a singular component is present in the electron spectrum of an icosahedral quasicrystal, we calculated the density of electron states for the approximants under consideration. The density of states for the three lowest approximants, 1/1, 2/1, and 3/2, was calculated by the tetrahedron method using energy levels at 40 **k**-points in the irreducible part of the Brillouin zone of the corresponding approximant. In the calculation of the density of states for the 5/3-approximant we used four **k**-points in the irreducible part of the Brillouin zone. Owing to the relatively small volume of the Brillouin zone of the 5/3-approximant (as compared to the Brillouin zone volume for the first approximant), the reduction in the number of **k**-points has little effect on the resulting density of states.

This is confirmed by the small difference between the density of state curves for the 5/3-approximant calculated using one **k**-point ($\mathbf{k}=0$) and four **k**-points in the irreducible part of the Brillouin zone. The curves of the density of electron states for optimal approximants of higher orders are less smooth and more ''spiky.'' This confirms the conclusion given above about the presence of a singular component in the electron spectrum of icosahedral quasicrystals, since the density of states would otherwise converge to a smooth curve.

Moreover, the calculations indicate that the spectrum smoothness is different in different energy regions. The spectrum is smoother at low energies, while strong oscillations are mainly seen at higher energies. The width of the smooth section, however, decreases with increasing approximation order. It seems probable, therefore, that in the thermodynamic limit (in a quasicrystal), strong oscillations in the density of states, indicating the presence of a singular component in the spectrum, occur throughout the whole energy range, which means that the bands of the energy spectrum in the quasicrystal are flat, i.e., they have low dispersion, and the group velocities of electrons are very low. This is obviously why the electrical conductivity strongly depends on the composition of quasicrystalline alloys.[6]

Our calculations are in good agreement with Carlsson's calculations[20] of the density of electron states for the first four periodic approximants in the weak-binding approximation: we have detected gaps in the low-energy region for the

5/3-approximant, although no such gaps have been detected in the lower-order approximants.

The calculations of the density of electron states allowed us to determine the Lebesgue measure of energy spectra of the corresponding approximants. The Lebesgue measure of an energy spectrum was calculated to be the total length of the permitted segments in the energy range $[-6;6]$ (according to a well-known theorem on spectral boundaries,[30] the energy levels of the system under consideration lie within the range $-6 \leqslant E \leqslant 6$). The Lebesgue measures of the spectra for the 1/1-, 2/1-, 3/2-, and 5/3 approximants normalized to $B = \epsilon_n - \epsilon_1$ (owing to the fast convergence of the integrated density of states, the value of B for the 5/3 approximant was adopted for the various approximants; see Fig. 1) proved to be 0.95, 0.98, 0.98, and 0.97, respectively. We therefore conclude that the Lebesgue measure of the energy spectrum depends weakly on approximant size in three-dimensional icosahedral quasicrystals, unlike the case of a one-dimensional quasicrystal, in which the Lebesgue measure of the energy spectrum of a periodic approximant decreases with system size as a power law.[8–10]

The small differences among the Lebesgue measures of the latter two approximants analyzed in our calculations indicates that the energy spectrum of three-dimensional icosahedral quasicrystal should occupy a finite-width band of the energy scale. In addition, normalized energy-spectrum Lebesgue measures close to 1 in the approximants under consideration, along with a tendency for the results to converge, suggest that the spectrum of an icosahedral quasicrystal does not contain notable gaps. These results corroborate the ''non-Cantorian'' nature of the spectrum based on the integrated density of states for the first four optimal periodic approximants (Fig. 1).

Figure 4 shows $I_8(\gamma)$ calculated for three optimal cubic approximants, 2/1, 3/2, and 5/3. The graph demonstrates convergence of $I_8(\gamma)$ in a narrow range about $\gamma \approx -2.3$ in the thermodynamic limit. Moreover, the graph shows $I_8(\gamma)$ converging to a step function with the jump at $\gamma \approx -2.3$ as the approximant order increases. This means that in the thermodynamic limit ($N \to \infty$), the $2p$-norms ($p=4$) of almost all wave functions in an icosahedral quasicrystal depend on system size as $\|\psi\|_8 \approx N^{\gamma}$, where $\gamma \approx -2.3$.

For delocalized states, the $2p$-norm of the wave function depends on system size as $\|\psi\|_{2p}^{\text{ext}} \approx N^{1-p}$, as follows from Eq. (7), and exponentially decaying localized functions have the $2p$-norm $\|\psi\|_{2p}^{\text{exp.loc}} \approx 1$. The calculated $2p$-norms ($p=4$) of the wave functions in an icosahedral quasicrystal, as a function of system size, therefore rule out both delocalized and exponentially localized states. It is known, nonetheless, that the dependence of the $2p$-norm of the wave function on system size described by the function $N^{\gamma(p,\alpha)}$ applies to wave functions whose squared amplitudes decay as power-law functions[10,11] ($\alpha \geqslant 0$ is the localization index; $\alpha = 0$ and $\alpha \to \infty$ characterize delocalized and exponentially localized states, respectively). We have therefore assumed in this work that, as in a two-dimensional quasicrystal,[10,11]

$$|\psi|^2 \propto |\mathbf{r}|^{-2\alpha}. \tag{9}$$

In calculating the $2p$-norm of the wave function (9) over sufficiently large quasiperiodic fragments, we have found out that the function $\gamma(p,\alpha)|_{p=4}$ for $-3 \leqslant \gamma \leqslant -1.7$ is described by the curve in Fig. 5. These results were obtained numerically for the first five approximants of an icosahedral quasicrystal in order to demonstrate their convergence to the thermodynamic limit (the graph shows calculations for the fifth approximant, whose basis contains about ten thousand atoms). Figure 5 shows that $\gamma \approx -2.3$ corresponds to a localization index $\alpha \approx 0.75$. According to the classification based on the normalization integral, wave functions with such behavior must be classified as critical, since functions satisfying the condition $|\psi|^2 \propto |\mathbf{r}|^{-2\alpha}$, $\alpha > 0$, can be normalized in the three-dimensional case only if $\alpha > 3/2$.

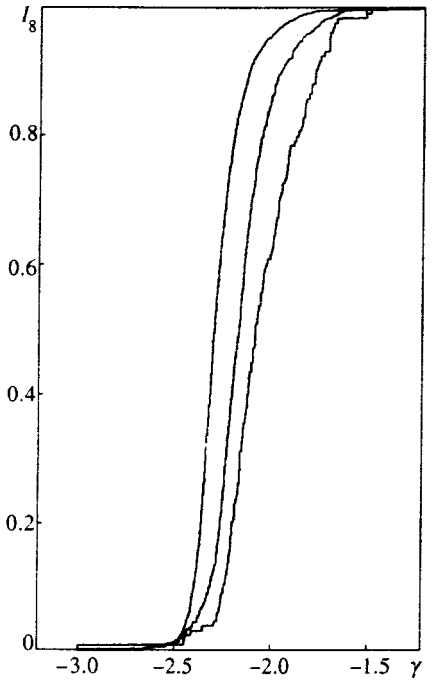

FIG. 4. Fraction $I_8(\gamma)$ of states with $\|\psi\|_8 \leqslant N^{\gamma}$ for periodic approximants 2/1, 3/2, and 5/3. Curves with smaller steps correspond to approximants of higher orders.

Figures 6, 7, and 8 show $D(\beta)$, $F(\beta)$, and $I_8(\gamma)$ at different degrees of chemical substitutional disorder characterized by the parameter δ and concentration of atoms of the second component B (the 5/3-approximant has been used). The graphs indicate that substitutional disorder makes the spectrum smoother than in a quasicrystal without substitu-

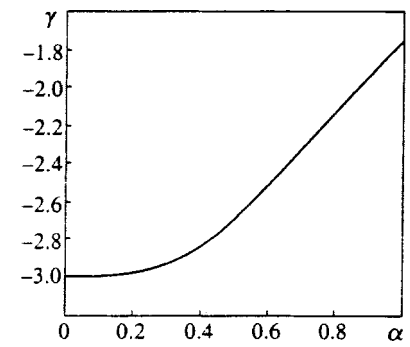

FIG. 5. Function $\gamma(p,\alpha)|_{p=4}$ for $-3 \leqslant \gamma \leqslant -1.7$.

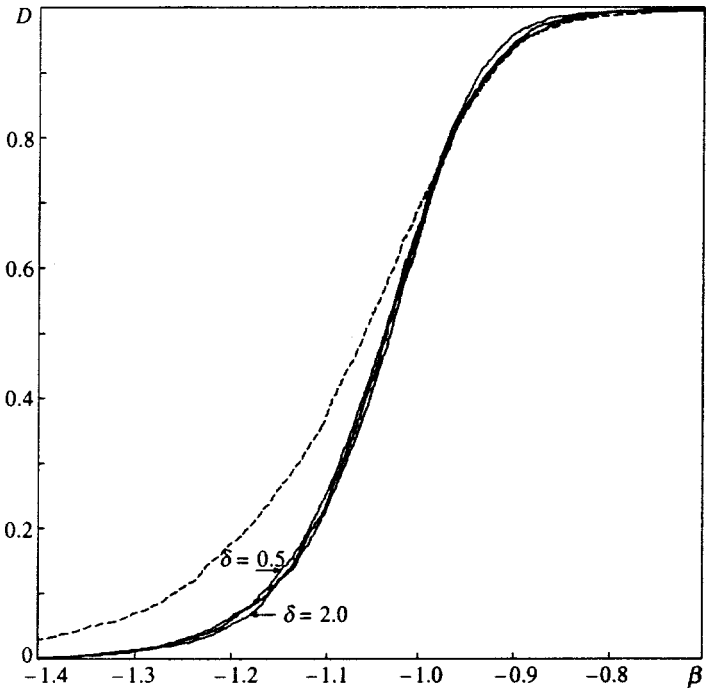

FIG. 6. $D(\beta)$ calculated for the 5/3-approximant at $\delta=0.5$, $\delta=1$, and $\delta=2$ (concentration C_B of B atoms and δ_1 are constant and equal to 50 at.% and 0, respectively). The dashed line shows the curve for the ideal case ($\delta_1=0$, $C_B=0$, and $\delta=0$).

tional disorder ($D(\beta)$ and $F(\beta)$ converge to a step function with jump at $\beta=-1$ as δ increases), and there is a tendency toward localization (the curves of $I_8(\gamma)$ shift to higher γ).

Similar effects, namely smoother spectra and a tendency toward localization, were also detected when the increase in chemical disorder was characterized by the parameter δ_1 and concentration of the second component B. ''Smoothing'' of the electron spectrum due to a higher degree of chemical disorder is illustrated by Fig. 9, which shows the integrated density of states for the 5/3-approximant at various concentrations of B atoms (the calculations were performed using four **k**-points in the irreducible part of the Brillouin zone).

4. DISCUSSION

Our results show that the electronic properties of a three-dimensional icosahedral quasicrystal with the structure of the Amman–MacKay network are substantially different from those of a one-dimensional quasicrystal (the Fibonacci chain). As in a one-dimensional quasicrystal, the spectrum of a three-dimensional quasicrystal contains a singular component. The structure of the electron spectrum of an icosahedral quasicrystal, however, is not self-similar, unlike the spectrum of one-dimensional quasicrystals, owing to the different topology of the system; as a result, the Lebesgue measure of

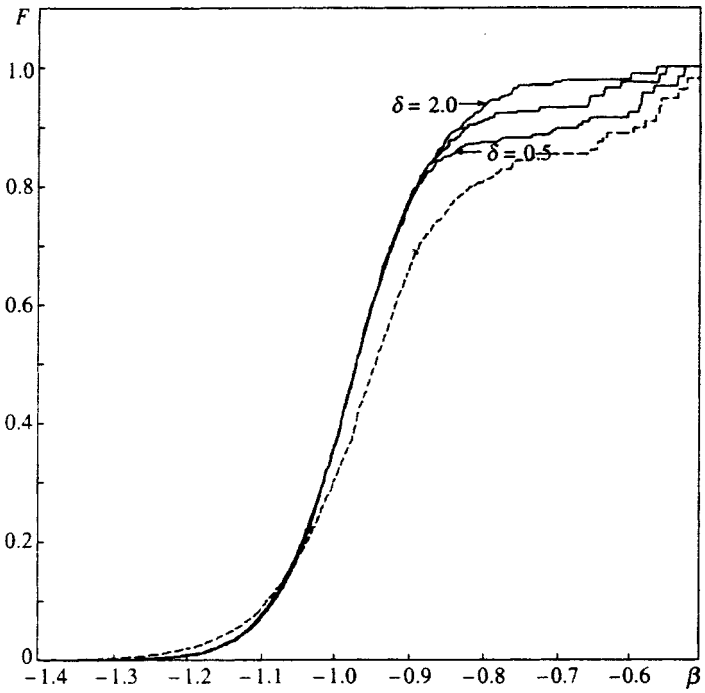

FIG. 7. Curves of $F(\beta)$ calculated for the 5/3-approximant at $\delta=0.5$, $\delta=1$, and $\delta=2$ (concentration C_B of B atoms and parameter δ_1 are constant and equal to 50 at.% and 0, respectively). The dashed curve refers to the ideal case ($\delta_1=0$, $C_B=0$, and $\delta=0$).

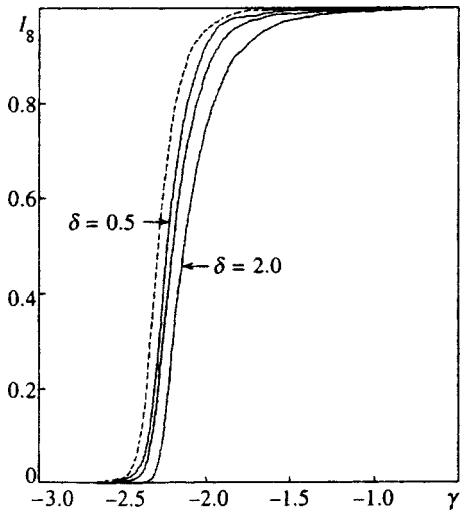

FIG. 8. Curves of $I_8(\gamma)$ calculated for the 5/3-approximant at $\delta=0.5$, $\delta=1$, and $\delta=2$ (concentration C_B of B atoms and parameter δ_1 are constant and equal to 50 at.% and 0, respectively). The dashed curve applies to the ideal case ($\delta_1=0$, $C_B=0$, and $\delta=0$).

the electron spectrum in a three-dimensional icosahedral quasicrystal is nonzero. Moreover, the spectrum of an icosahedral quasicrystal does not contain large gaps, as follows from our calculations, and the quasiperiodicity of the structure gives rise to large oscillations in the density of states throughout the entire energy range of the spectrum. This shape of the electron spectrum of a quasicrystal is consistent with strong electron scattering at the Fermi level (all electron states have zero group velocity) due to the everywhere dense set of reciprocal lattice vectors in the quasicrystal.[31]

Most of the wave functions in an icosahedral quasicrystal are ''critical'' (according to the classification based on the normalization integral) and fall off as a power-law function of the distance. This is in good agreement with previous results[32] in which electron and phonon spectra in a self-similar cluster package of pseudo-MacKay icosahedrons were analyzed in an AlPdMn alloy. Contrary to the claim made in Ref. 32, however, wave functions based on the self-similarity of a structure with localization parameter $\alpha=2.5$ are not ''critical,'' strictly speaking, but are instead localized

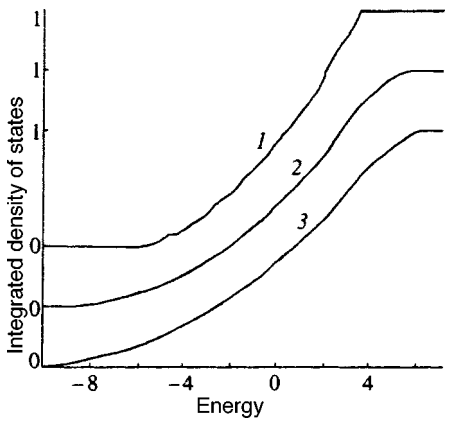

FIG. 9. Integrated density of electron states for an icosahedral quasicrystal with (*1*) $C_B=0$ at.%, (*2*) 30 at.%, and (*3*) 50 at.% in arbitrary units.

and fall off as a power law (in accordance with the classification based on the normalization integral; see above). An important difference from the one- and two-dimensional quasiperiodic packages (Fibonacci chains and Penrose tiling, respectively) is the single value of the localization index α for most of the wave functions, rather than a continuous spectrum of α.[10] When chemical disorder is introduced, the electron spectrum becomes ''smoother'' and wave functions tend to be localized (i.e., the localization index α increases with the degree of chemical disorder).

5. CONCLUSIONS

We have demonstrated that the electron spectrum of a perfect (defect-free) quasicrystal in three dimensions contains a singular (nonsmooth) component, and the wave functions are critical. Chemical disorder (substitutional disorder) makes the electron spectrum smoother and tends to localize electron states. The results enable us to concisely address the electrical conductivity of a three-dimensional quasicrystal (a more complete analysis will be given in a separate publication). Although the number of publications on this topic is relatively large, they have not yet given a clear explanation of the reason for the low conductivity of quasicrystals and its temperature dependence.

Different solutions to the problem of conductivity of icosahedral quasicrystals obtained in the weak-binding approximation yield inconsistent results: according to Kitaev,[33] it should be finite, and according to Sokoloff (a solution based on perturbation theory),[34] it should be infinite. Burkov *et al.*[35,36] modeled electron scattering in a three-dimensional quasicrystal with a multicomponent Fermi surface that contains a large but finite number of electron and hole ''pockets.'' These authors noted that in this case the Anderson localization conditions apply to electrons in ''pockets'' of the Fermi surface, which should lead to a temperature dependence $\sigma\propto\sqrt{T}$ of the conductivity (due to electron–electron scattering) when the electron states are localized within the valleys, and $\sigma\propto T$ when intervalley scattering is turned on at higher temperatures (phonon scattering).[37] By considering a quasicrystal to be a structural limit in a sequence of optimal approximants with increasing period, we previously concluded[31] that all electron states on the Fermi surface suffer strong scattering in a three-dimensional quasicrystal, that the wave functions should have oscillations on all scales (this is confirmed by analysis of the critical properties of the wave functions), and that hopping conductivity should occur.

Hopping conductivity was also suggested for the self-similar structure of an icosahedral AlPdMn quasicrystal whose basic structural unit was a 51-atom pseudo-MacKay icosahedron,[32] and for icosahedral quasicrystals in AlCuFe.[38] One experiment demonstrated that the low-temperature conductivity of $Al_{70.5}Pd_{21}Re_{8.5-x}Mn_x$ icosahedral quasicrystals is described by the Mott law for hopping conductivity with variable hop distance:

$$\sigma=\sigma_0\exp\left(-\left[\frac{T_0}{T}\right]^p\right),$$

where $p=1/2$ for $x=2$, 2.5, 3, 3.5, and $p=1/4$ for $x=0$ (the temperature ranged between 0.45 and 10 K).[39] This result indicates the existence of a mobility threshold and a ''metal glass'' state. It would be pointless to discuss the temperature dependence in detail, for it deserves a separate publication, but this result favors the critical properties of wave functions and their tendency to be localized when disorder is introduced into the system (especially near band edges).

If wave functions decay with distance according to a power law, $\psi \propto 1/r^{\alpha}$, the hop frequency $1/\tau$ (τ is the relaxation time) in the hopping mode as a function of temperature has the form $1/\tau \propto T^{\alpha/n}$ if the hop activation barrier ΔE is related to the hop length λ as $\Delta E = \text{const} \cdot \lambda^{-n}$, where n is a positive integer[32] (this condition is valid, as follows from experiment,[39] at temperatures that are not too low, when the Mott law for the hopping conductivity holds). Moreover, the ''smoothing'' of the electron spectrum due to disorder leads to a higher density of states at the Fermi level, given the experimental fact that the Fermi level in stable icosahedral quasicrystals is located at a local minimum of the density of states, i.e., a ''pseudogap''). Thus, the Einstein conductivity $\sigma \propto e^2 N(E_f) \lambda^2/\tau$ applied to quasicrystals in the hopping conductivity mode[31] yields a conductivity that increases with temperature and the degree of disorder in the system.

In conclusion, we thank P. A. Korzhavyi and D. V. Livanov for interest in this work and valuable remarks. This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-16143).

*)E-mail: olenev@trf.misa.ac.ru

[1] D. Shechtman, I. Blech, D. Gratias, and J. W. Cahn, Phys. Rev. Lett. **53**, 1951 (1984).
[2] D. Levine and P. J. Steinhardt, Phys. Rev. Lett. **53**, 2477 (1984).
[3] S. J. Poon, Adv. Phys. **41**, 303 (1992).
[4] F. S. Pierce, S. J. Poon, and Q. Guo, Science **261**, 737 (1993); C. Berger, T. Grenet, P. Lindqvist, P. Lanco, J. C. Grieco, G. Fourcaudot, and F. Cyrot-Lackmann, Solid State Commun. **87**, 977 (1993).
[5] H. Akiyama, Y. Honda, T. Hashimoto, K. Edagawa, and A. Takeuchi, Jpn. J. Appl. Phys., Part 2 **7**, L1003 (1993); Y. Honda, K. Edagawa, A. Yoshika, T. Hashimoto, and S. Takeuchi, Jpn. J. Appl. Phys., Part 1 **9**, 4929 (1994).
[6] Z. M. Stadnik, G. W. Zhang, A.-P. Tsai, and A. Inoue, Phys. Rev. B **51**, 11358 (1995).
[7] T. Fujiwara, S. Yamamoto, and G. T. Laissardiere, Phys. Rev. Lett. **71**, 4166 (1993).
[8] P. A. Kalugin, A. Yu. Kitaev, and L. S. Levitov, JETP Lett. **41**, 145 (1985).
[9] M. Kohmoto, B. Sutherland, and C. Tang, Phys. Rev. B **35**, 1020 (1987).
[10] H. Tsunetsugu, T. Fujiwara, K. Ueda, and T. Tokihiro, Phys. Rev. B **43**, 8879 (1991).
[11] H. Tsunetsugu, T. Fujiwara, K. Ueda, and T. Tokihiro, J. Phys. Soc. Jpn. **55**, 1420 (1986).
[12] G. Kasner, H. Schwabe, and H. Böttger, Phys. Rev. B **51**, 10454 (1995).
[13] M. A. Marcus, Phys. Rev. B **34**, 5981 (1986).
[14] K. Niizeki and T. Akamatsu, J. Phys.: Condens. Matter **2**, 2759 (1990).
[15] D. V. Olenev and Yu. Kh. Vekilov, JETP Lett. **63**, 113 (1996).
[16] D. V. Olenev and Yu. Kh. Vekilov, JETP Lett. **64**, 590 (1996).
[17] T. Fujiwara and T. Yokokowa, Phys. Rev. Lett. **63**, 333 (1991).
[18] G. T. Laissardiere and T. Fujiwara, Phys. Rev. B **50**, 5999 (1994).
[19] D. N. Davydov, D. Mayou, C. Berger, C. Gignoux, A. Neumann, A. G. M. Jansen, and P. Wyder, Phys. Rev. Lett. **77**, 3173 (1996).
[20] A. E. Carlsson, Phys. Rev. B **47**, 2515 (1993).
[21] A. P. Smith and N. W. Ashcroft, Phys. Rev. Lett. **59**, 1365 (1987).
[22] J. Friedel, Helv. Phys. Acta **61**, 538 (1988).
[23] V. G. Vaks, V. V. Kamyshenko, and G. D. Samolyuk, Phys. Lett. A **132**, 131 (1988).
[24] M. A. Fradkin, JETP Lett. **49**, 705 (1989).
[25] D. V. Olenev, P. A. Korzhavyi, and Yu. Kh. Vekilov, Zh. Éksp. Teor. Fiz. **104**, 4130 (1993) [JETP **77**, 998 (1993)].
[26] V. Elser, Phys. Rev. B **32**, 4892 (1985).
[27] C. L. Henley, Phys. Rev. B **34**, 797 (1986).
[28] L. J. Shaw, V. Elser, and C. L. Henley, Phys. Rev. B **43**, 3423 (1991).
[29] L. Yamamoto and T. Fujiwara, Phys. Rev. B **51**, 8841 (1995).
[30] H. Ehrenreich and L. Schwartz, *The Electronic Structure of Alloys*, Academic Press, New York (1976).
[31] Yu. Kh. Vekilov, P. A. Korzhavyi, and D. V. Olenev, JETP Lett. **62**, 372 (1995).
[32] C. Janot, Phys. Rev. B **53**, 181 (1996).
[33] A. Yu. Kitaev, JETP Lett. **48**, 298 (1988).
[34] J. B. Sokoloff, Phys. Rev. B **36**, 6361 (1987).
[35] S. E. Burkov, A. A. Varlamov, and D. V. Livanov, JETP Lett. **62**, 361 (1995).
[36] S. E. Burkov, A. A. Varlamov, and D. V. Livanov, Phys. Rev. B **53**, 11504 (1996).
[37] B. L. Altshuler and A. G. Aronov, *Electron–Electron Interaction in Disordered Systems*, A. L. Efros and M. Pollak (eds.), Elsevier Science Publishers B.V., Amsterdam (1985).
[38] D. Mayou, C. Berger, F. Cyrot-Lackmann, T. Klein, and P. Lanco, Phys. Rev. Lett. **70**, 3915 (1993).
[39] Q. Guo and S. J. Poon, Phys. Rev. B **54**, 12793 (1996).

Translation provided by the Russian Editorial office.

Magnetic susceptibility anisotropy and low-dimensional antiferromagnetism of CuO

T. I. Arbuzova,*) I. B. Smolyak, A. A. Samokhvalov, and S. V. Naumov

Institute of the Physics of Metals, Russian Academy of Sciences, Ural Branch, 620219 Ekaterinburg, Russia
(Submitted 17 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 1026–1035 (March 1998)

The anisotropy of the magnetic susceptibility χ and the influence of oxygen vacancies in CuO single crystals on it are investigated. The temperature dependences of $\chi(T)$ along the a, b, and c axes in the range $60<T<600$ K and the behavior of the field dependence of the magnetization $\sigma(H)$ above and below the Néel temperature T_N are plotted for a crystal before and after heat treatment. The $\chi(T)$ curves have the form characteristic of low-dimensional systems, which become three-dimensional when the temperature is lowered. The character of the $\chi(T)$ curves remains unchanged after annealing. Oxygen vacancies have practically no influence on the a-axis magnetic susceptibility, but they alter the absolute values of the b- and c-axis susceptibilities. The significant effects of reducing the oxygen concentration include a decrease in the magnitude of the low-temperature anomaly (increase) in χ and an increase in the minimum value of χ. The results of the calculations of the exchange parameter I/k and the g factor are discussed in terms of the Heisenberg and Ising models for a one-dimensional system. © *1998 American Institute of Physics.* [S1063-7761(98)01903-9]

1. INTRODUCTION

Cupric oxide (CuO) occupies a special place among the semiconductor $3d$ oxides because of the uniqueness of some of its physical properties. Unlike NiO, CoO, FeO, and MnO (which have fcc lattices), CuO has a low-symmetry monoclinic crystal structure and temperature behavior of the magnetic susceptibility χ that is unusual for three-dimensional (3D) antiferromagnets. In polycrystals the susceptibility increases, rather than decreases, above the Néel temperature T_N, passes through a broad maximum near $T=550$ K, and then slowly decreases as the temperature rises.[1,2] Its electrical and optical properties also have several special features.[3]

Copper compounds are usually good examples of isotropic Heisenberg systems. The formation of linear (1D) chains and planar two-dimensional (2D) magnets is typical of them. Their formation can be due to the Jahn–Teller effect resulting from the low-symmetry environment of the Cu^{2+} ions. The data on the susceptibility,[1,4–9] magnetization,[10–13] and specific heat[14,15] of CuO indicate strong spin correlations and a possibility of low-dimensional antiferromagnetic ordering above T_N. The dimensionality of the magnetic system influences the thermodynamic functions more strongly than does the structure of the crystal lattice. When the dimensionality decreases from 3D to 1D, the short-range-order effects are enhanced. The characteristic features of low-dimensional systems include flattened maxima of the susceptibility χ and the specific heat C at temperatures comparable to the exchange interaction parameter.

Cupric oxide (CuO) has a simple chemical formula and can serve as a model object for high-T_c superconductors. Many of the properties of CuO resembles those of $YBa_2Cu_3O_6$ and La_2CuO_4. They all contain identical structural fragments, i.e., $Cu(O)_4$ parallelograms formed by Cu–O chains, are p-type semiconductors with a low charge-carrier mobility and antiferromagnets with fairly high Néel temperatures $T_N>200$ K, and have similar values of the saturation magnetic moment per Cu^{2+} ion, which are considerably less than the theoretical value. Strong spin correlations are observed in the temperature range $T>T_N$.

The exchange interaction between the Cu^{2+} ions (their electronic configuration is $3d^9$, and the spin $S=1/2$) is mediated mainly by O^{2-} ions. Defects in the oxygen sublattice should have an appreciable influence on the magnetic properties of CuO, and the intentional introduction of such defects can be a good way to reveal the special features of the magnetic bonds and ordering. However, it is known that CuO has a very narrow region of oxygen homogeneity, whose size cannot be determined by ordinary methods.[16] At the same time, the existing disparities between the published data[6–9] on both the magnitude and the character of the variation of $\chi(T)$ in the principal crystallographic directions can be attributed to the differences in the quality of the samples and their defect density. Therefore, for CuO magnetic measurements, particularly measurements of $\chi(T)$, are an effective tool, which makes it possible to investigate the influence of oxygen vacancies and other defects on the magnetic properties.

2. SAMPLES AND MEASUREMENT METHOD

Cupric oxide single crystals were grown by forming a solution in a melt in the CuO–BaO system in ZrO_2 crucibles. A mixture of CuO and $BaCO_3$ powders (of very-high-purity grade) in the required proportions was heated in air to 1000 °C. The mixture melted completely. The melt was held at 1000 °C for 6 h and then cooled to 890 °C at the rate of 1 °C/h. Below 890 °C the system was cooled to room temperature together with the furnace. The crystals were mechanically extricated from the crucibles. To obtain samples

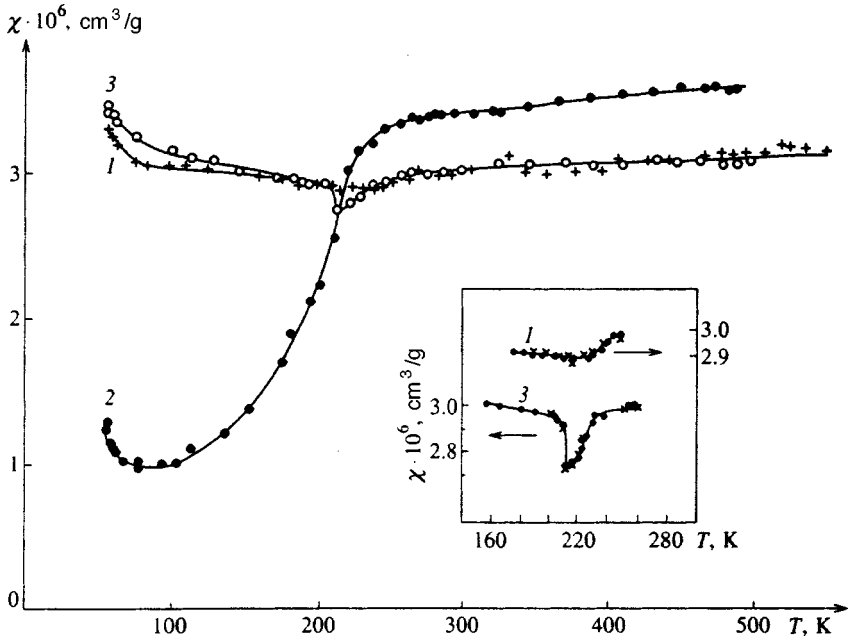

FIG. 1. Temperature dependence of the *a*- (*1*), *b*- (*2*), and *c*-axis susceptibility (*3*) for a stoichiometric CuO single crystal. Inset: plots of $\chi(T)$ near T_N for $\mathbf{H}\|\mathbf{a}$ (*1*) and $\mathbf{H}\|\mathbf{c}$ (*3*) during heating (○) and cooling (+).

with different defect densities, the single crystals were subjected to prolonged anneals for 20–80 h at 500 °C in oxygen (nominally stoichiometric samples) and at 950 °C with sudden cooling in liquid nitrogen (samples with oxygen defects). Such heat treatments do not alter the lattice constants of the crystals.

X-ray diffraction investigations of the samples were performed at room temperature on a DRON-2 diffractometer using chromium K_α radiation. The single crystals had a monoclinic $C2/c(15)$ lattice[1] with the parameters $a=4.677$ Å, $b=3.422$ Å, $c=5.129$ Å, and $\beta=99.50°$. The CuO single crystals had the shape of quadrangular prisms measuring up to $3\times3\times10$ mm^3. The natural large faces were oriented in the (110) plane.

The magnetic measurements were performed on a Faraday magnetic balance in the temperature range 60–600 K and in the range of magnetic fields $H\leqslant13$ kOe. The sensitivity of the balance permits measurement of the susceptibility down to 10^{-8} cm^3/g. The temperature dependence of the susceptibility was measured mainly in a 9 kOe field. The measurement error was less than 3%.

3. MAGNETIC PROPERTIES OF CuO

The first thorough measurements of the susceptibility of a CuO polycrystal over the broad temperature range $4.2<T<1100$ K were performed by O'Keefe and Stone.[1] According to their data, the susceptibility $\chi=2\times10^{-6}$ cm^3/g remains constant as the temperature rises up to 140 K. A somewhat different type of dependence of $\chi(T)$, viz., an increase in the susceptibility as the temperature drops below 100 K, was obtained in Refs. 4–6, 8, 9, and 14.

We investigated a large number of CuO polycrystals and single crystals. Our single crystals exhibited an increase in χ as T decreases in the low-temperature region. The magnitude of the change in χ, the temperature range with a constant value of χ, and its minimum value depend on the defect density in the samples. Decreasing the number of oxygen vacancies by additional annealing in a stream of O_2 reduced the minimum value to $\chi=1.0\times10^{-6}$ cm^3/g. As the concentration of oxygen vacancies increased, the maximum on the $\chi(T)$ curve in the high-temperature region shifted to a lower value of T and became less pronounced.

The temperature dependence of the magnetic susceptibility of the CuO single crystals in the principal crystallographic directions was studied in Refs. 6, 7, and 9. They point out the anisotropy of χ both in the range $T<T_N$ and considerably above T_N. The anisotropy of $\chi_\|$ and $\chi_\perp$ in the region of magnetic ordering is characteristic of collinear antiferromagnets. Above T_N the susceptibility should be isotropic, and the anisotropy of χ can be observed only in highly anisotropic antiferromagnets.[17] In any case, for a fixed value of the magnetic moment of the ions the susceptibility should decrease with increasing temperature in the region of magnetic disorder for $T>T_N$; however, in CuO it increases.

The temperature dependence of the susceptibility of CuO has the form that is typical of low-dimensional (1D or 2D) antiferromagnetic systems, which undergo a phase transition to a 3D state with long-range interaction order when the temperature is lowered.[18] To isolate the influence of oxygen on the magnetic properties of CuO, we investigated the $\chi(T)$ curves of a single crystal weighing 15.6 mg and measuring $2\times1.7\times2.5$ mm^3 before and after heat treatment.

Figure 1 presents plots of the temperature dependence of the susceptibility in the principal directions *a* (χ_a), *b* (χ_b), and *c* (χ_c) in a field $H=8.9$ kOe for a CuO single crystal annealed in a stream of oxygen. In the magnetically ordered region ($T<213$ K) the spin of the Cu^{2+} ion is directed along the [010] axis; therefore, χ_b is the parallel susceptibility, and χ_a and χ_c comprise the perpendicular susceptibility.[10–12] The parallel susceptibility should tend to zero as $T\to0$ in a collinear antiferromagnet, but a finite value of χ_b is always observed in CuO. This is an indication of partial activation of the orbital angular moment and spin-orbit coupling. As the

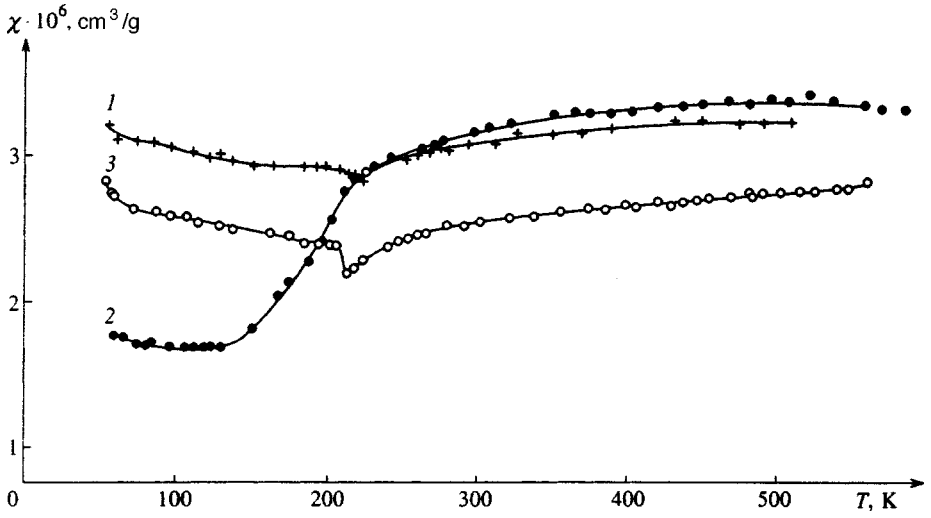

FIG. 2. Temperature dependence of the *a*- (*1*), *b*- (*2*), and *c*-axis susceptibility (*3*) for a CuO single crystal annealed in air.

temperature rises in the range $T>140$ K, the value of χ_b increases quite dramatically. However, at $T_N=230$ K the susceptibility peak typical of Néel antiferromagnets is not observed, and only the slope of the $\chi_b(T)$ curve changes. As the temperature rises further, the increase in χ_b is fairly weak. The values of χ_a and χ_c remain almost constant and close to one another in the temperature range investigated. A small dip in χ was discovered for the $\mathbf{H}\|\mathbf{c}$ direction in all the single crystals. According to neutron-diffraction investigations, at $T=212$ K the magnetic structure of CuO undergoes a transition from collinear antiferromagnetic ordering to helical ordering with a magnetic moment in the ac plane.[10] A second specific-heat peak is observed at the same temperature (the first is observed at T_N), indicating a phase transition. Hysteresis of the intensity of the magnetic reflections was observed in the transition region at 212–230 K;[11] therefore, we performed finer measurements of the susceptibility in both the heating and cooling regimes. No hysteretic phenomena were discovered in the behavior of $\chi(T)$ along the a, b, and c axes near 212 K (see the inset in Fig. 1).

To increase the number of oxygen vacancies, the CuO single crystal was annealed in air at 950 °C and investigated again. Figure 2 presents plots of the temperature dependence of χ_a, χ_b, and χ_c for the same crystal annealed in air. It follows from a comparison of Figs. 1 and 2 that the form of the temperature dependences of the susceptibility is maintained. Oxygen vacancies have practically no influence on the $\chi_a(T)$ curve, but they cause significant changes in the absolute values of χ_b and χ_c. In the defect-containing single crystal the parallel susceptibility increases for $T<T_N$ and decreases for $T>T_N$, i.e., the changes in the absolute value of $\chi_b(T)$ become weaker. Inversion of the $\chi_a(T)$ and $\chi_c(T)$ curves is observed at low temperatures. In the stoichiometric crystal $\chi_c>\chi_c$ is obtained for $T<T_N$, and these susceptibilities practically coincide at $T>T_N$. In the defect-containing single crystal $\chi_c<\chi_a$ is observed over the entire range of variation of the temperature. In addition, the temperature of the minimum of χ_c remains equal to 213 K before and after annealing. The most significant effects of the oxygen vacancies are a decrease in the magnitude of the low-temperature increase in the susceptibility along all the axes and an increase in the minimum value of χ_b for $T<140$ K.

The temperature dependences of the susceptibility for the single crystal with oxygen vacancies are in good agreement with the curves obtained in Ref. 7 with respect to the character of the variation of χ and its absolute values. Above T_N, χ_b has the largest value, and χ_c has the smallest value. However, the claim made by Kobler and Chattopadhyay in Ref. 7 that the negative jump $-\Delta\chi_c$ at $T=212$ K in CuO is equal to the positive jump $\Delta\chi_b$ and that there is no variation of the a-axis susceptibility is incorrect. This behavior of the susceptibility is characteristic of their sample. Unlike Kobler and Chattopadhyay,[7] we observed small, but noticeable discontinuities on the course of the $\chi_a(T)$ curve near $T=220$ K. The $\chi(T)$ curves for the principal axes in Ref. 9 differ somewhat from our data and the data in Ref. 7. More specifically, at $T>200$ K the c-axis susceptibility has the largest value, and $\chi_b\simeq\chi_a$. An analysis of the data on $\chi(T)$ for different single crystals and the results in Refs. 5–9 reveals that the sharpest changes in χ occur in the noncollinear phase at 212–230 K. The quality of the single crystals has a weak influence on T_N, but can alter the relationship between the values of χ along the a, b, and c axes.

The disparities in the literature in regard to the temperature and field dependences of the magnetization of CuO are probably attributable to technological details. A nonlinear dependence of the magnetization σ on the field was observed for $T\leqslant 250$ K in several polycrystals[5,19] and single crystals.[8] The small spontaneous magnetization is attributed to the weak ferromagnetism appearing in the antiferromagnet when there is a deviation of the spins from the collinear direction (tilting of the spins, canted antiferromagnetism). It is not clear whether the weak ferromagnetism is an inherent property of CuO or is caused by defects, particularly by oxygen vacancies. We recorded the field dependence of the magnetization $\sigma(H)$ for polycrystals and single crystals in the range of fields $1.5<H<15$ kOe at various temperatures. Although we cannot evaluate the behavior of $\sigma(H)$ in weak fields, extrapolation of the magnetization to $H=0$ yields information on the spontaneous magnetization and weak ferromagnetism.

In single crystals the spontaneous magnetization differs

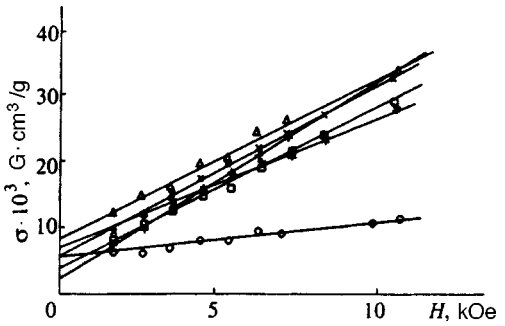

FIG. 3. Field dependence of the *a*- (×, □), *b*- (○, +), and *c*-axis magnetization (△, *) for a stoichiometric CuO single crystal at $T=56$ K (×, ○, △) and $T=300$ K (□, +, ∗).

from zero both above and below T_N. Figure 3 presents plots of $\sigma(H)$ for a stoichiometric single crystal along the a, b, and c axes at 56 and 300 K. The spontaneous magnetization varies from 1.5×10^{-3} to 8×10^{-3} G·cm^3/g. A similar dependence of $\sigma(H)$ was observed for the a direction at $T=115$ K in Ref. 8. In a single crystal with oxygen vacancies the $\sigma(H)$ curves remain qualitatively unchanged, and the residual magnetization is maintained in the same range. The small value of the residual magnetization indicates weak canting of the Cu^{2+} spins, which is not associated with oxygen vacancies.

4. DISCUSSION OF RESULTS

The temperature dependence of the susceptibility of CuO has the form characteristic of low-dimensional magnetic systems. Below $T_{N2}=213$ K CuO is a collinear antiferromagnet, but above $T_{N1}=230$ K strong spin correlations are observed, and the short-range order remains intact at least up to 450 K. It should be noted that the helical magnetic structure discovered in the intermediate range $212<T<230$ K can be realized in both three-dimensional and one-dimensional systems. The nature of the long-range magnetic order in low-dimensional systems differs from the usual order in three-dimensional magnetic arrays, as is manifested by the appreciable decrease in the mean spin due to low-energy excitations. The decrease in the mean spin is inversely proportional to the value of the spin and the number of interacting neighbors.

The magnetic structure of CuO can be represented in the form of Cu–O–Cu chains parallel to the [101] direction with strong antiferromagnetic superexchange I_1 through O^{2-} ions within a chain and a weaker ferromagnetic interaction I_2 between chains. The relation between I_1 and I_2 is of great importance. If I_2 is significantly smaller than I_1, then as the temperature at which kT becomes comparable to I_2 rises, the interaction between chains vanishes and CuO can go over to a one-dimensional antiferromagnetic state. If the interaction between chains in the ac plane is nonzero, the probability of two-dimensional antiferromagnetic ordering is not ruled out. The interaction within a chain leads to spin correlation and short-range-order effects. One manifestation of the short-range-order effects in the low-dimensional systems is the flattened maximum on the $\chi(T)$ curve near $T\approx I/k$.

At the present time there is no clear verdict regarding the dimensionality (1D or 2D) of the magnetic system in CuO and the model (the Heisenberg, Ising, or XY model) that should be used to describe it. Since CuO is a system which initially consists of independent antiferromagnetic chains and since Cu^{2+} ions are often described by the Heisenberg model, we attempted to describe the temperature dependence of the susceptibility in the range $T>T_N$ using a one-dimensional isotropic chain with $S=1/2$.[20] The best agreement of the experimental $\chi(T)$ curve for a polycrystal with the calculated curves for different values of the exchange parameter I/k and the g factor was obtained for $I/k=430$ K and $g=1.97$.

In single crystals magnetic susceptibility anisotropy is observed above T_N in parallel and perpendicular fields relative to the direction of the spin. In the Heisenberg model the exchange within a chain is isotropic; therefore, the calculated susceptibility should also be isotropic. The anisotropy of χ in this model can also be related to the anisotropy of the g factor. In this case the maximum susceptibility, which is specified by the relation

$$\chi_{\max}|I|/Ng^2\mu_B^2=0.07346,$$

depends on the g factor, and the temperature of the susceptibility maximum for all the field directions should be the same. In CuO the value of the g factor is unknown, since single crystals do not exhibit an EPR signal in the temperature range 80–450 K.[21] Weak anisotropy of the g factor of Cu^{2+} ions is usually observed in compounds: $g_{\parallel,\perp}=1.9-2.4$. The small g-factor anisotropy ($g_\parallel=2.02$ and $g_\perp=1.88$ for $I/k=390$ K) can account for the observed anisotropy of $\chi_\parallel$ and $\chi_\perp$ in the restricted range $T>400$ K, but the entire temperature dependence of the susceptibility above T_N is not described by the 1D Heisenberg model. Treatment of the experimental plots of $\chi(T)$ along the a, b, and c axes for stoichiometric and defect-containing single crystals using the numerical results of Bonner and Fisher[20] for a Heisenberg chain with $S=1/2$ gave strongly differing values of the exchange parameters and the g factor. Therefore, CuO is not an isotropic system, and the 1D Heisenberg model is inapplicable.

An anisotropic antiferromagnetic chain can be described by the Ising model or the XY model.[18,22] We estimated the exchange parameters and the g factor for a CuO single crystal before and after annealing using the theoretical formulas for an Ising chain. The calculations showed that the values of I/k and the g factor for the different axes differ strongly, i.e., the Ising model is also inapplicable. It should, however, be noted that the exchange parameters for the a and c directions are fairly close and are significantly greater than the value of I/k for the b axis. Therefore, the strongest interaction between Cu ions occurs in the ac plane, in which the Cu–O–Cu chains lie.

The plots of the temperature dependence of the susceptibility of CuO are very similar to the $\chi(T)$ curves for the 1D antiferromagnet $CsCoCl_4$,[23] which can be described by the XY model. The XY model is characterized by strong g-factor anisotropy ($g_\perp \gg g_\parallel$) and the absence of susceptibility anisotropy in the xy plane. It can be assumed that CuO is either a

one-dimensional compound corresponding to the XY model or a two-dimensional compound. The magnetic properties of 1D and 2D compounds do not differ strongly; therefore, it is often difficult to correctly choose the model. The estimates of I/k and the g factor based on different models should be treated with caution, since it is usually assumed in the calculations that the spin of the Cu^{2+} ions is directed along the b axis. However, the transition from three-dimensional to low-dimensional antiferromagnetism is accompanied by a change in the direction of the spin. This is indicated by the neutron-diffraction data and the relation between the values of χ along the principal axes. In anisotropic low-dimensional models the parallel susceptibility should be significantly smaller than the perpendicular component, and the maximum of $\chi_\perp$ shifts to a lower temperature in comparison to $\chi_\parallel$. As we see from Figs. 1 and 2, χ_b has the largest value above T_N. In this temperature range the susceptibility along the principal axes is apparently a result of the superposition of $\chi_\parallel$ and $\chi_\perp$.

Cupric oxide (CuO) is a frustrated antiferromagnet because of the closeness of the values of the competitive ferromagnetic exchange J_1 between nearest neighbors and the antiferromagnetic interaction J_2 between next-nearest neighbors. According to the theoretical models,[24] a first-order phase transition from Néel ordering to helical ordering should be observed in a frustrated 2D Heisenberg antiferromagnet with $S=1/2$ at $J_2/J_1=1.1$, and a first-order transition from helical antiferromagnetism to a quantum spin-fluid state should be observed at $J_2/J_1=1.65$. Qualitatively similar behavior is observed for the magnetic system in CuO. The transition from a 3D state to a low-dimensional state with strong spin correlation in a definite direction apparently does not occur at a critical point, but is spread over a temperature range. This is indirectly supported by the nonzero residual magnetization and the weak variation of χ_a, χ_b, and χ_c with the temperature at $T>T_N$. Microscopic magnetically ordered regions are apparently preserved above T_N. The absence of dramatic changes along the course of the $\chi_a(T)$ curve is not an indication of very strong spin correlations.[7] It can be attributed to the fact that the a-axis susceptibility is equal to $\chi_\perp$ below 212 K, while the component $\chi_\parallel$ becomes important above 212 K.

As we have already noted above, defects influence the magnitude of the susceptibility and the form of the $\chi_\parallel(T)$ and $\chi_\perp(T)$ curves. As the number of oxygen vacancies increases, the maximum of χ shifts to a lower value of T. This can be caused by the cleavage of antiferromagnetic bonds and by a decrease in the effective exchange. The appearance of small regions with broken antiferromagnetic bonds increases the susceptibility. In polycrystalline CuO films, in which there is a possibility for a larger number of defects than in bulk samples, the Néel temperature is considerably lower ($T_N<160$ K).[25] The ''paramagnetic'' increase in the susceptibility for CuO single crystals and polycrystals at low temperatures is attributed to the weak ferromagnetism appearing because of the canting of the spins near defects and to the appearance of free Cu^{2+} ions.[4,7,8] In our opinion, the weak ferromagnetism is, in fact, attributable to a deviation of the direction of the spins from collinear ordering due to frustration of the exchange interactions, but the formation of isolated paramagnetic Cu^{2+} centers is unlikely. A paramagnetic resonance signal was not observed for any of the polycrystals and single crystals not containing extrinsic impurities. The absence of an EPR signal is attributed to the long-range antiferromagnetic order at $T<T_N$ and to the strong spin correlations of the Cu^{2+} ions above T_N. It can be seen from Figs. 1 and 2 that the presence of defects in CuO single crystals does not cause enhancement of the paramagnetic increase in χ. The decrease in the low-temperature anomaly in a defect-containing crystal is fictitious because of the general increase in χ. An increase in the susceptibility with decreasing T, which is qualitatively similar to the behavior of $\chi_\perp$ and $\chi_\parallel$ as a function of the temperature, is possible in alternated chains[20] and frustrated antiferromagnets. The low-temperature increase in χ can also be caused by a repeated transition to a spin-glass state or a state of low-dimensional ordering. In the closely related compound $La_2CuO_{4-\delta}$ a decrease in the intensity of the 3D neutron Bragg scattering and, accordingly, a sharp increase in the 2D quasielastic scattering were observed when the temperature was lowered below 30 K.[26] The probability of such a transition in CuO has not been ruled out.

5. CONCLUSIONS

The magnetic susceptibility of CuO is anisotropic over the entire temperature range investigated. The plots of $\chi(T)$ along the principal axes indicate 3D antiferromagnetism below T_N and low-dimensional ordering of a special type (a quantum spin-fluid state) above T_N. Of all the known low-dimensional magnets, CuO and the closely related perovskite-like compounds have the highest Néel temperatures (T_N) and, accordingly, the strongest exchange interaction. Defects lead to partial cleavage of the antiferromagnetic bonds. Oxygen vacancies do not qualitatively influence the form of $\chi(T)$, but change the absolute values of χ_b and χ_c. The noticeable effects of oxygen vacancies include a decrease in the magnitude of the low-temperature increase in χ and weaker changes in $\chi_b(T)$. For $T>T_N$ the b-axis susceptibility χ_b has the largest value, indicating a change in the direction of the spin of the Cu^{2+} ions. Treatment of the experimental data on the basis of the 1D Heisenberg and Ising models has shown that these models do not describe the behavior of $\chi(T)$ in CuO.

We thank V. V. Dyakin and B. V. Karpenko for their assistance in treating the experimental results.

This work was carried out as part of the Russian State Program (012) ''Surface Atomic Structures'' (Project No. 95-2.10).

*)E-mail: magsemi@ifm.e-burg.su

1)The number in parentheses is the designation of the group in the international classification.

[1] M. O'Keefe and F. S. Stone, J. Phys. Chem. Solids **23**, 261 (1962).
[2] T. I. Arbuzova, A. A. Samokhvalov, I. B. Smolyak *et al.* JETP Lett. **50**, 34 (1989)].
[3] A. A. Samokhvalov, N. A. Viglin, B. A. Gizhevskiĭ *et al.*, Zh. Éksp. Teor. Fiz. **103**, 951 (1993) [JETP **76**, 463 (1993)].

[4] M. S. Seehra, Z. Feng, and G. R. Gopalakrishnan, J. Phys. C: Solid State Phys. **21**, 1051 (1988).
[5] K. Muraleedharan, C. K. Subramaniam, N. Venkataramani *et al.*, Solid State Commun. **76**, 727 (1990).
[6] T. I. Arbuzova, A. A. Samokhvalov, I. B. Smolyak, *et al.*, J. Magn. Magn. Mater. **95**, 198 (1991).
[7] U. Kobler and T. Chattopadhyay, Z. Phys. B **82**, 383 (1991).
[8] C. B. Azzoni, A. Paleari, and G. B. Paravicini, J. Phys.: Condens. Matter **4**, 1359 (1992).
[9] F. Marabelli, G. B. Paravicini, and P. Wachter, Solid State Commun. **86**, 131 (1993).
[10] J. B. Forsyth, P. J. Brown, and B. M. Wanklyn, J. Phys. C: Solid State Phys. **21**, 2917 (1988).
[11] B. X. Yang, T. R. Thurston, J. M. Tranquada, and G. Shirane, Phys. Rev. B **39**, 4343 (1989).
[12] B. X. Yang, J. M. Tranquada, and G. Shirane, Phys. Rev. B **38**, 174 (1988).
[13] T. Chattopadhyay, G. J. McIntyre, P. J. Brown, and J. B. Forsyth, Physica C **170**, 371 (1990).
[14] A. Junod, D. Eckert, G. Triscone *et al.*, J. Phys.: Condens. Matter **1**, 8021 (1989).
[15] J. W. Loram, K. A. Mirza, C. P. Joyce, and A. J. Osborne, Europhys. Lett. **8**, 263 (1989).
[16] Yu. V. Levinskiĭ, *Phase Diagrams of Binary Metallic Systems, Vol. 1* [in Russian], Metallurgiya, Moscow (1990).
[17] A. S. Borovik-Romanov, ''Antiferromagnetism,'' in *Antiferromagnetism and Ferrites. Results of Science* [in Russian], Moscow (1962), p. 7.
[18] R. L. Carlin, *Magnetochemistry*, Springer-Verlag, Berlin–New York (1986) [Russ. transl., Mir, Moscow (1989)].
[19] B. Roden, E. Braun, and A. Freimuth, Solid State Commun. **64**, 1051 (1987).
[20] J. C. Bonner and M. E. Fisher, Phys. Rev. **135**, A640 (1964).
[21] N. A. Viglin, S. V. Naumov, and A. A. Samokhvalov, Fiz. Tverd. Tela (St. Petersburg) **38**, 1277 (1996) [Phys. Solid State **38**, 706 (1996)].
[22] L. Y. de Jongh and A. R. Miedema, Adv. Phys. **23**, 1 (1974).
[23] P. M. Duxbury, J. Oitmaa, M. N. Barber *et al.*, Phys. Rev. B **24**, 5149 (1981).
[24] M. Albrecht and F. Mila, Europhys. Lett. **34**, 145 (1996).
[25] M. Sohma and K. Kawaguchi, J. Appl. Phys. **77**, 1189 (1995).
[26] Y. Endoh, K. Yamada, R. J. Birgeneau *et al.*, Phys. Rev. B **37**, 7443 (1988).

Translated by P. Shelnitz

Light-induced evaporation and condensation growth of aerosol particles

V. G. Chernyak*) and O. V. Klitenik

A. M. Gor'kiĭ Ural State University, 620083 Ekaterinburg, Russia
(Submitted 22 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 1036–1047 (March 1998)

We analyze the possibility of resonant optical radiation inducing evaporation and condensation growth of aerosol particles suspended in a vapor–gas mixture. The molecules of the vapor absorb radiation selectively as to velocity. We examine the Knudsen regime, in which the particle radius is much smaller than the mean free path of the molecules. © *1998 American Institute of Physics.* [S1063-7761(98)02003-4]

1. INTRODUCTION

Studies of phenomena related to the effect of optical radiation on submicrometer particles are of fundamental importance in astrophysics and are of interest in connection with laser monitoring of the atmosphere. In particular, the evaporation of aerosols in the field of optical radiation is one of the main processes determining the time it takes an aerodispersion system to clear up.

Suppose that a fine macroscopic particle is suspended in a mixture consisting of its own vapor and a noncondensing gas. A system is in phase equilibrium, of course, if the chemical potentials, temperatures, and pressures of the components of the gaseous phase and the particle are the same.

When the particle is irradiated by optical radiation, it heats up, thus destroying the phase equilibrium, and begins to evaporate. From the standpoint of phenomenological thermodynamics, breakdown of the phase equilibrium of a system in a field of optical radiation is only possible if the particle and vapor–gas medium absorb light differently. But can there be evaporation and all the more condensation growth of a particle if the system does not absorb radiation, or if such absorption is a peripheral factor rather than a decisive one?

Microscopic analysis suggests at least three reasons why such evaporation and condensation growth are possible.

Suppose that the radiation is a traveling light wave whose frequency is close to that of an electronic or vibrational–rotational transition of a vapor molecule. Because of the Doppler effect, the radiation is absorbed selectively as to velocity. The effective resonant velocity interval $\Delta\mathbf{v}$ is determined by the condition that $\mathbf{k}\cdot\Delta\mathbf{v}\sim\Gamma$, where $\mathbf{k}$ is the wave vector, and Γ is the homogeneous halfwidth of the absorption line. The molecules that absorb radiation become excited. As a result, the velocity distribution of the vapor molecules in the interval $\Delta\mathbf{v}$ has a Bennett dip, and the corresponding molecules are in the excited state, forming a Bennett peak.[1] The position of the Bennett peak and dip is determined by the sign and value of the offset Ω of the radiation frequency from the center of the absorption line.

If excited and unexcited vapor molecules interact differently with the molecules of the buffer (nonabsorbing and noncondensing) gas, the rates of decay of the Bennett peak and of smoothing-out of the Bennett dip are different. As a result, the overall velocity distribution function for the vapor molecules is non-Maxwellian. At nonzero offsets Ω, this leads to light-induced drift of the vapor[2] and the aerosol particle.[3] Another consequence, directly related to the problem being discussed, is a change in vapor temperature that is not, however, accompanied by a change in the temperature of the vapor–gas mixture.[4] Depending on the size of Ω, the temperature of the vapor can be higher or lower than the equilibrium temperature of the system. In the first case the aerosol particle evaporates, and in the second it grows as a result of condensation of the cooled vapor.

Another possible reason for the breakdown of phase equilibrium is the dependence of the collision rate on the quantum state of the molecules. To be more specific, let us examine the case in which the collision cross section for the molecules that have absorbed radiation grows. This leads to an increase in the number of vapor molecules that started off from the surface of the particle, but as a result of the first collisions in the Knudsen layer changed velocity and ended up on the interphase boundary. In other words, there is partial screening of the boundary layer and, as a result, the number of vapor molecules colliding with the particle surface per unit time and having a definite condensation probability increases.

The number of molecules that evaporate from the surface per unit time depends only on the type of substance, temperature, and curvature of the surface. As a result, the dynamic equilibrium between evaporation and condensation is violated in favor of the latter, i.e., the aerosol particle grows. On the other hand, an increase in the transport cross section of the molecules due to absorption of radiation reduces the probability of these molecules being shifted from the volume of the gas to the surface of the particle. This violates dynamic equilibrium in favor of evaporation. Thus, variation in the collision rate of the excited molecules generates two competing factors: one stimulating the condensation growth of the aerosol particle and the other stimulating evaporation. It is probably impossible to predict a priori in which direction the process will proceed. Establishing the conditions in which evaporation or growth occurs constitutes the goal of the theory.

Another reason phase equilibrium might be destroyed

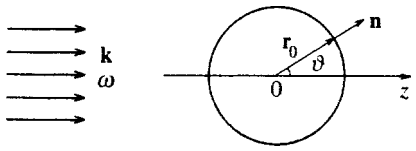

FIG. 1. Problem geometry.

could be the dependence of the nature of the interaction of vapor molecules and the surface of the particle on the quantum state of the molecules. If the condensation coefficient of the molecules that have absorbed radiation increases, condensation processes are dominant and the particle grows. In the opposite case, in which the probability of condensation of the excited molecules is lower than that of the unexcited molecules, the particle evaporates.

The kinetic theory of radiative evaporation of an aerosol particle absorbing radiation was developed in Ref. 5. In the present paper we develop a theoretical model that describes the evaporation and condensation growth of a single nonabsorbing particle suspended in a vapor–gas mixture, in which the vapor molecules are excited by resonant optical radiation selectively as to velocity. We examine the Knudsen regime, in which the mean free path of the molecules is much greater than the size of the particle.

2. STATEMENT OF THE PROBLEM

Let us examine a spherical aerosol particle suspended in a mixture consisting of the vapor of this particle and a noncondensing gas. The radius r_0 of the particle is much greater than the radius of a critical nucleation center. The system is in thermodynamic equilibrium at temperature T_0.

We irradiate the system with monochromatic light, a traveling light wave propagating along the z axis (Fig. 1). Let the radiation be absorbed by the vapor molecules in an electronic or vibrational–rotational transition from the ground state n to the first excited state m. The frequency ω of the radiation is slightly offset from ω_{mn}, the center frequency of the absorption line, i.e., $\Omega=\omega=\omega_{mn}$ ($|\Omega|\ll\omega,\omega_{mm}$). Molecules that have absorbed radiation become excited. This alters their transport characteristics and their probability of being captured by the interphase plane. Thus, the gaseous phase can be interpreted as a three-component mixture consisting of the excited and unexcited components of the vapor and the buffer (nonabsorbing and noncondensing) gas. The components of the vapor consist of molecules of equal mass m_1 but different diameters $d_n \neq d_m$. As a result of stimulated transitions and the radiative decay of the excited level, these components are constantly swapping molecules.

Resonant interaction between light and vapor destroys the state of phase equilibrium, and evaporation or condensation growth of the aerosol begins. Obviously, the decrease (increase) in particle mass per unit time, i.e., the evaporation (condensation) rate, is determined by the mass vapor flux at the particle's surface:

$$\frac{d}{dt}\left(\frac{4}{3}\pi r_0^3\rho\right)=-\langle J\rangle, \tag{1}$$

where ρ is the density of the particle, J is the radial component of the mass vapor flux at the particle surface, and angle brackets $\langle\cdots\rangle$ always denote integration over the surface S of the particle (Fig. 1):

$$\langle J\rangle=\int_{(S)} J\ dS=2\pi r_0^2\int_0^\pi J(\theta)\sin\ \theta\ d\theta. \tag{2}$$

If the velocity distribution functions f_m and f_n for the excited and unexcited vapor molecules are known, the flux J can be calculated

$$J=m_1\int v_r(f_m+f_n)d\mathbf{v}. \tag{3}$$

When the absorbing molecules are approximated by a two-level model, the velocity distribution functions for the vapor molecules, f_m and f_n, and the molecules of the buffer gas, f_2, satisfy the following system of kinetic equations[1,4]:

$$\frac{\partial f_m}{\partial t}+\mathbf{v}\cdot\nabla f_m=\frac{\chi(\mathbf{v})\Gamma_m(f_n-f_m)}{2}-\Gamma_m f_m+S_m, \tag{4}$$

$$\frac{\partial f_n}{\partial t}+\mathbf{v}\cdot\nabla f_n=-\frac{\chi(\mathbf{v})\Gamma_m(f_n-f_m)}{2}+\Gamma_m f_m+S_n, \tag{5}$$

$$\frac{\partial f_2}{\partial t}+\mathbf{v}\cdot\nabla f_2=S_2. \tag{6}$$

Here

$$S_i=S_{im}+S_{in}+S_{i2},\quad i=m,n,2,$$

$$\chi(\mathbf{v})=\frac{4|G_{mn}|^2\Gamma}{\Gamma_m[\Gamma^2+(\Omega-\mathbf{k}\cdot\mathbf{v})^2]},\quad G_{mn}=\frac{E_0 d_{mn}}{2\hbar}, \tag{7}$$

where Γ is the homogeneous halfwidth of the absorption line, Γ_m is the rate of radiative decay of the excited level, $\chi(\mathbf{v})$ is the probability of absorption per unit time for molecules with a given velocity $\mathbf{v}$ (the absorption rate for such molecules), d_{mn} is the dipole matrix element of the $m\to n$ transition, E_0 is the amplitude of the electric field of the light wave, S_{ij} are Boltzmann collision integrals, and $\hbar$ is Planck's constant.

To specify the boundary conditions for Eqs. (4)–(6) we must fix the distribution functions $f_i^+(\mathbf{r}_0,\mathbf{v})$ ($i=m,n,2$) for molecules emitted from the particle surface. Generally, the temperature T_s of the surface differs from the equilibrium temperature T_0, a fact that can be explained either by the absorption of light by the particle or by removal of the latent heat of evaporation.[5] Let us assume that the fraction α_i of vapor molecules in the ith state is evaporated with a Maxwellian distribution f_i^s corresponding to the surface temperature T_s and number density n_{is}, while the fraction $(1-\alpha_i)$ is elastically reflected without experiencing condensation in the ith state with a distribution function f_i^r. We also assume that vapor molecules reflected in the ith state experience complete energy accommodation in collisions with the interphase surface, and have a Maxwellian distribution function f_i^r corresponding to temperature T_s and number density n_{ir}. Ignoring inelastic collisions with the surface, for vapor molecules we have

$$f_i^+(\mathbf{r}_0,\mathbf{v})=\alpha_i f_i^s(\mathbf{v})+(1-\alpha_i)f_i^r(\mathbf{v}),$$

$$i=m,n,\quad \mathbf{v}\cdot\mathbf{n}>0,$$

$$f_i^{s,r}=n_{is,r}\left(\frac{m_1}{2\pi k_B T_s}\right)^{3/2}\exp\left(-\frac{m_1 v^2}{2k_B T_s}\right), \tag{8}$$

where m_1 is the mass of a vapor molecule, and k_B is Boltzmann's constant.

We assume that molecules of the buffer gas are diffusely scattered by the surface, with their energy being completely accommodated and their velocity distribution function being Maxwellian:

$$f_2^+(\mathbf{r}_0,\mathbf{v})=n_{2r}\left(\frac{m_2}{2\pi k_B T_s}\right)^{3/2}\exp\left(-\frac{m_2 v^2}{2k_B T_s}\right),\quad \mathbf{v}\cdot\mathbf{n}>0, \tag{9}$$

where n_{2r} is the number density of the buffer gas molecules scattered by the surface, m_2 is the mass of a buffer gas molecule, and $\mathbf{n}=\mathbf{r}_0/r_0$ is the outward-directed unit normal (Fig. 1).

We assume that light is absorbed only by vapor molecules with velocities in the range $\mathbf{k}\cdot\Delta\mathbf{v}\sim\Gamma$, and that it does not interact with molecules of the condensed phase. In this case only unexcited molecules evaporate from the surface. Then in the boundary condition (8) we must put $n_{ms}=0$, while n_{ns} is the equilibrium density n_s of saturated vapor corresponding to surface temperature T_s. With allowance for the curvature of the particle surface, Thomson's formula[6] yields

$$n_s=n_{s0}\exp\left(\frac{2m_1\sigma}{\rho k_B T_s r_0}\right), \tag{10}$$

where n_{s0} is the density of saturated vapor above a flat surface at temperature T_s, and σ is the surface tension.

In (8), the unknown number densities n_{ir} of the excited ($i=m$) and unexcited ($i=n$) vapor molecules reflected from the surface of the aerosol particle can be found from the balance relations for the mass vapor fluxes at the interphase boundary:

$$N_i^+=(1-\alpha_i)|N_i^-|+\alpha_i N_i^s,\quad i=m,n. \tag{11}$$

Here N_i^+, N_i^-, and N_i^s are the numbers of vapor molecules in the ith state emitted by, incident upon, and evaporated from unit surface area per unit time:

$$N_i^{+,s}=\int_{\mathbf{n}\cdot\mathbf{v}>0} v_r f_i^{+,s}(\mathbf{r}_0,\mathbf{v})d\mathbf{v},$$

$$|N_i^-|=\int_{\mathbf{n}\cdot\mathbf{v}<0}|v_r| f_i^-(\mathbf{r}_0,\mathbf{v})d\mathbf{v},\quad N_m^s=0. \tag{12}$$

The distribution functions f_i^+ and f_i^s of the of the emitted and evaporated vapor molecules are given by (8), and that of the incident molecules, f_i^-, can be found by solving the kinetic equations (4) and (5).

In the boundary condition (9), the unknown number density n_{2r} of reflected buffer-gas molecules can be found from the nonpercolation condition

$$N_2^+=|N_2^-|. \tag{13}$$

The mass vapor flux at the particle surface is

$$J=m_1\sum_{i=m,n}(N_i^+-|N_i^-|) \tag{14}$$

or, if we take (11) and (12) into account,

$$J=m_1\sum_{i=m,n}\alpha_i(N_i^s-|N_i^-|)$$

$$=\alpha_n m_1 N_n^s-m_1\sum_{i=m,n}\alpha_i|N_i^-|. \tag{15}$$

We now examine quasistationary evaporation (growth), in which the rate at which the particle radius varies, $\dot{r}_0$, is low compared to the vapor velocity U_1. Equations (1), (2), and (15) imply that $\dot{r}_0/U_1\sim\rho_1/\rho$, where ρ_1 and ρ are the mass densities of vapor and particle. Thus, far from the critical point the quasistationary approximation is valid. In the kinetic equations (4)–(6) we can then ignore the time derivatives of the distribution function.

We limit ourselves to the analysis of the free-molecule evaporation (growth) regime, in which the mean free path of molecules is much greater than the particles' radius. In this case we can ignore the perturbation of the distribution function for the molecules incident upon a particle due to collisions with emitted particles. If we assume that the intensity of the light is independent of position and the transverse size of the light beam is much greater than the mean free path of the molecules, the distribution functions produced by intermolecular collisions, i.e., far from the particle, are also spatially homogeneous. Then in Eqs. (4)–(6), which the distribution functions f_i^- of the incident molecules must satisfy, we can ignore spatial derivatives.

We examine the case of small values of $\chi(\mathbf{v})$, which is common for vibrational–rotational transitions, while for electronic transitions this case is realized at low light intensities I. What is important is that for $\chi(\mathbf{v})\ll 1$ the state of the system is close to equilibrium, so that the molecular velocity distribution functions can be written as perturbed Maxwellian distributions:

$$f_i(\mathbf{v})=f_{i0}[1+h_i(\mathbf{v})],$$

$$f_{i0}=n_i\left(\frac{m_i}{2\pi k_B T_0}\right)^{3/2}\exp\left(-\frac{m_i v^2}{2k_B T_0}\right), \tag{16}$$

where n_i is the number density of molecules of species i, and T_0 is the equilibrium temperature of the vapor–gas mixture.

The component temperatures differ little from the equilibrium temperature T_0:

$$T_i=T_0(1+\tau_i),\quad i=m,n,2, \tag{17}$$

where τ_i is the unknown perturbation of the ith component.

Under these assumptions, the linearized variants of Eqs. (4)–(6) assume the form

$$\frac{1}{2}\Gamma_m\chi(\mathbf{v})\left(\frac{n_n}{n_m}-1\right)-\Gamma_m(1+h_m)+L_{mm}+L_{mn}+L_{m2}=0, \tag{18}$$

$$-\frac{1}{2}\Gamma_m\chi(\mathbf{v})\left(1-\frac{n_m}{n_n}\right)+\frac{n_m}{n_n}\Gamma_m(1+h_m)+L_{nm}+L_{nn}+L_{n2}=0, \tag{19}$$

$$L_{2m}+L_{2n}+L_{22}=0. \tag{20}$$

We consider only elastic collisions between molecules, and for the linearized collision integrals L_{ij} we make approximations that ensure accurate values for the first thirteen moments of L_{ij} (see Ref. 7). Obviously, in an approximation linear in $\chi(\mathbf{v})$ we can assume that $n_m \ll n_n$. Then the solution of Eqs. (18)–(20) takes the form

$$h_m=\frac{\Gamma_m}{\Gamma_m+\gamma_m}\left(\frac{1}{2}\frac{n_n}{n_m}\chi(\mathbf{v})-1+\frac{A_m}{\Gamma_m}\right), \tag{21}$$

$$h_n=\frac{\Gamma_m}{\Gamma_m+\gamma_m}\frac{\gamma_m}{\gamma_n}\left(-\frac{1}{2}\chi(\mathbf{v})+\frac{n_m}{n_n}\right)+\frac{\Gamma_m}{\Gamma_m+\gamma_m}\frac{n_m}{n_n}\frac{A_m}{\gamma_n}+\frac{A_n}{\gamma_n}, \tag{22}$$

$$h_2=\frac{A_2}{\gamma_2}. \tag{23}$$

Here we have introduced the notation

$$A_i=\sum_{j=m,n,2}A_{ij},\quad \mathbf{c}_i=\frac{\mathbf{v}}{\overline{v}_i},\quad \overline{v}_i=\left(\frac{2k_BT_0}{m_i}\right)^{1/2},$$

$$A_{ij}=\left(c_i^2-\frac{3}{2}\right)\left[\gamma_{ij}\tau_i^\infty-2\frac{m_{ij}}{m_j}(\tau_i^\infty-\tau_j^\infty)\nu_{ij}^{(1)}\right],$$

$$m_{ij}=\frac{m_im_j}{m_i+m_j},\quad \gamma_i=\gamma_{im}+\gamma_{in}+\gamma_{i2},\quad i=m,n,2. \tag{24}$$

The expression for the rate $\nu_{ij}^{(1)}$ is given in Appendix A.

The number density and temperature perturbation of the ith component are defined to be

$$n_i=\int f_i\,d\mathbf{v},$$

$$\tau_i^\infty=\frac{T_i-T_0}{T_0}=\pi^{-3/2}\int\left(\frac{2}{3}c_i^2-1\right)h_i\exp(-c_i^2)d\mathbf{c}_i. \tag{25}$$

Note that generally the expression for A_{ij} in (24) incorporates terms containing the macroscopic velocities of the vapor and the buffer gas, partial heat fluxes, and components of the stress tensor. These terms contribute only to the local vapor flux $J(\theta)$ and do not alter $\langle J\rangle$, the flux averaged over the surface. The point here is that the light-induced fluxes of matter and energy, which are directed along the z axis (Fig. 1), provide a contribution to $J(\theta)$ proportional to $\cos\theta$, and the components of the stress tensor provide a contribution proportional to $(3\cos^2\theta-1)$. The surface integral (2) of these terms vanishes. This situation is typical of axisymmetric problems of gas kinetics in an approximation linear in the Mach number. The intensity of evaporation of a particle enveloped in a slow flow of the vapor–gas mixture will be the same as in a medium at rest. In our case this means that the rate of evaporation or condensation growth of a particle can be considered independently of light-induced drift[2] and photophoresis.[3]

3. RADIAL VAPOR FLUX $\langle J\rangle$

In an approximation linear in $\chi(\mathbf{v})$ we have $n_m \ll n_n$, i.e., $n_n \approx n_1$. Then Eqs. (16), (21), and (25) yield

$$n_m=\frac{n_1}{2\pi^{3/2}}\chi_0,\quad \chi_0=\int\chi(\mathbf{v})\exp(-c_1^2)d\mathbf{c}_1,$$

$$\tau_m^\infty=\frac{1}{3\pi^{3/2}}\frac{n_1}{n_m}\frac{\Gamma_m\chi_2}{\Gamma_m+\nu_{mn}^{(1)}+2\nu_{m2}^{(1)}m_{12}/m_2},$$

$$\chi_2=\int\left(c_1^2-\frac{3}{2}\right)\chi(\mathbf{v})\exp(-c_1^2)d\mathbf{c}_1. \tag{26}$$

Both χ_0 and χ_2, which depend on the parameters $x=\Omega/k\overline{v}_1$ and $y=\Gamma/k\overline{v}_1$, have a simple form for both inhomogeneous ($y\ll 1$) and homogeneous ($y\gg 1$) line broadening. For intermediate values of y, a numerical calculation is carried out in Appendix B.

The relative light-induced cooling (heating) of the vapor and heating (cooling) of the buffer gas are, respectively,

$$\tau_1^\infty=\frac{n_m\tau_m^\infty+n_n\tau_n^\infty}{n_1}=\frac{n_2}{n_1+n_2}\frac{n_m}{n_1}\tau_m^\infty\left(1-\frac{\Omega_{2m}^{(1,1)}}{\Omega_{2n}^{(1,1)}}\right),$$

$$n_2\tau_2^\infty=-n_1\tau_1^\infty. \tag{27}$$

Here the $\Omega_{ij}^{(l,r)}$ are the Chapman–Cowling integrals,[8] which depend on the model of the pairwise interaction potential. An expression for molecules approximated by the hard-sphere model is given in Appendix A.

The mean mass vapor flux at the particle surface is

$$\langle J\rangle=\frac{\rho_1\overline{v}_1}{2\pi^{1/2}}\left(G_s\Delta\alpha+G_c\frac{\Delta\gamma}{\gamma_n}+\frac{\alpha_n}{2}\tau_1^\infty\right). \tag{28}$$

Here

$$G_s=\frac{\gamma_m}{\Gamma_m+\gamma_m}\frac{\chi_0}{2\pi^{3/2}}+\frac{\Gamma_m}{\Gamma_m+\gamma_m}\times\left(\frac{\chi_1^-}{\pi}+\frac{1}{2}\frac{n_m}{n_1}\frac{\nu_{m2}^{(1)}}{\Gamma_m}\frac{m_1-m_2}{m_1+m_2}\tau_m^\infty\right), \tag{29}$$

$$G_c=\frac{\alpha_n\Gamma_m}{\Gamma_m+\gamma_m}\left(\frac{\chi_0}{2\pi^{3/2}}-\frac{\chi_1^-}{\pi}-\frac{\chi_2}{6\pi^{3/2}}\right), \tag{30}$$

$$\chi_1^-=\frac{1}{2}\int_0^\pi\sin\theta\,d\theta\int_{c_{1r}<0}|c_{1r}|\chi(\mathbf{v})\exp(-c_1^2)d\mathbf{c}_1,$$

$$\Delta\alpha=\alpha_n-\alpha_m,\quad \Delta\gamma=\gamma_n-\gamma_m. \tag{31}$$

In (28) the quantity $\chi_1^-(x,y)$ is calculated in Appendix B.

In Eq. (28), the kinetic coefficient G_s characterizes the surface mechanism of particle evaporation (growth), which is related to the difference in the evaporation–condensation coefficients for excited and unexcited vapor molecules. The kinetic coefficient G_c characterizes the contribution related to the difference in rates of collisions of excited and unex-

cited vapor molecules. Finally, the third term in (28) describes the rate of condensation growth (evaporation) of the aerosol particle due to light-induced cooling (heating) of the vapor.

If we choose effective rates of collisions of the $i-j$ in the form $\gamma_{ij}=\nu_{ij}^{(1)}$ and assume that the interacting molecules are hard elastic spheres and that the effective interaction radii of excited and unexcited vapor molecules with the buffer-gas molecules are commensurate, i.e. $(d_{n2}-d_{m2})/d_{n2}\ll 1$, the expression (28) for the mean mass flux of the vapor becomes

$$\langle J\rangle=\frac{\rho_1\overline{v_1}}{2\pi^{1/2}}\left(G_s\Delta\alpha+G_d\frac{\Delta d}{d_{n2}}\right),\quad \Delta d=d_{n2}-d_{m2}. \tag{32}$$

Here

$$G_d=\alpha_n\frac{n_m}{n_1}\frac{n_2}{n_1+n_2}\tau_m^{\infty}+2BG_c,$$

$$B=\frac{(m_1/2)^{1/2}n_1d_nd_{n2}+m_{12}^{1/2}n_2d_{n2}^2}{(m_1/2)^{1/2}n_1d_n^2+m_{12}^{1/2}n_2d_{n2}^2}. \tag{33}$$

The second term in parentheses in (32) characterizes the bulk evaporation (growth) mechanism of the particle, a mechanism related to the transport properties of excited and unexcited vapor molecules.

The expressions for the kinetic coefficients G_s and G_d become much simpler for inhomogeneous and homogeneous line broadening

3.1. Inhomogeneous broadening

Such broadening $(y\ll 1)$ is characteristic of tenuous gases. For $x<1$, i.e., $\Omega<k\overline{v_1}$, we obtain

$$G_s=\frac{2\pi^{1/2}G_{mn}^2}{3\Gamma_mk\overline{v_1}}\left(\frac{2\Gamma_m+3\gamma_m}{\Gamma_m+\gamma_m}+\frac{4\Gamma_m}{\Gamma_m+\gamma_m}\frac{m_2-m_1}{m_2+m_1}\frac{\nu_{m2}^{(1)}}{\Gamma_m+\nu_{mn}^{(1)}+2\nu_{m2}^{(1)}m_{12}/m_2}\right), \tag{34}$$

$$G_d=\frac{2\pi^{1/2}\alpha_nG_{mn}^2}{\Gamma_mk\overline{v_1}}\left(2B\frac{\Gamma_m}{\Gamma_m+\gamma_m}+\frac{1}{3}\frac{n_2}{n_1+n_2}\frac{\Gamma_m}{\Gamma_m+\nu_{mn}^{(1)}+2\nu_{m2}^{(1)}m_{12}/m_2}\right). \tag{35}$$

3.2. Homogeneous broadening

In this case $y\gg 1$ and, to within terms of order y^{-1}, we have

$$G_s=\frac{2G_{mn}^2}{\Gamma_m\Gamma},\quad G_d=0. \tag{36}$$

Thus, when homogeneous broadening occurs, we can ignore the contribution of the bulk evaporation (growth) mechanism.

4. DISCUSSION

The kinetic coefficients G_s and G_d, which characterize the surface and bulk evaporation (condensation growth) mechanisms of an aerosol particle, are proportional to the radiation intensity and depend on the ratio of the rate of radiative decay of an excited level, Γ_m, to the intermolecular collision rate γ_m. The value of Γ_m/γ_m, which depends on gas pressure, has a stronger effect on the bulk component of the radial vapor flux.

The frequency offset Ω from the center of the absorption line determines the light-induced cooling or heating of the vapor, i.e., the sign of relative variation of the vapor temperature τ_1^{∞}. From (29) we see that the sign of τ_1^{∞} is determined by the signs of Δd and χ_2. Suppose that the effective diameter of molecules that have absorbed light increases $(\Delta d<0)$. Then as Fig. 4b shows, for small values of Ω the vapor heats up, while for large values of Ω it cools off. Thus, there exists an inversion value $\Omega_{\rm inv}$ that depends on the parameter $y=\Gamma/k\overline{v_1}$. For inhomogeneous broadening $(y\ll 1)$, $\Omega_{\rm inv}=7.5\Gamma$, while for $y=2$ we have $\Omega_{\rm inv}\approx 0.75\Gamma$. When $\Delta d<0$, the third term in parentheses on the right-hand side of the Eq. (28) for the radial vapor flux describes evaporation of a particle if $\Omega<\Omega_{\rm inv}$ and condensation growth if $\Omega>\Omega_{\rm inv}$.

The rate of particle evaporation or growth is strongly affected by Ω. In contrast to the phenomenon of light-induced drift,[2] the direction of the radial vapor flux is independent of the sign of Ω, i.e., the kinetic coefficients G_s and G_d are even functions of Ω. Here the evaporation (growth) rate peaks at exact resonance, $\Omega=0$. Figures 2 and 3 depict the dependence of

$$G_s^*=G_s\left(\frac{4|G_{mn}|^2}{\Gamma_mk\overline{v_1}}\right)^{-1},$$

$$G_d^*=G_d\left(\alpha_n\frac{\Gamma_m}{\Gamma_m+\gamma_m}\frac{4|G_{mn}|^2}{\Gamma_mk\overline{v_1}}\right)^{-1}$$

on the offset parameter x for inhomogeneous broadening $(y=0.1)$ with $m_1\approx m_2$. At other values of the parameters both G_s^* and G_d^* vary, but qualitatively the dependence on x remains the same. The curve representing $G_s(x)$ (Fig. 2) is a Lorentzian, and the kinetic coefficient G_s is always positive. The curve representing $G_d(x)$ is plotted in Fig. 3. We see

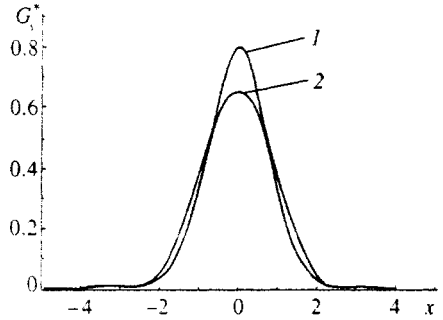

FIG. 2. Dependence of the kinetic coefficient G_s on the offset parameter x at $y=0.1$ and $m_1\approx m_2$. Curve *1* corresponds to $\gamma_m\gg\Gamma_m$, and curve *2* to $\gamma_m\ll\Gamma_m$.

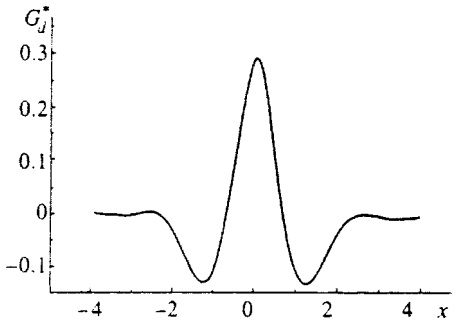

FIG. 3. Dependence of the kinetic coefficient G_d on the offset parameter x at $y=0.1$, $m_1 \approx m_2$, and $n_1 \ll n_2$.

that at $\Omega=\Omega_{\text{inv}}$, G_d changes sign. The inversion value of the offset, Ω_{inv}, which depends on the broadening parameter y, coincides with the value that determines the sign of the relative variation of the temperature, τ_1^{∞}.

When $\Omega<\Omega_{\text{inv}}$, both G_s and G_d are positive. Hence for $\Delta\alpha>0$ and $\Delta d>0$ the particle evaporates, and at $\Delta\alpha<0$ and $\Delta d<0$ it grows. If $\Delta\alpha$ and Δd have opposite signs, the surface and volume bulk of the vapor flux $\langle J\rangle$ compete, and there is no way to predict the direction of the phase transition in the general case. To make such a prediction, we must have numerical estimates of the coefficients G_s and G_d for specific systems and specific physical conditions. When $\Omega>\Omega_{\text{inv}}$, the particle grows if $\Delta\alpha<0$ and $\Delta d>0$ and evaporates if $\Delta\alpha>0$ and $\Delta d<0$.

We now make some numerical estimates for a specific system under typical experimental conditions. We take a sodium particle whose diameter r_0 is 1 μm and immerse it in a mixture of sodium vapor and the inert gas argon at temperature T_0 equal to 371 K. The saturated vapor pressure is 1.5×10^{-5} Pa. The effect of surface curvature can be neglected. The molecular characteristics are: $m_1=3.82\times10^{-26}$ kg, $m_2=6.64\times10^{-26}$ kg, $d_1=3.0$ Å, and $d_2=3.4$ Å. With the pressure of the buffer gas (argon) being about 1300 Pa, the Knudsen number Kn is approximately 10, which corresponds to the free-molecule regime.

The source of light is a tunable dye laser emitting in the vicinity in the D_1 and D_2 lines of sodium (wave $\lambda\sim600$ nm). The radiated power is about 10 mW and the beam diameter is about 1 mm. At this intensity the Rabi frequency G_{mn} is of order 10^8 Hz. The radiative decay rate $\Gamma_m=6\times10^7$ Hz (see Ref. 9). Finally, the parameter $\Gamma/k\overline{v_1}\approx0.01$ corresponds to inhomogeneous line broadening.

Numerical estimates of the kinetic coefficients yield

$$G_s\approx1.10, \quad G_d\approx1.22.$$

If we assume that $\Delta\alpha=0.01$ and $\Delta d/d_{n2}=0.01$, the particle loses approximately 5×10^{-19} g per second.

In conclusion we note that experimental studies at a variety of buffer-gas pressures and values of Ω might become a source of accurate information about evaporation–condensation coefficients and transport characteristics of excited molecules.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-01-00756).

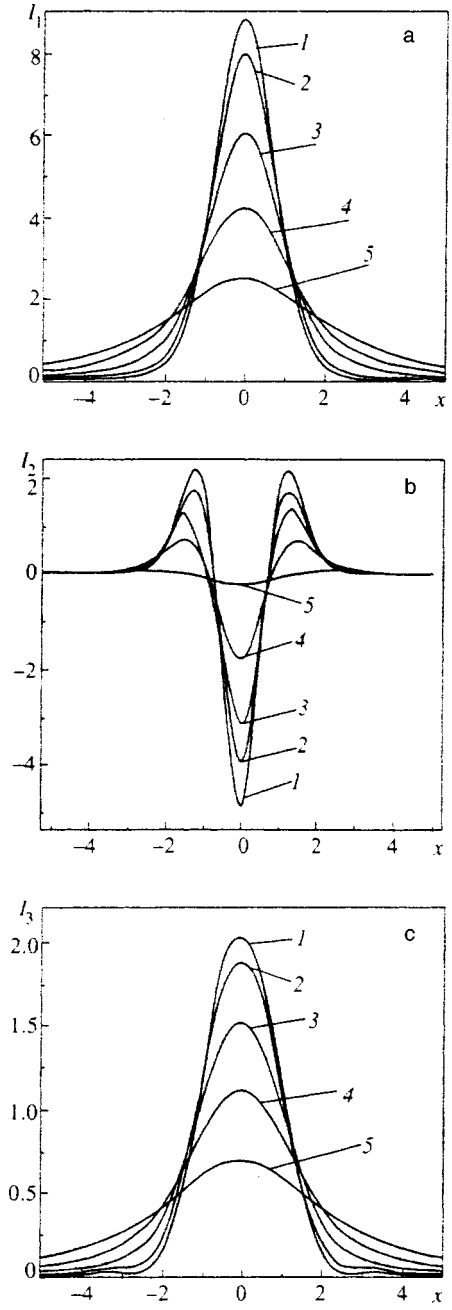

FIG. 4. The x-dependence of I_1 (a), I_2 (b), and I_3 (c) for various values of $y=0.1$ (curves *1*), 0.2 (curves *2*), 0.5 (curves *3*), 1 (curves *4*), and 2 (curves *5*).

APPENDIX A

In Eq. (24), the expression for the rate $\nu_{ij}^{(1)}$ taken from Ref. 7 has the form

$$\nu_{ij}^{(1)}=\frac{16}{3}\frac{m_{ij}}{m_i}n_j\Omega_{ij}^{(1,1)}.$$

Here the $\Omega_{ij}^{(l,r)}$ are the Chapman–Cowling integrals,[8] which depend on the shape of the pair interaction potential between molecules of species i and j. For hard spheres of diameter d we have

$$\Omega_{ij}^{(l,r)}=\left(\frac{k_BT}{2\pi m_{ij}}\right)^{1/2}\frac{r+1!}{2}\left(1-\frac{1}{2}\frac{1+(-1)^l}{l+1}\right)\pi d_{ij}^2,$$

$$d_{ij}=\frac{d_i+d_j}{2}.$$

APPENDIX B

Here we find the dependence of χ_0, χ_1^-, and χ_2 on the offset and broadening parameters x and y. It is convenient to write (26) and (31) in the form

$$\chi_0=\frac{4G_{mn}^2}{\Gamma_m k\bar{v}_1}I_1(x,y), \quad \chi_2=\frac{4G_{mn}^2}{\Gamma_m k\bar{v}_1}I_2(x,y),$$

$$\chi_1^-=\frac{4G_{mn}^2}{\Gamma_m k\bar{v}_1}I_3(x,y),$$

where $x=\Omega/k\bar{v}_1$, and $y=\Gamma/k\bar{v}_1$.

For inhomogeneous broadening ($y\ll 1$) with $x\ll 1$ we have

$$I_1(x,y)=\pi^2, \quad I_2(x,y)=-\frac{\pi^2}{2}, \quad I_3(x,y)=\frac{\pi^{3/2}}{3}.$$

For homogeneous broadening ($y\gg 1$) with $x<y$ we have, to within terms of order y^{-1},

$$I_1(x,y)=\frac{\pi^{3/2}}{y}, \quad I_2(x,y)=0, \quad I_3(x,y)=\frac{\pi}{2y}.$$

The results of numerical calculations for various values of the parameters x and y are plotted in Fig. 4.

*)E-mail: vladimir.chernyak@usu.ru

[1] S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in the Spectra of Atoms and Molecules* [in Russian], Nauka, Novosibirsk (1979).
[2] F. Kh. Gel'mukhanov and A. M. Shalagin, JETP Lett. **29**, 711 (1979).
[3] V. G. Chernyak, Zh. Éksp. Teor. Fiz. **96**, 878 (1989) [Sov. Phys. JETP **69**, 498 (1989)].
[4] F. Kh. Gel'mukhanov and L. V. Il'ichev, Khim. Fiz. **3**, 1544 (1984).
[5] V. G. Chernyak, Izv. Ross. Akad. Nauk, Fiz. Atmos. Okeana **31**, 800 (1995).
[6] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, 2nd ed., Pergamon Press, Oxford (1969).
[7] F. J. McCormack, Phys. Fluids **16**, 2095 (1973).
[8] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-uniform Gases*, Cambridge Univ. Press, London (1952).
[9] A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions*, Springer-Verlag, Berlin (1985).

Translated by Eugene Yankovsky

Evolution of the ESR spectrum at the metal–insulator transition in quasi-one-dimensional systems

V. A. Atsarkin and V. V. Demidov*)

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia
(Submitted 22 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 1048–1057 (March 1998)

The evolution of ESR spectra in metals containing both conduction electrons and localized paramagnetic centers with inhomogeneous broadening of the magnetic resonance is treated theoretically. It is shown that such a spectrum can be effectively narrowed when a relaxation bottleneck is present for an arbitrary distribution of inhomogeneous broadening. The temperature dependence of the ESR spectrum in the polymeric phase of RbC_{60} is investigated experimentally in the region of the metal–insulator transition. Application of the calculations to the experimental data demonstrates good agreement with the physical model if it is assumed that the given material is a quasi-one-dimensional system. [S1063-7761(98)02103-9]

1. INTRODUCTION

The shape of the ESR line in metals containing localized paramagnetic centers in addition to their conduction electrons is usually analyzed on the basis of the Barnes–Plefke theory.[1] The role of these localized centers is played by various impurities possessing unpaired spin or by the conduction electrons, but localized for one reason or another.

For a graphic description of the kinetics of the various processes taking place in such systems, it is convenient to represent the delocalized and localized spins in the form of two subsystems, respectively e and s. Figure 1 schematically depicts the direct interaction between the subsystems with the corresponding rates V_{es} and V_{se} and their direct coupling with the lattice with rates V_{eL} and V_{sL}. The dynamics of such subsystems are typically investigated with the help of the Bloch–Hasegawa equations for the magnetizations of the localized and delocalized spins.[1,2] In the case where the exchange interaction of the subsystems with each other exceeds their partial lattice coupling ($V_{es}, V_{se} \gg V_{eL}, V_{sL}$), coupled oscillations of the spin magnetizations of the conduction electrons and the localized spins take place, i.e., the phenomenon of a relaxation bottleneck is observed.[1] In this case the system is characterized by a single Lorentzian ESR line. In the opposite limit the spectrum divides into two lines with central frequencies ω_e and ω_s characterizing each of the subsystems individually.

The evolution of the ESR spectrum in such systems has been investigated in detail in many studies (see the corresponding references in the review in Ref. 1). However, in most experiments the picture is blurred by the huge width of the resonance of the conduction electrons. Nevertheless, there is a sufficient number of objects in which the ''intrinsic'' width of the ESR line of the conduction electrons is comparable with the width of the ESR line of the localized centers. Here, first of all, we should mention quasi-one-dimensional organic conductors[3,4] and the recently discovered conducting polymer phases of the two alkali fullerides RbC_{60} and CsC_{60}.[5–8] The quasi-one-dimensionality of the last two compounds has been the growing subject of discussion, but the relatively narrow ESR line of the conduction electrons (2–6 G) provides additional confirmation of the low dimensionality of these compounds. The main interest in quasi-one-dimensional conductors is motivated by the metal–insulator transition, which in turn is accompanied by an abrupt change in the ESR spectrum.

An important problem of the ESR spectroscopy of one-dimensional systems is inhomogeneous (not coupled with the rate V_{sL}) broadening of the ESR line of the localized spins. In 1984 Tagirov and Trutnev extended the Barnes–Plefke theory to the case of an arbitrary distribution of resonant frequencies of the localized spins with distribution function $g(\Omega)$ (Ref. 9), where

$$\int_{-\infty}^{\infty} g(\Omega)\, d\Omega = 1. \tag{1}$$

They showed that if the coupling rate V_{se} (see Fig. 1) exceeds some characteristic width of inhomogeneous broadening of the resonant frequencies $\tilde{\Delta}$, these frequencies are averaged or, what is the same thing, a narrowing of the spectrum by rapid exchange between the subsystems takes place. However, this narrowing has been graphically demonstrated only for the case of finite second moments M_2 of the function $g(\Omega)$, where

$$M_2 = \int_{-\infty}^{\infty} \Omega^2 g(\Omega)\, d\Omega. \tag{2}$$

In this case the contribution of inhomogeneous broadening is decreased to the value M_2/V_{se}. In the following section we calculate the ESR spectrum over the wide range of variation of the rates V_{se}, V_{es} characteristic of the metal–insulator transition and show that an analogous inhomogeneous line narrowing effect should be observed in the more general case of an arbitrary function $g(\Omega)$ satisfying condition (1).

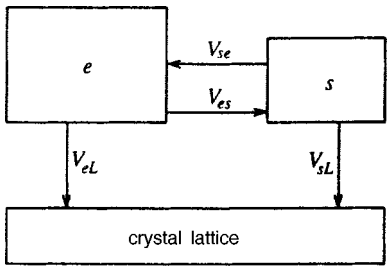

FIG. 1. Diagram of the coupling of the spin subsystems of the conduction electrons and localized spins between themselves and with the crystal lattice.

In the third section we present our experimental data on the temperature dependence of ESR spectra along with calculations to help interpret these spectra in the polymer phase of RbC_{60} in the region of the metal–insulator transition.[10]

2. CALCULATION OF THE ESR SPECTRUM AT THE METAL–INSULATOR TRANSITION

As the starting point of our ESR spectrum calculations we employ the kinetic equations for the magnetizations of the delocalized and localized spins from Ref. 9, augmented for completeness by terms allowing for indirect spin–lattice relaxation of the localized spins.

Following the recommendations of Ref. 9 explicitly, we obtain an exact solution (within the framework of the model) for the imaginary part of the dynamic susceptibility of the delocalized and localized spins in the form

$$\chi''(\omega)=\frac{PY+RX}{X^2+Y^2}. \tag{3}$$

Here we have introduced the notation

$$P\equiv\omega_s\chi_s\delta_e+\omega_e\chi_e q_s+2\omega_s\omega_e\lambda\chi_s\chi_e-b_s(\Sigma_e+B_s)-b_e(\Sigma_s+B_e),$$

$$R\equiv\omega_s\chi_s(\Sigma_e+B_s)+\omega_e\chi_e(\Sigma_s+B_e)+b_s(\delta_e+\omega_e\lambda\chi_e)+b_e(q_s+\omega_s\lambda\chi_s),$$

$$X\equiv\omega_s\omega_e\lambda^2\chi_s\chi_e+\Sigma_e\Sigma_s-B_eB_s-q_s\delta_e,$$

$$Y\equiv q_s\Sigma_e+\delta_e\Sigma_s+\lambda(\omega_e\chi_eB_e+\omega_s\chi_sB_s),$$

$$\delta_e=\omega_e\frac{\chi_e}{\chi_e^0}-\omega,\quad \delta_s=\omega_s\frac{\chi_s}{\chi_s^0}-\omega,\quad \kappa\equiv\frac{\omega_s}{\omega_e}.$$

$$q_s=\frac{A}{A^2+B^2},\quad A\equiv\int_{-\infty}^{\infty}\frac{(\Omega+\delta_s)g(\Omega)}{(\Omega+\delta_s)^2+W_s^2}\,d\Omega,$$

$$B\equiv\int_{-\infty}^{\infty}\frac{W_sg(\Omega)}{(\Omega+\delta_s)^2+W_s^2}\,d\Omega,$$

$$\Sigma_s\equiv\frac{B}{A^2+B^2}+\kappa\lambda\chi_e^0V_{es},\quad \Sigma_e\equiv W_e+\kappa^{-1}\lambda\chi_s^0V_{se},$$

$$W_e\equiv V_{es}+V_{eL},\quad W_s\equiv V_{se}+V_{sL},$$

$$B_e\equiv\kappa V_{es}+\lambda\chi_s^0W_s,\quad B_s\equiv\kappa^{-1}V_{se}+\lambda\chi_e^0W_e,$$

$$b_e\equiv\chi_e^0W_e-\kappa^{-1}\chi_s^0V_{se},\quad b_s\equiv\chi_s^0W_s-\kappa\chi_e^0V_{es},$$

ω_e,ω_s and χ_e,χ_s are the resonant frequencies and static susceptibilities for the delocalized and localized centers, and λ is a constant of the molecular field introduced to allow for the mutual influence of the e and s subsystems.[1] Note should be made of the difference between the partial susceptibilities χ_e^0 and χ_s^0 on the one hand and the total static susceptibilities χ_e and χ_s, on the other. The coupling between them is obvious from the relations for the magnetizations M_e and M_s in an external magnetic field H_0:

$$M_e=\chi_e^0(H_0+\lambda M_s),\quad M_s=\chi_s^0(H_0+\lambda M_e).$$

The kinetic coefficients

$$V_{es}=\frac{8\pi c}{3\hbar}\,S(S+1)\rho J^2, \tag{4}$$

$$V_{se}=\frac{4\pi}{\hbar}\,(\rho J)^2k_BT \tag{5}$$

are determined by the Overhauser and Korringa coupling mechanisms of the conduction electron subsystems with localized spins and vice versa.[1] In relations (4) and (5) ρ is the density of states of the conduction electrons at the Fermi level, J is the exchange integral, S and c are the spin and concentration of the localized spins, and k_B is the Boltzmann constant. The kinetic coefficients are related to the partial static susceptibilities by the principle of detailed balance:

$$V_{es}\chi_e^0=V_{se}\chi_s^0. \tag{6}$$

Mathematical analysis of expression (3) shows that under conditions in which the bottleneck effect is strong and for

$$\alpha\equiv\int_{-\infty}^{\infty}\frac{\Omega^2g(\Omega)}{(\delta_s+\Omega)^2+W_s^2}\,d\Omega\ll1 \tag{7}$$

the imaginary part of the dynamic susceptibility of the total spin system is described by a single Lorentzian line with half-width

$$\Delta_L=\frac{\chi_s^0V_{sL}+\chi_e^0V_{eL}}{\chi_s^0+\chi_e^0}+\alpha\,\frac{V_{es}V_{se}}{V_{es}+V_{se}}. \tag{8}$$

Note that if inequality (7) is satisfied it is necessary to have $V_{se}\gg\tilde{\Delta}$. It is clear from Eq. (8) that for $V_{se}\geqslant V_{es}$ narrowing of the spectrum is possible for any form of inhomogeneous broadening. For example, for $M_2/V_{se}^2\ll1$ the well-known result

$$\Delta_L=\frac{\chi_s^0V_{sL}+\chi_e^0V_{eL}}{\chi_s^0+\chi_e^0}+\frac{\chi_s^0}{\chi_s^0+\chi_e^0}\,\frac{M_2}{V_{se}}. \tag{9}$$

obtains. In the case when the second moment of $g(\Omega)$ diverges, formula (9) is inapplicable and the expression for Δ_L is determined by the specific form of $g(\Omega)$. In particular, for a Lorentzian dependence

$$g(\Omega)=\frac{\tilde{\Delta}/\pi}{\tilde{\Delta}^2+\Omega^2} \tag{10}$$

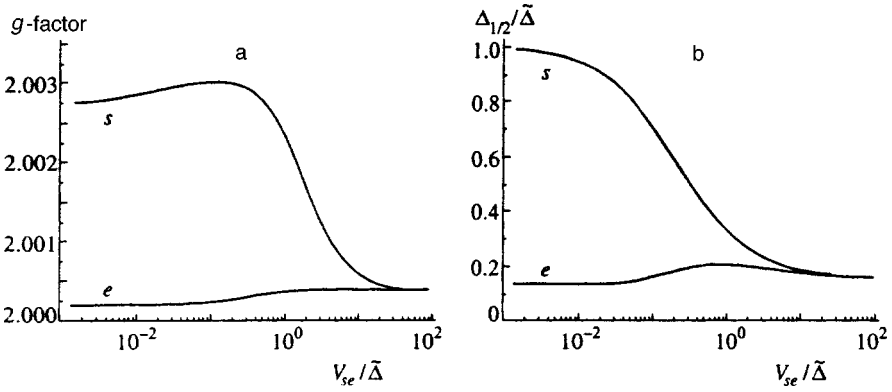

FIG. 2. Calculated values of the dependence of the g-factor (a) and the half-width (b) of the two Lorentzian components of the ESR spectrum on the parameter V_{se} in the case in which the inhomogeneous broadening is described by a Lorentzian.

inequality (7) reduces simply to $V_{se} \gg \tilde{\Delta}$ and we now have

$$\Delta_L = \frac{\chi_s^0 V_{sL} + \chi_e^0 V_{eL}}{\chi_s^0 + \chi_e^0} + \frac{\chi_s^0}{\chi_s^0 + \chi_e^0}\,\tilde{\Delta}. \tag{11}$$

Thus, in the case $\chi_s^0 < \chi_e^0$ or, what is the same thing, for $V_{se} > V_{es}$, the inhomogeneous broadening will be narrowed by rapid exchange between the subsystems.

Thus, it is clear from Eqs. (9) and (11) that under conditions of strong coupling between the s and e subsystems their total ESR spectrum can narrow abruptly. With weakening of the interaction of the s and e subsystems each of them becomes more independent, and in the limit in which the coupling is switched off completely the ESR spectrum consists of two lines: a Lorentzian line, describing the resonance of the conduction electrons, and a line with line shape $g(\Omega)$ (under the condition $\tilde{\Delta} \gg V_{sL}$), characterizing the resonance of the localized spins.

The intermediate case is of greatest interest, at least from the point of view of the metal–insulator transition. Here it is not possible to obtain an analytical formula. Therefore we carried out a numerical calculation of the frequency dependence of the imaginary part of the dynamic susceptibility $\chi''(\omega)$ of a model system. We assumed that the system contains conduction electrons with g-factor $g_e = 2.0002$ and activation behavior of the static susceptibility:

$$\chi_e(T) = \chi_e(T_0)\exp(-E_A/k_BT),$$

and localized spins with susceptibility obeying the Curie law:

$$\chi_s(T) = \chi_s(T_0)(T_0/T).$$

In this case the temperature dependence of the kinetic coupling coefficients of the subsystems was taken to have the following form:

$$V_{es}(T) = V_{es}(T_0)\exp(-E_A/k_BT),$$

$$V_{se}(T) = V_{se}(T_0)(T/T_0)\exp(-2E_A/k_BT).$$

The physical meaning of the parameters E_A, $\chi_e(T_0)$, $\chi_s(T_0)$, $V_{es}(T_0)$, and $V_{se}(T_0)$ (T_0 is some arbitrary temperature) is obvious, and their numerical values were chosen for convenience of an illustrative example. Although the above dependences are of a specifically model character, they nevertheless are close to the real situation of the metal–insulator transition. As is well known, the metal–insulator transition is accompanied by the appearance of an energy gap $E_g = 2E_A$ near the Fermi surface. As a result, the carrier concentration begins to fall abruptly as the temperature is lowered, as a rule according to an activation law with activation energy E_A. For this reason an activation law appears in the temperature dependence of V_{es}, V_{se}, and χ_e. In addition, in our model calculation we used the Lorentzian distribution function $g(\Omega)$ given by Eq. (10) in order to convincingly demonstrate the narrowing effect in this nontrivial case.

The spectrum calculated for each specific temperature was approximately described by a sum of two Lorentzians. Of course, in the presence of coupling between the e and s subsystems the total spectrum cannot be represented in the form of a sum of two independent ESR lines. However, the proposed approximation proved to be quite successful. In particular, in the example given below, in which the calculated curve was fitted by a sum of two Lorentzians, the ratio of the rms error to the peak-to-peak amplitude of the calculated curve was less than 0.2% over the entire range of variation of V_{se}. Thus, there was hardly any difference between the fits and the calculated spectra. Figure 2 plots the dependence of the g-factors and half-widths $\Delta_{1/2}$ of each of the Lorentzian components on the parameter V_{se}. All frequencies were normalized to the inhomogeneous half-width $\tilde{\Delta}$. It can be seen that for $V_{se} > 20\tilde{\Delta}$ the ESR spectrum is described by a single Lorentzian. After that it becomes complicated, but, as was noted above, it is well described by a sum of two Lorentzians over the entire range of variation of V_{se}. In this case, one Lorentzian is monotonically broadened out to the value $\tilde{\Delta}$ and its g-factor tends to the g-factor of the localized spins. At the same time, the parameters of the second curve gradually approach values corresponding to the conduction electrons. Such behavior of the separate Lorentzian components justifies associating them with the ESR lines of the s and e subsystems.

To summarize here, the numerical calculations graphically demonstrate a significant convergence of the resonant frequencies of the main body of localized spins (the effect of spectrum narrowing) due to a strong exchange interaction of the e and s subsystems. In this case, the second moment of the distribution function of the local magnetic fields in which the localized spins initially find themselves can be as large as desired.

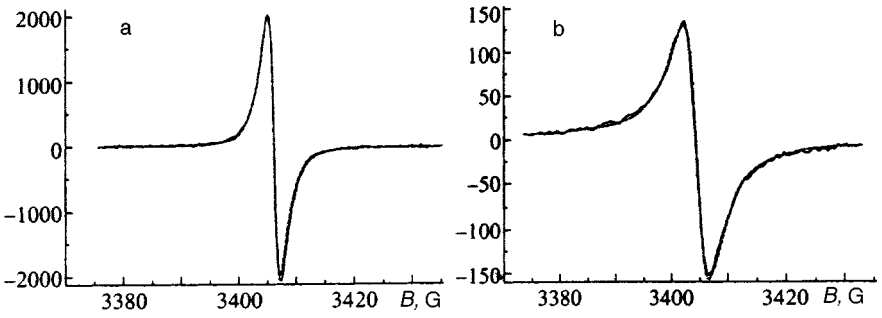

FIG. 3. Example of fitting of the experimental ESR spectra (the plotted signals are normalized to a reference level). The solid curves correspond to experiment, and the dashed curves—to calculation with the parameters given in the text, at $T=60.6$ K (a) and 30.1 K (b).

3. THE ESR SPECTRUM IN THE POLYMERIC PHASE OF RbC_{60} AT THE METAL–INSULATOR TRANSITION

We performed an experimental study of ESR spectra and various relaxation characteristics in the polymer phase of RbC_{60}[1)] in which the temperature was lowered from 300 to 4.2 K. The ESR spectra were recorded on a Bruker ER-200 spectrometer at a frequency $\omega/2\pi=9.56$ GHz. The temperature of the samples was varied with the help of an Oxford attachment to the ESR 900 cryostat (relative accuracy of temperature setting ± 0.1 K) and was additionally monitored by the ESR signal from a reference piece of ruby. The relaxation characteristics were recorded on the custom-built spectrometer described in Ref. 10. The main, general results were published in Ref. 11; here we will dwell in detail only on the evolution of the ESR spectrum in this material at the metal–insulator transition. The point here is that at high enough temperatures the ESR spectrum, usually ascribed to the conduction electrons, is ideally described by a Lorentzian. However, as the temperature is lowered below 55 K the ESR line becomes asymmetric and wide wings appear in it (Fig. 3). The total intensity I, that is, the area under the resonance curve, identified with the total spin susceptibility, falls abruptly (Fig. 4). The anomalous falloff of the susceptibility in RbC_{60} was observed earlier and interpreted as a metal–insulator transition with transition temperature $T_c=55$ K (Ref. 5). The presence of such a transition is also indicated by studies of relaxation characteristics[10] and the electrical conductivity of this material[12,13] in the same temperature interval.

We think that in addition to free carriers, localized paramagnetic centers also participate in the formation of the ESR spectrum in the polymer phase of RbC_{60}, where the ends of broken polymer chains of C_{60} molecules may play the role of such centers. Thus, the given system can be analyzed in terms of the model of two coupled spin subsystems considered in the previous section.

In the course of the analysis, we set $\lambda=0$. Test calculations with $\lambda\neq 0$ showed that the influence of this constant reduces for the most part to a mutual modification of the central resonant frequencies of both spin subsystems and, consequently, is not fundamental for the derivations that follow. As the temperature dependence of the static susceptibility of the localized spin we used the Curie law: $\chi_s(T)=\chi_s(T_c)(T_c/T)$. Here we assumed that the localized spins undergo inhomogeneous broadening with a Gaussian distribution

$$g(\Omega)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{\Omega^2}{2\sigma^2}\right), \tag{12}$$

where $\sigma=\tilde{\Delta}/\sqrt{2\ln 2}$. To calculate the static susceptibility of the delocalized spins we followed the indication of quasi-

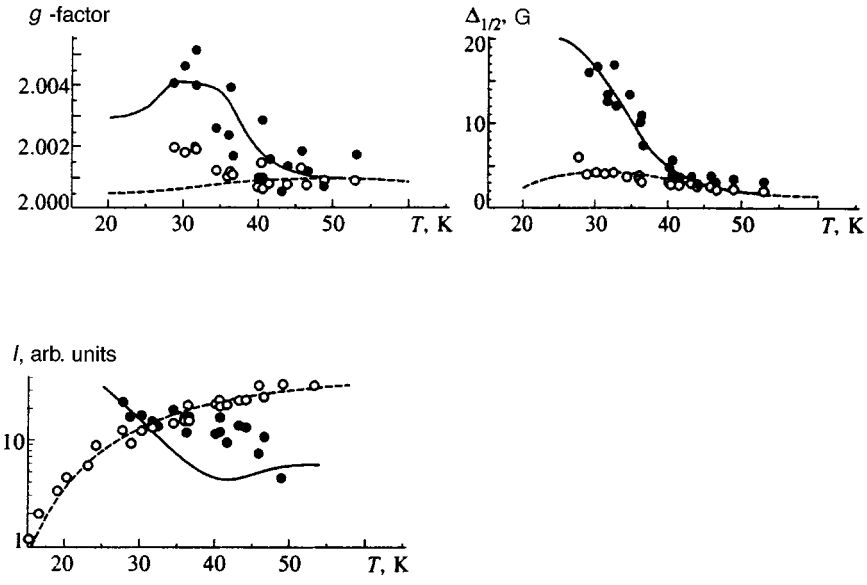

FIG. 4. Comparison of the temperature dependence of the experimental and calculated parameters of the ESR spectrum in RbC_{60}: empty and filled circles—experimental values of the parameters of the individual components corresponding to the e- and s-subsystems. Dashed curves—calculations with the parameters given in the text.

one-dimensionality of the electrical conductivity in this material stated in Ref. 12 and applied the formula obtained for the susceptibility of a quasi-one-dimensional conductor with energy gap $E_g=2E_A$ near the Fermi surface[14]

$$\chi_e(T)=2\chi_e(T_c)\int_D^\infty \frac{x}{\sqrt{x^2-D^2}}\frac{e^x}{(1+e^x)^2}\,dx, \tag{13}$$

where $D\equiv E_A/k_BT$. As the temperature is lowered, the latter formula tends to the expression

$$\chi_e(T)=2\chi_e(T_c)\sqrt{D}e^{1-D},$$

which approximates expression (13) well already for $D>4$. Thus, the temperature dependence of $\chi_e(T)$, in addition to its activation dependence, acquires the additional factor of $T^{-1/2}$. This is the reason why we took the kinetic coefficient for the Korringa relaxation in the form

$$V_{se}=V_{se}(T_c)\exp\left(-\frac{2E_A}{k_BT}\right), \tag{14}$$

which was previously successfully used for this material.[11] The temperature behavior of V_{es} was determined from the principle of detailed balance (6). As for the temperature dependence of the energy gap $E_A(T)$, there are as of yet no generally accepted theoretical indications in this regard. An analysis of our experimental data based on the described model required that the gap open up rapidly as the temperature decreases. Therefore, for definiteness in our calculations we took $E_A(T)$ to have the form

$$E_A(T)=\begin{cases}0, & T\geqslant T_c\\ D(0)[1-(T/T_c)^6] & T<T_c\end{cases}, \tag{15}$$

proceeding from the best fit of the calculated curves to the experimental data. From the same considerations we adopted the following values for the other parameters: $\kappa=1.0013$, $\Delta(0)/k_B=80$ K, $\chi_e(T_c)/\chi_s(T_c)=4.0$, $\omega_s/\sigma=150$, $V_{se}(T_c)/\sigma=30$, $V_{eL}/\sigma=0.11$, and $V_{sL}/\sigma=0.11$. In addition, for best fit of the calculated curves to the experimental data we assigned a monotonic twofold increase of the rate V_{sL} with increase in temperature from 40 to 55 K. A sample comparison of experimental and calculated ESR spectra is shown in Fig. 3. It can be seen that the model spectra satisfactorily describe the experimental data.

As in the previous section, further processing of the calculated spectra reduced to the representation of the calculated curves as the sum of two Lorentzians. Note that despite the Gaussian form of the function $g(\Omega)$, the calculated spectra are quite well decomposed into Lorentzian components all the way down to $T=30$ K (an example of such a decomposition is shown in Fig. 5), and only at lower temperatures is the Gaussian nature of the inhomogeneous broadening manifested. This latter circumstance lowers the reliability of the interpretation of the wide spectral component at $T<30$ K, but as before, the description of the narrow spectral component governed by the conduction electrons remains in force.

Figure 4 plots the dependence of the g-factors, the halfwidths, and the intensities of the two Lorentzian components into which both the experimental and calculated spectra were

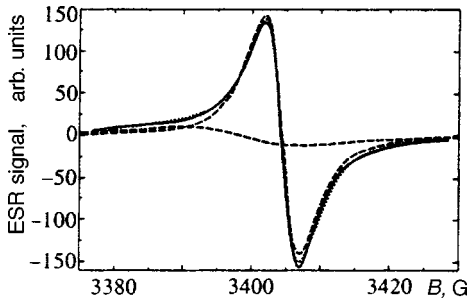

FIG. 5. Example of decomposition of the calculated ESR spectrum into a sum of two Lorentzians. Solid curve—calculated spectrum for $T=30.1$ K (see comparison with experiment in Fig. 3), dashed lines—Lorentzians of the individual components, dotted line—sum of these components.

decomposed. It can be seen from this figure that the calculated curves satisfactorily describe the experimental data. The divergence in the behavior of the g-factor of the narrow Lorentzian component at $T<35$ K can be ascribed to neglecting the parameter λ, which was mentioned above. We have difficulty in explaining the underestimated values of the intensities of the wide component at $T>35$ K in terms of the model even though note should be made of the significant error in determining the amplitude of the given component in this temperature range.

At the same time, it should be noted that in fact only four quantities κ, $V_{se}(T_c)$, $\chi_e(T_c)/\chi_s(T_c)$, and $\Delta(0)$ play the role of adjustable parameters. Here the quantity κ is uniquely determined from the experimental ratio of the g-factors of both components when the bottleneck effect is absent (at low temperatures), while $\Delta(0)$ is determined from the experimental falloff of the intensity of the narrow component at $T<30$ K.

Transforming from the dimensionless quantities to the real ones gives the following parameter values: $g_s=2.0031$, $g_e=2.0004$, $\sigma/4\pi=63.7$ MHz, $V_{se}(T_c)=1.2\times10^{10}$ s^{-1}, $V_{eL}=4.5\times10^7$ s^{-1}, $V_{sL}=9.0\times10^7$ s^{-1}, $\chi_e(T_c)=8.8\times10^{-8}$ emu/mol, $\chi_s(T_c)=2.2\times10^{-8}$ emu/mol, and $\Delta(0)/k_B=80$ K.

In conclusion, let us recapitulate the main results of our work. In metals containing both conduction electrons and localized paramagnetic centers the ESR spectrum in the presence of a strong bottleneck is described by a single Lorentzian regardless of the form of the distribution of the local magnetic fields in the sample. In this case an effective decrease of the inhomogeneous contribution to the combined ESR spectrum of coupled spin systems of conduction electrons and localized paramagnetic centers is possible for any distribution function $g(\Omega)$. In the intermediate region of the transition from strong bottleneck conditions to the case of independent spin subsystems the total ESR spectrum can be described with good accuracy by a sum of two Lorentzians, each of which is associated with one type of center, and such a representation is also valid when the inhomogeneous broadening is Gaussian. Application of our calculations to experimental data on the temperature dependence of the ESR spectra in the polymer phase of RbC_{60} yields good agreement between our results and the model of a quasi-one-dimensional metal containing two efficiently exchange-

coupled subsystems, conduction electrons and impurity paramagnetic centers.

ACKNOWLEDGMENTS

We are greatly indebted to L. Forró and A. Jánossy for kindly supplying the RbC_{60} samples for this study and for helpful discussions. We thank S. N. Artemenko, S. Zaĭtsev-Zotov and N. N. Kirova for interesting remarks made during discussions of this work. This work was supported by a grant from the Russian Fund for Fundamental Research (Grant No. 96-02-19719).

*)E-mail: demidov@mail.cplire.ru

1)The material was kindly provided by L. Forró (Département de Physique, École Polytechnique Fédérale de Lausanne, Switzerland).

[1] S. E. Barnes, Adv. Phys. **30**, 801 (1981).
[2] H. Hasegawa, Progr. Theor. Phys., Osaka **27**, 483 (1959).
[3] G. Sachs, E. Pöhlmann, and E. Dormann, J. Magn. Magn. Mater. **69**, 131 (1987).
[4] D. Jérome and H. J. Schulz, Adv. Phys. **31**, 299 (1982).
[5] O. Chauvet, G. Oszlányi, L. Forró, P. W. Stephens, M. Tegze, G. Faigel, and A. Jánossy, Phys. Rev. Lett. **72**, 2721 (1994).
[6] S. Pekker, L. Forró, L. Mihály, and A. Jánossy, Solid State Commun. **90**, 349 (1994).
[7] A. Jánossy, O. Chauvet, S. Pekker, J. R. Cooper, and L. Forró, Phys. Rev. Lett. **71**, 1091 (1993).
[8] J. Robert, P. Petit, J.-J. André, and J. E. Fischer, Solid State Commun. **96**, 143 (1995).
[9] L. R. Tagirov and K. F. Trutnev, Zh. Éksp. Teor. Fiz. **86**, 1092 (1984) [Sov. Phys. JETP **59**, 638 (1984)].
[10] V. A. Atsarkin, G. A. Vasnëva, and V. V. Demidov, Zh. Éksp. Teor. Fiz. **108**, 927 (1995) [JETP **81**, 509 (1995)].
[11] V. A. Atsarkin, V. V. Demidov, and G. A. Vasneva, Phys. Rev. B **56**, 9448 (1997).
[12] F. Bommeli, L. Degiorgi, P. Wachter, Ö. Legeza, A. Jánossy, G. Oszlányi, O. Chauvet, and L. Forró, Phys. Rev. B **51**, 14 794 (1995).
[13] J. Hone, M. S. Fuhrer, K. Khazeni, and A. Zettl, Phys. Rev. B **52**, R8700 (1995).
[14] D. C. Johnston, Phys. Rev. Lett. **52**, 2049 (1984).

Translated by Paul F. Schippnick

Influence of magnetic field on the electronic specific heat of the organic metal $(BEDT\text{-}TTF)_2KHg(SCN)_4$

A. Kovalev

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Russia

H. Mueller

European Synchrotron Radiation Facility, F-38043 Grenoble, France

M. V. Kartsovnik

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Russia; Walther-Meissner-Institut, Walther-Meissner-Str. 8, D-85748 Garching, Germany

(Submitted 17 August 1997)
Zh. Éksp. Teor. Fiz. **113**, 1058–1063 (March 1998)

Specific heat measurements of a single crystal of the organic metal $(BEDT\text{-}TTF)_2KHg(SCN)_4$ have been carried out at low temperatures and under a magnetic field of up to 14 T. A jump in the specific heat of about 0.1 J/mol·K, which corresponds to the antiferromagnetic phase transition, has been observed. The magnetic field is found to decrease the transition temperature at any field orientation. The strongest effect was found to take place in the field direction along the highly conducting ac plane. © *1998 American Institute of Physics.* [S1063-7761(98)02203-3]

1. INTRODUCTION

The organic metals $(BEDT\text{-}TTF)_2MHg(SCN)_4$, where BEDT-TTF stands for bis(ethylenedithio)-tetrathiafulvalene and M=K, Tl, and Rb are of significant interest due to their unusual properties in magnetic fields at low temperatures.[1] The Fermi surface characteristic of these compounds consists of a cylindrical (quasi-two-dimensional, Q2D) part and slightly corrugated open (quasi-one-dimensional, Q1D sheets. The instability of the Q1D conducting band against the spin density wave (SDW) formation is thought to be the reason for a phase transition occurring below 10 K in these compounds. The strong anisotropic change of the magnetic susceptibility[2] resembles that previously observed in the purely Q1D conductors such as $(TMTSF)_2X$,[3] although the magnitude of the magnetic moment modulation extracted from the μSR experiment[4] is considerably smaller, $\mu_{SDW} = 0.003\mu_B$ (here μ_B is the Bohr magneton).

According to the B-T phase diagram originally proposed by Sasaki *et al.*[5] for the M=K salt on the basis of magnetoresistance studies at different temperatures, the phase transition into the SDW state is gradually suppressed by the magnetic field applied perpendicular to the highly conducting planes, and the normal metallic state is stabilized above 23 T in the entire temperature region. Many of the magnetoresistance anomalies have been explained in the framework of the SDW model[6] and the phase diagram.[5] However, several serious problems remain. For example, the effect of magnetic field on the SDW ground state in these compounds near the transition temperature T_p is not clearly understood. For the quasi-one-dimensional $(TMTSF)_2X$ salts, a magnetic field applied perpendicular to the highly conducting planes is known to stimulate rather than suppress the SDW instability due to the effective enhancement of the one-dimensional character of the electron motion. The theoretical model developed for one-dimensional systems[7,8] and applied to $(BEDT\text{-}TTF)_2KHg(SCN)_4$ (Ref. 9) predicts a gradual increase of the SDW transition temperature with magnetic field. Other models considering magnetic breakdown between Q1D and Q2D parts of the Fermi surface[10,11] or fluctuation effects[12] have been proposed as possible explanation for the field-induced reentrant transition. We note that all the theories mentioned above consider the effect of the field directed perpendicular to the highly conducting planes as the most significant effect. However, from the experimental point of view, no agreement has been established as to whether the perpendicular field component has the main effect on the low-temperature antiferromagnetic state near T_p.

The B-T phase diagram was recently revised on the basis of the comparative magnetoresistance and magnetic torque studies by Sasaki *et al.*[9] and Kartsovnik *et al.*[13] Both investigations proposed that the low-temperature state differs from the normal metallic state even in the high-field region. However, no agreement has been reached as to the magnetic field effect in the low-field range. According to the diagram proposed by Sasaki *et al.*,[9] at least two successive transitions take place upon cooling the sample in a finite field, where the magnetically ordered state is stabilized by the field. In contrast, only one transition was found by Kartsovnik *et al.*[13] in the temperature dependences of the resistance and torque at fields below 10 T; the transition temperature shifted to slightly lower temperatures as the field was raised.

Thus, the effect of magnetic field on the low-temperature state of α-$(BEDT\text{-}TTF)_2MHg(SCN)_4$ is still an open question and further detailed studies are required in order to

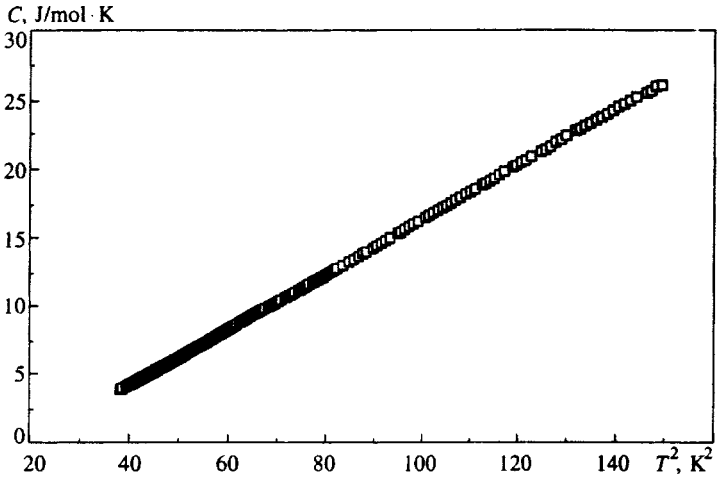

FIG. 1. Specific heat of $(BEDT\text{-}TTF)_2KHg(SCN)_4$ versus temperature squared.

clarify the problem. We report here the results of an experimental study of the heat capacity of a α-$(BEDT\text{-}TTF)_2KHg(SCN)_4$ single crystal as a function of temperature under a magnetic field up to 14 T. A single prominent anomaly corresponding to one phase transition has been observed at all applied fields. The anomaly gradually shifts down as the field increases. Tilting the field from the direction normal to the highly conducting planes increased the magnitude of the shift.

2. EXPERIMENTAL

The single crystal used for the experiment was obtained by galvanostatic electrolysis ($j=1.0\ \mu A/cm^2$) of a solution of KSCN, $Hg(SCN)_2$ and 18-crown-6 (10 mmol/l each) in 1,1,2-trichloroethane/MeOH (abs.) (9:1; 25 ml) in the presence of BEDT-TTF (13 mg) at a temperature of 25 °C.

To measure specific heat we used, the standard ac modulation technique.[14] The magnitude of the temperature modulation was 0.5–2%, the modulation frequency was $\nu=0.04$ Hz and the calorimeter-to-bath relaxation time τ was a few seconds in the studied temperature region of 6–14 K so that $2\pi\nu\tau>1$. The calorimeter consisted of a bare chip of Cernox thermoresistor and a carbon heater. The thermal link between the calorimeter and the bath was provided by 50-μm manganin wires, which were used also as electrical leads to the thermometer and heater. A single crystal of $(BEDT\text{-}TTF)_2KHg(SCN)_4$ with the mass of 0.9 mg was fixed by applying a small amount of Apiezon N grease to the thermometer and a smaller amount of the same grease was used to fix the heater to the sample. The absolute value of the specific heat of the sample was determined with an accuracy of 5% and was approximately a factor of 4 larger than the admixture. The temperature error due to the magnetoresistance of the Cernox thermoresistor did not exceed 0.1 K at 14 T and was taken into account in the data analysis.

3. RESULTS AND DISCUSSION

The total specific heat at zero field is presented in Fig. 1 (here 1 mol$=N_A$ formula units, where N_A is the Avogadro number). The specific heat is plotted versus temperature squared. Since the plot is almost linear in such coordinates, we can approximate the monotonic background as: $C=a+bT^2$. The specific heat, after subtraction of the background at zero magnetic field and at two orientations of the magnetic field $B=14$ T, is shown in Fig. 2; the data are offset for clarity. The phase transition is manifested by a peak-like feature in the heat capacity. At zero field the maximum is at 8 K and it shifts to lower temperature in a magnetic field. The relatively large width of the anomaly is probably caused by the strong fluctuations. The magnitude of the specific heat jump at zero field is evaluated as $C_0=0.1$ J/mol·K or about 1% of the total specific heat. This value agrees with our previous measurements,[15] but it is about a factor of 3 smaller than the value reported by Henning *et al.*[16] On the other hand, Nakazawa *et al.*[17] may not have observed the anomaly since the scattering of their experimental data exceeded 1%. Figure 3 shows the field dependence of the maximum position at two field orientations: **B** is perpendicular to the highly conducting **ac** plane and **B** is parallel to this plane. We see that in both cases the shift of the maximum position is proportional to the magnetic field squared, $T_p(B)=D(\theta)B^2$,

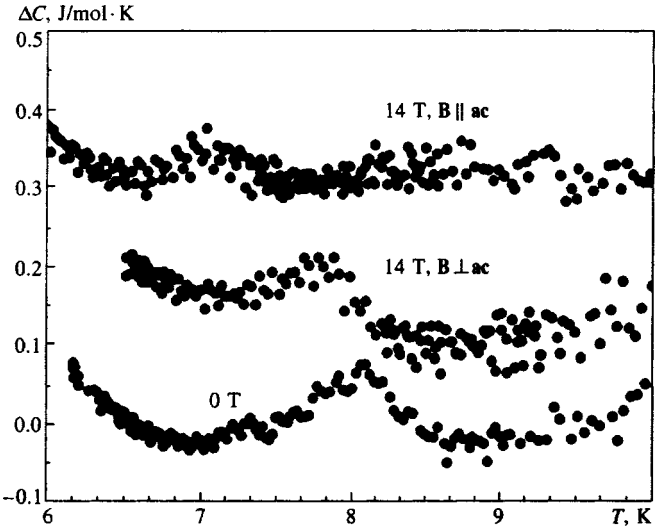

FIG. 2. Specific heat of $(BEDT\text{-}TTF)_2KHg(SCN)_4$ after subtraction of the monotonic background at zero field and at field $B=14$ T directed parallel and perpendicular to the **ac** plane.

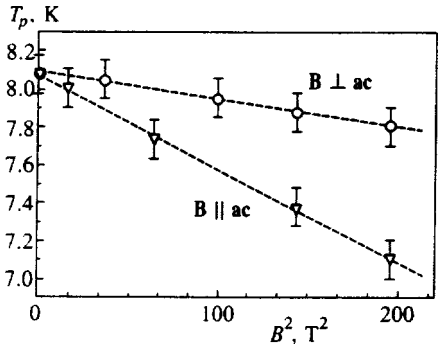

FIG. 3. Field dependence of T_p versus magnetic field squared at two field directions.

where θ is the angle between the normal to the **ac** plane and the magnetic field direction. For the field orientations shown in Fig. 3 we find

$$D(\theta=0°)=(1.5\pm0.4)\times10^{-3}\ \ \mathrm{K/T^2} \tag{1}$$

and

$$D(\theta=90°)=(5.0\pm0.5)\times10^{-3}\ \ \mathrm{K/T^2}.$$

The angular dependence of the maximum position is shown in Fig. 4 for $B=14$ T. The plot demonstrates an approximately linear dependence on the square of the cosine of the angle θ between the normal to the **ac** plane and the field direction.

Assuming that the phase transition is second order, we can relate the behavior of the specific heat to the magnetic susceptibility by using the thermodynamic identity

$$C_i-C_j=-T\mu_0\left(\frac{\partial H_{ij}}{\partial T}\right)_H\left[\left(\frac{\partial M_i}{\partial T}\right)-\left(\frac{\partial M_j}{\partial T}\right)\right]_H$$

$$=-\frac{T\mu_0}{2}\left(\frac{\partial H_i^2}{\partial T}\right)_H\left[\left(\frac{\partial \chi_i}{\partial T}\right)-\left(\frac{\partial \chi_j}{\partial T}\right)\right]_H, \tag{2}$$

where μ_0 is the magnetic permeability of vacuum. $H_{ij}(T)$ is the phase-separation line, and $C_i(C_j)$ and $\chi_i(\chi_j)$ are the specific heat and susceptibility of the phase $i(j)$, respectively. Although in our case the transition is significantly broadened in temperature, we will attempt to evaluate $\Delta(\partial\chi/\partial T)_H$ in the antiferromagnetic state from Eq. (2) using our data. For **B** parallel to the **ac** plane we estimate

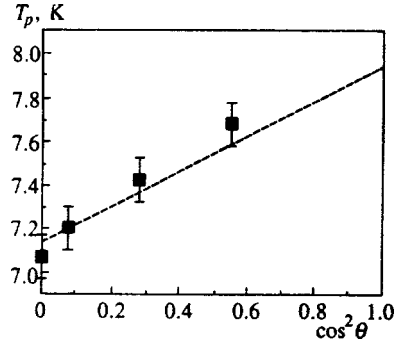

FIG. 4. Angular dependence of T_p versus cosine of the angle between the normal to the **ac** plane and field direction.

$$\Delta\left(\frac{\partial\chi}{\partial T}\right)_H\approx2.6\times10^{-7}\ \ \mathrm{K^{-1}}$$

in the limit $B\to0$, in agreement with the susceptibility data from Ref. 2 and Ref. 18. For **B** perpendicular to the **ac** plane we estimate

$$\Delta\left(\frac{\partial\chi}{\partial T}\right)_H\sim0.8\times10^{-8}\ \ \mathrm{K^{-1}}\ \text{as}\ \ B\to0.$$

We note that no significant change in the susceptibility in that direction was found in the magnetization measurements[2] in the field $B=5$ T. In principle, this may be explained by a dependence of the transition temperature, which is much weaker than B^2 in fields below 5 T. Such an assumption is consistent with our data, as can be seen in Fig. 3. The uncertainty in the determination of the peak position does not allow us to make a definite conclusion about the field effect below 5 T.

Thus, our specific heat data show only one phase transition in a magnetic field in the range $0<B<14$ T. In contrast with the conclusion made by Sasaki *et al.*,[9] no evidence of two successive transitions below 14 T was found. Therefore, we affirm that in the field range studied by us the only effect of magnetic field is that it gradually shifts the transition to a lower temperature. An important result is that the shift is dependent on the field direction.

The fact that the field parallel to the **ac** plane has a stronger effect than the field perpendicular to the **ac** plane seems to be consistent with the assumption about the easy-axis or the easy-plane antiferromagnetic order.[19 1)] For the field along the **ac** plane we can compare the obtained shift of the transition temperature with that predicted, in the mean field approximation, for the SDW state with an easy-axis ordering under the magnetic field parallel to the spin alignment direction given by[20]

$$T_p(B)=T_0\ \exp(-7\zeta(3)b^2), \tag{3}$$

where $b=\mu_BB/2\pi kT$. In low fields this expression reduces to[21]

$$\frac{T_p(B)-T_p}{T_p}\approx0.2\left(\frac{\mu_BB}{k_BT_p}\right)^2, \tag{4}$$

which has a parabolic field dependence, in agreement with our data. Substituting $T_p=8$ K, we obtain from Eq. (4) the shift of the transition temperature 0.01 K/T^{-2}, in reasonable agreement with the experimental value (1).

The influence of the magnetic field perpendicular to the **ac** plane is rather different in our measurements than expected for the conventional SDW state. According to the theoretical prediction,[7,8] it also differs from the results reported by Sasaki *et al.*[9] As mentioned above, other theoretical models[10,11] predict a decrease of the transition temperature for this field direction. Probably the total effect of the magnetic field is a superposition of one of those mechanisms and the spin-field interaction.

In conclusion, we have found an anomaly in specific heat of (BEDT-TTF)$_2$KHg(SCN)$_4$ which corresponds to the antiferromagnetic ordering phase transition. This anomaly is

shifted toward low temperatures in an applied magnetic field. The shift occurs at any direction of the magnetic field; the strongest effect is observed for the field parallel to the **ac** plane. We suggest that for this field orientation the shift may be understood as resulting from the destruction of the SDW phase via the interactions of the electron spins with the external magnetic field. The reason for the decrease in the transition temperature in a field perpendicular to the **ac** plane is yet to be clarified.

This work was supported in part by the Russian Fund for Fundamental Research (No. 96-02-17475) and INTAS (Grant No. 93-2400-EXT).

[1)]We note that although the SDW transition in Q1D organic metals is commonly characterized by an easy-axis antiferromagnetic ordering, recent torque experiments pointed out the possibility of a more complicated spin alignment in the **ac** plane.

[1] For an overview see, e.g., J. Wosnitza, *Fermi Surfaces of Low-Dimensional Organic Metals and Superconductors*, p. 80–99, Springer-Verlag, Berlin-Heidelberg (1996); M. V. Kartsovnik and V. N. Laukhin, J. Phys. I France **6**, 1753 (1996); S. Uji, T. Terashima, H. Aoki, J. S. Brooks *et al.*, Phys. Rev. B **54**, 9332 (1996), and the references cited there.
[2] T. Sasaki, H. Sato, and N. Toyota, Synth. Met. **41–43**, 2211 (1991).
[3] T. Ishiguro and K. Yamaji, *Organic Superconductors*, Springer-Verlag, Berlin (1990).
[4] F. L. Pratt, T. Sasaki, and N. Toyota, Phys. Rev. Lett. **74**, 3892 (1995).
[5] T. Sasaki and N. Toyota, Solid State Commun. **82**, 447 (1992).
[6] M. V. Kartsovnik, A. E. Kovalev, and N. D. Kushsch, J. Phys. I France **3**, 1187 (1993).
[7] L. P. Gor'kov and A. G. Lebed', J. Phys. (France) Lett. **45**, L-433 (1984).
[8] G. Montambaux, Phys. Rev. B **38**, 4788 (1988).
[9] T. Sasaki, A. G. Lebed', T. Fukase, and N. Toyota, Phys. Rev. B **54**, 12969 (1996).
[10] T. Osada, S. Kagoshima, and N. Miura, Synth. Met. **70**, 931 (1995).
[11] K. Kishigi and K. Machida, J. Phys. Soc. Jpn. **64**, 3853 (1995).
[12] R. H. McKenzie, Phys. Rev. Lett. **74**, 5140 (1995).
[13] M. V. Kartsovnik, W. Biberacher, E. Steep, P. Christ *et al.*, Synth. Met. **86**, 1933 (1997).
[14] P. F. Sullivan and G. Seidel, Phys. Rev. **173**, 679 (1968).
[15] A. Kovalev and H. Mueller, Synth. Met. **86**, 4668 (1996).
[16] P. H. Henning, J. S. Brooks, J. E. Crow, Y. Tanaka *et al.*, Solid State Commun. **95**, 691 (1995).
[17] Y. Nakazawa, A. Kawamoto, and K. Kanoda, Phys. Rev. B **52**, 12890 (1995).
[18] P. Crist, W. Biberacher, H. Mueller, and K. Andres, Solid State Commun. **91**, 451 (1994).
[19] P. Christ, W. Biberacher, W. Bensch, H. Mueller *et al.*, Synth. Met. **86**, 2057 (1997).
[20] A. Bjelis and D. Zanchi, Phys. Rev. B **49**, 5968 (1994).
[21] G. Sarma, J. Phys. Chem. Solids **24**, 1029 (1963).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Contribution of Andreev reflection to the increase in the resistance of the normal metal in a bimetallic $N-S$ structure

Yu. N. Chiang*) and O. G. Shevchenko

B. I. Verkin Physicotechnical Institute of Low Temperatures, National Academy of Sciences of Ukraine, 310164 Kharkov, Ukraine

(Submitted 17 September 1997)
Zh. Éksp. Teor. Fiz. **113**, 1064–1070 (March 1998)

We study the resistive properties of 3D normal-metal–superconductor systems in the pure mean-free-path limit $l_{N,S} \gg \xi(T)$ ($l_{N,S}$ are the mean free paths in the metals, and ξ is the coherence length) at liquid helium temperatures. In contrast to the situation where $l \ll \xi$, which is common in experiments involving either sandwiches or mesoscopic samples, here the $N-S$ system exhibits unusual temperature behavior that cannot be described by existing theories of boundary resistance. What is most remarkable is a rise in normal resistance in regions that do not incorporate the $N-S$ boundary as the temperature decreases, with asymptotic behavior resembling that of the temperature curve of the gap of a superconductor in contact with a normal metal. We show that this effect, not observed earlier in 3D systems, is due to the nonequivalence of the cross sections of scattering by normal-metal impurities of electron and hole excitations in conditions of Andreev reflection. We also show that in standard measurements of the contribution of the $N-S$ boundary lying between the test contacts, this effect is masked by accompanying effects, the proximity effect and the boundary resistance, whose estimate requires taking into account the presence on the $N-S$ boundary of an electrostatic barrier of the Schottky type, a barrier that redistributes the probabilities of ordinary and Andreev reflections of quasiparticles in the nonequilibrium conditions due to current flow. [S1063-7761(98)02303-8]

Usually, when measuring the temperature dependance of the resistance of a pure normal metal, one expects that a decrease in the temperature T leads to a decrease in the metal's resistance, with a saturation plateau reached as $T \to 0$. In this paper we report on results of experiments that demonstrate the opposite behavior as a current is introduced into the normal metal through a contact with a pure superconductor, with the area of the $N-S$ contact being so large that tunneling properties play no role.

We have studied three-dimensional (3D) $N-S$ systems, which guarantees that several conditions are met simultaneously: that the bulk characteristics of the initial metal are retained, e.g., the macroscopically large mean free paths l, which ensures a certain freedom in positioning the test contacts within regions of the size of the mean free path surrounding the $N-S$ boundary; that the pure limit $l \gg \xi_{S,N}$ (here ξ_S and ξ_N are the coherence lengths in the superconductor and the normal metal, respectively) is achieved so that the proximity effect can be easily identified; and, finally, that a $N-S$ contact can be manufactured that is sure to have nontunnel characteristics (it is difficult to meet these conditions in thin-film systems, which are usually used in studies of the $N-S$ boundary). The samples were bimetallic $N-S$ systems consisting of single-crystal copper with tin fused on one of the faces, with mean free paths ranging from 10 to 100 μm in both metals. The area of the $N-S$ boundary varied within $(l-10l)\times(l-10l)$. The test contacts in the normal region of the system (copper) were placed at distances $L_2=(0.5-2)l$ for the contact closest to the boundary and, accordingly, at distances $L_1=(3-10)l$ for the second contact, farthest from the boundary.

Figure 1 (curve *1*) depicts the temperature curve for the resistance between contacts in the normal region of an $N-S$ system for one sample (with $L_1 \approx 3l$, $L_2 \approx 0.5l$, and $l \approx 20$ μm) in a temperature interval containing the transition temperature T_c of tin. Figure 2 (curve *1*) depicts the temperature curve for the resistance of a region containing the $N-S$ boundary for the same sample at temperatures below T_c. For the other samples the corresponding temperature curves have a similar shape. The first of these curves, measured for test contacts that were located at great distances from the $N-S$ boundary, clearly correlates with the transition of tin into the superconducting state, as a comparison with Fig. 2 clearly shows, but has a shape that is unusual for a normal metal: as the temperature decreases, the resistance grows, and this increase in resistance is fairly rapid near T_c (we believe that this effect is not described in the literature).

To understand the nature of this effect, we first turn to the theory of the boundary resistance of $N-S$ systems that allows for the contribution of this resistance to the measured temperature curves in Figs. 1 and 2.

As is known,[1] an $N-S$ boundary raises the resistance of a system consisting of a normal metal and a type I superconductor, measured between test points positioned in such a way that the boundary lies between the points. The effect manifests itself most vividly near the transition temperature

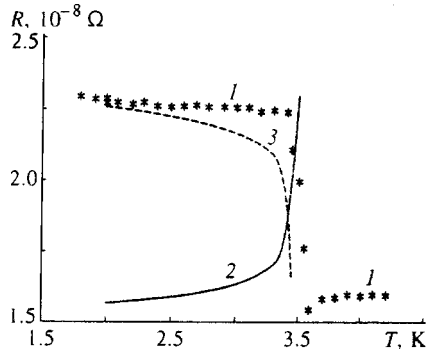

FIG. 1. Temperature behavior of the resistance of the normal metal in a 3D structure below T_c of the superconductor: *1*) experimental points; *2*) boundary resistance ($z \neq 0$); *3*) contribution of Andreev reflection.

T_c of the superconductor ($T \leqslant T_c$), where an extremely strong temperature dependence of the excess resistance is predicted. Many researchers believe (see, e.g., Refs. 2 and 3) that this resistance is the resistive contribution of a layer of the superconductor with a thickness on the scale of the rms of the inelastic and elastic mean free paths of electrons in this metal, i.e., it is assumed that there exists an extended layer of the superconductor with nonzero potential, although the distribution of the total current outside the Meissner layer is described by the same equation as in the absence of an electric field in the superconductor.[2]

Now, variations in the mean free path in the normal metal are allowed here, since the excess resistance is assumed to develop exclusively in the superconducting half-space. However, earlier we observed (see Ref. 4) a decrease in the resistance of an $N-S$ double system in the transition of the superconductor to the normal state, and this effect suggested that the excess resistance may be the result of variations of the electron mean free path in the normal-metal half-space. The experiment described in Ref. 4 was stimulated by one of the first such ideas, based on the contribution of non-Andreev reflection of quasiparticles ''gliding'' along the $N-S$ boundary[5] and having energy lower than the gap energy.

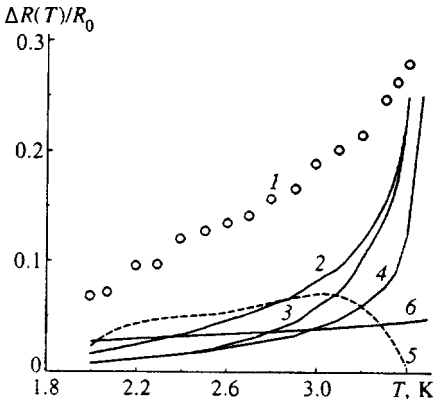

FIG. 2. Resistance of the region of the $N-S$ system incorporating the $N-S$ boundary, below T_c of the superconductor: *1*) experimental points; *2*) contribution of the boundary resistance ($z \neq 0$); *3*) contribution of the boundary resistance at $z=0$; *4*) results of calculations of the boundary resistance by the theories of Artemenko *et al.*[2,3,13] and Hsiang and Clarke theory;[3] *5*) contribution of Andreev reflection; *6*) contribution of the proximity effect.

Later certain indications emerged[6–8] that pointed to the possibility of other variations in the effectiveness of scattering in the nonequilibrium region of the normal metal adjacent to the superconductor, a layer whose thickness is of the order of the mean free path. Generally, two mechanisms can be distinguished here. One is related to the special features of the scattering by the impurities closest to the boundary in the Andreev reflection of quasiparticles with energies lower than the gap energy.[7,8] The second is related to the special features of the reflection of quasiparticles with energies higher than the gap energy by the potential barrier of electrostatic origin, provided that such a barrier can form at the boundary between two conductors, one of which is a superconductor.[9] Blonder *et al.*[9] developed the method that allows for the second mechanism and assumed that such a barrier forms at random, while the height of the barrier was assumed fixed.

There is, however, the possibility that the contact potential difference plays an important role in forming a potential barrier between two metals at a temperature at which one metal passes to the superconducting state. If the fields in the metals on both sides of the boundary do not differ too strongly, as is the case with metals whose conductances are of the same order, the contact potential difference probably cannot lead to a sizable potential barrier, so that such a barrier can be ignored. But when the fields differ considerably, as in the case when a semiconductor is in contact with an ordinary metal, the contact potential difference, lumped in the boundary region of the conductor with the lower conductance, bends the bands and initiates the formation of a potential, or Schottky, barrier.

A similar situation emerges at the boundary between a normal metal and a superconductor, whose electric potential and, hence, the contact potential difference acquire an addition, in comparison to the normal state of the superconductor, of the order of the Bernoulli potential Δ^2/ε_F (Refs. 10 and 11), i.e., of order of the variation in the energy of the ground state of the metal that becomes a superconductor.[12] At atomic distances this energy corresponds to a field strength of several volts per centimeter, which is quite sufficient for the formation of a barrier of the Schottky type with a height that cannot be ignored in large-area contacts with high transmittance (in 3D contacts).

Let us use the method of calculating the boundary resistance developed by Blonder *et al.*[9] for narrow channels. This approach is justified, since the resistance of the boundary region (no thicker than the coherence length of the normal metal) in the vicinity of the $N-S$ boundary is an order of magnitude higher than the resistance of the normal metal, so that all the voltage is actually lumped in this region as in a narrow channel, and our system can be interpreted as an $N-I-S$ structure. Naturally, the large lateral dimensions of the $N-S$ contact ($\leqslant 10l$) introduce certain limitations, so that the theory developed by Blonder *et al.*[9] in our case gives the qualitative picture of the behavior of the resistance rather than the quantitative.

Calculations of the current via probability relations for quasiparticle states within the framework of the Bogolyubov

equation and the boundary conditions of Andreev and ordinary reflections at the $N-S$ boundary (for more details see Ref. 9), done with allowance for the barrier mentioned earlier, lead to the following expression obtained entirely in the normal half-space (relative to the $N-S$ boundary) without allowance for nonequilibrium processes in the superconductor:

$$I_{N-S}=2N(\varepsilon_F)e^2v_FV\frac{[\Delta(T)/k_BT]\exp[\Delta(T)/k_BT]+C}{z_0^2+4z^2}, \tag{1}$$

where $N(\varepsilon_F)$ is the density of states at the Fermi surface, e is the electron charge, v_F is the Fermi velocity, eV is electron energy at the $N-S$ boundary of the normal half-space, $\Delta(T)$ is the energy gap in the superconductor, k_B is the Boltzmann constant, z is the dimensionless height of the barrier that forms at the normal-metal–superconductor junction (for more details see Ref. 9): $z=k_BN(\varepsilon_F)\Delta(T)/2\varepsilon_F$ ($z\sim 1$ for the copper–tin pair as $T\to 0$), and z_0 is the height of a barrier of static origin, which is temperature-independent (much smaller than unity for 3D contacts); C is a constant such that $C\ll 1$ for $z\leqslant 1$ and $C\sim 1$ for $z\gg 1$.

For $T\ll T_c$ this result is identical (if we put $z=0$) to the one obtained from studies of nonequilibrium processes in a superconductor,[2,3] when the total current is calculated as the sum of two contributions, the contribution related to charge imbalance, and the contribution that gets converted directly into supercurrent. This suggests that the model of Blonder *et al.*[9] is correct; in it a variation in the conductance of the $N-S$ system in a nonequilibrium situation (in the presence of a current) is the result of two different probabilities of occupation of the branches of the drastically different spectra of excitations in the normal and superconducting half-spaces even in the ballistic regime of crossing the $N-S$ boundary. The reader will recall that the starting assumption in this model is that the field transferred to the superconductor by excitations with an energy $E>\Delta$ is balanced by the gradient of the chemical potential of the superconducting electrons outside the Meissner layer. We also note that, in contrast to the narrow-channel model used by Blonder *et al.*,[9] the given expression leads to no resistance for the $N-N$ boundary with z_0, $z=0$, as should be the case for 3D contacts, for which our calculations were done.

Equation (1) determines the contribution of the $N-S$ boundary to the limiting of the current in the system due to the flux of particles reflected by the boundary into the normal metal in the direction opposite that of the initial current. One should expect that this contribution is also present, with a weight $\exp(-L/l)$, in test contacts that are entirely inside the normal metal; here L is the distance from the $N-S$ boundary to the corresponding test contact in the normal metal. (But if the limiting of the current is related solely to the resistive contribution of the scattering processes in the superconducting half-space, it is clear that this has no effect on the resistance measured between test contacts in the normal metal.)

Equation (1) directly describes the current in an $N-S$ system in the constant-voltage regime (i.e., the internal resistance of the current source can be ignored), in which within the temperature range from T_c to T_0 the currents may assume any value from 0 to ∞. In real experiments, in which measurements are conducted in the constant-current regime, I_0, the maximum values of current cannot exceed I_0, so that the effective current at the $N-S$ boundary can be written as follows:

$$I_b^*(T)=I_b(T)[1+(I_b(T)/T_0)]^{-1}\equiv I_0-I_{0\ \mathrm{ref}}(T),$$

where $I_{0\ \mathrm{ref}}(T)$ is the flux of the charged particles reflected from the boundary into the normal metal. If we allow for the exponential nature of the damping of this flux (as noted earlier) as we move away from the boundary, for a point x_i of the normal half-space we can write

$$I_{0\ \mathrm{ref}}(T,x)=\alpha(x)I_{0\ \mathrm{ref}}(T),\quad \alpha(x)=1-\exp(-lx),$$

$$I^*(T,x_i)=I_0-\alpha(x_i)I_{0\ \mathrm{ref}}(T),$$

where x is the coordinate of the point measured from the $N-S$ boundary ($x=0$).

In practice, one usually measures the effective resistance R^*, which in the situations considered here exceeds the true value R_0 by

$$\frac{\Delta R^*}{R_0}\equiv\frac{R^*-R_0}{R_0}=I_0(x_i-x_k)\left[\int_0^{x_i}I^*(T,x)dx-\int_0^{x_k}I^*(T,x)dx\right]^{-1}. \tag{2}$$

Figures 1 and 2 (curves *2*) depict the temperature curves for $R^*(T)$ and $\Delta R^*/R_0$, calculated by Eq. (2) together with (1), and the corresponding experimental data for $x_i=46$ μm and $x_k=11$ μm (Fig. 1) and for $x_i=11$ μm and $x_k=0$ μm (Fig. 2). The latter case corresponds to a situation in which the $N-S$ boundary is between the test contacts. Figure 2 also depicts the temperature curve (curve *4*) calculated according to the theories of Artemenko *et al.*[2,3,13] and Hsiang and Clarke[3] for the case when one test contact is placed on the superconductor (these theories do not allow for excess resistance on test contacts in the normal metal), and the temperature curve for the boundary resistance, with $z(T)$ ignored. The coefficients of all the calculated curves are normalized in a unified manner to two limits: $R^*=R(T_c=3.5\ \mathrm{K})$ and $R^*(T\to 0)\to 0$.

Thus, there is no theory of excess resistance— including the theory that yields the greatest value of this resistance due to allowance for additional energy barriers at the $N-S$ boundary—that describes the results of real experiments, even at the qualitative level, for test contacts in the normal metal. Comparison of the temperature behavior of the resistance of the region containing the $N-S$ boundary (Fig. 2) and the resistance measured on the test contact deep inside the superconductor (i.e., far from the boundary) suggests that there is a 0.2-K difference in the observed transition (or critical) temperatures (3.5 K and 3.7 K). This explains the initial decrease in the resistance at $T=3.7$ K of the test contacts in the normal region (Fig. 1) as a manifestation of the proximity effect, before the entire superconductor, including its boundary layer (whose critical temperature differs somewhat from that of the bulk of the superconductor), becomes superconductive. Hence it is only natural to allow for the contribution

of the proximity effect in the region of the test contact as well, including the $N-S$ boundary. According to Ref. 12, this contribution in the normal region can be estimated at

$$R^*_{\rm prox}(T,x_N)\approx R_N\left\{1-\frac{1-\exp(-\Delta/k_BT)}{1+\exp(-\Delta/k_BT)}\,\frac{\xi_N(T)}{\xi(T)+\xi_N(T)}\right.$$
$$\left.\times\int_0^{x_N}\exp\left[-\frac{x}{\xi_N(T)}\right]\frac{dx}{x_N}\right\}, \qquad (3)$$

where x_N is the coordinate of the test contact in the normal region, whose resistance is R_N. Subtracting the contributions of the proximity effect (curve *6* in Fig. 2) and the boundary resistance from the experimental data, we arrive at the temperature dependence of the resistance of the region incorporating the $N-S$ boundary, which is depicted by curve *5* in Fig. 2. For test contacts that are entirely in the normal region, the same procedure yields the results depicted by curve *3* in Fig. 1.

Thus, notwithstanding the entirely different temperature curves for the resistance in different regions of an $N-S$ system, there is a general pattern that can be observed if the appropriate effects are taken into account; namely, there is a contribution that increases with decreasing temperature by a law characteristic of gap temperature behavior. It can be assumed that the nature of such behavior is related to the nonequivalence, predicted by Herath and Rainer,[6] of the cross section of scattering by normal-metal impurities near the $N-S$ boundary of electron and Andreev-reflected hole excitations. In their calculations, Kadigrobov *et al.*[14] found that allowing for this effect in multiple reflections and averaging over the angles can double the effective cross section of scattering by impurities in a layer of thickness of order $\xi_N\approx\hbar v_F/k_BT$, with $\Delta R^*/R\approx\xi_N/L$, where L is the distance from the $N-S$ boundary. For our test contacts in the normal region at $T=2$ K ($\xi_N\approx 6$ μm), $\xi_N/L_1=0.13$ and $\xi_N/L_2=0.55$, so that we can expect that on the average $\Delta R^*/R$ amounts to 33%, which is extremely close to the 44% observed in experiments (Fig. 1). Obviously, for a pair of test contacts encompassing the $N-S$ boundary, we must allow for the competing mechanism of conductance enhancement due to an appreciable change in the number of reflections by the boundary as the temperature drops, when the potential barrier becomes higher.[7]

The authors are grateful to A. M. Kadigrobov and A. N. Omel'yanchuk for useful discussions.

*)E-mail: chiang@ilt.kharkov.ua

[1] G. L. Harding, A. B. Pippard, and J. R. Tomlinson, Proc. R. Soc. London, Ser. A **340**, 1 (1974).
[2] S. N. Artemenko and A. F. Volkov, Zh. Éksp. Teor. Fiz. **72**, 1018 (1977) [Sov. Phys. JETP **45**, 533 (1977)].
[3] T. Y. Hsiang and J. Clarke, Phys. Rev. B **21**, 945 (1980).
[4] Yu. N. Tszyan (Chiang) and O. G. Shevchenko, Fiz. Nizk. Temp. **14**, 543 (1988) [Sov. J. Low Temp. Phys. **14**, 299 (1988)].
[5] A. M. Kadigrobov, Fiz. Nizk. Temp. **14**, 427 (1988) [Sov. J. Low Temp. Phys. **14**, 236 (1988)].
[6] J. Herath and D. Rainer, Physica C **161**, 209 (1989).
[7] B. J. van Wees, P. de Vries, P. Magnic, and T. M. Klapwijk, Phys. Rev. Lett. **69**, 510 (1992).
[8] A. M. Kadigrobov, Fiz. Nizk. Temp. **19**, 943 (1993) [Low Temp. Phys. **19**, 671 (1993)].
[9] G. E. Blonder, M. Tinkham, and T. M. Klapwijk, Phys. Rev. B **25**, 4515 (1982).
[10] A. N. Omel'yanchuk and S. I. Beloborod'ko, Fiz. Nizk. Temp. **9**, 1105 (1983) [Sov. J. Low Temp. Phys. **9**, 572 (1983)]; A. N. Ohmelyanchuk, in *Progress in High-Temperature Superconductivity*, Vol. 11, World Scientific, Singapore (1989), p. 99.
[11] Yu. N. Chiang and O. G. Shevchenko, Fiz. Nizk. Temp. **22**, 669 (1996) [Low Temp. Phys. **22**, 513 (1996)].
[12] P. G. de Gennes, *Superconductivity of Metals and Alloys*, W. A. Benjamin, New York (1966).
[13] S. N. Artemenko, A. F. Volkov, and A. V. Zaitsev, J. Low Temp. Phys. **30**, 487 (1978).
[14] A. Kadigrobov, R. Shekhter, and M. Jonson, Physica B **218**, 134 (1996).

Translated by Eugene Yankovsky

Experimental investigation of the behavior of the specific heat in finite systems in the vicinity of the critical mixing point

V. P. Voronov and V. M. Buleĭko

Institute for Oil and Gas Problems, Russian Academy of Sciences, 117917 Moscow, Russia
(Submitted 29 September 1997)
Zh. Éksp. Teor. Fiz. **113**, 1071–1080 (March 1998)

The specific heat of a 2,6-lutidine–water mixture is measured in the vicinity of the lower critical mixing point in the bulk and in porous media. The results of the measurements are interpreted by finite-size scaling. In particular, a universal function of the ratio t/t^* [t is the dimensionless deviation of the temperature from the specific-heat maximum $T_m(L)$, and t^* is the characteristic dimensionless temperature, which depends on the pore size] describing the behavior of the specific heat in the vicinity of the critical point in porous media with different pore sizes is obtained. The results obtained are consistent with the predictions of finite-size scaling and with the data from a numerical calculation of the specific heat of the finite three-dimensional Ising model. © *1998 American Institute of Physics.* [S1063-7761(98)02403-2]

1. INTRODUCTION

In recent years there has been a great deal of interest in the investigation of phase transitions in liquids placed in a porous medium. There are several reasons for this. First, a porous medium is characterized by a large surface, and, therefore, surface phenomena play an essential role. Second, a porous medium is a finite system of size L (the characteristic pore size), which can be commensurate with the correlation length ξ_c. The thermodynamic properties of a substance in a porous medium clearly depend on the relation between L and ξ_c and can differ significantly from the bulk properties ($L \gg \xi_c$). In the vicinity of a second-order phase transition, where the correlation length increases without bound as the critical temperature T_c is approached, i.e.,

$$\xi_c = \xi_0 [(T - T_c)/T_c]^{-\nu},$$

such a situation can occur in an experimentally achievable region near T_c even for characteristic system sizes $L \sim 10^4$ Å.

The experimental data[1–7] obtained by studying phase transitions in porous media are generally interpreted within the theory of phase transitions in a random field, where $L \ll \xi_c$ and the porous medium plays the role of the random field. Only in Ref. 8, in which the specific heat of ^{4}He placed in a porous medium was measured in the vicinity of the λ transition, were the experimental data compared with the predictions of finite-size scaling theory.[9,10]

In a bulk system the behavior of the specific heat can be described in a broad vicinity of the critical point by the expression

$$C_p = A^{\pm} |\tau|^{-\alpha} + D^{\pm} |\tau|^{\Delta - \alpha} + B^{\pm} + E\tau, \tag{1}$$

where $\tau = (T_c - T)/T_c$, α is the specific-heat critical exponent, and $B^{\pm} + E\tau$ is the regular part of the specific heat. The second term in (1) is the nonasymptotic Wegner correction, and $\Delta = 0.5$. The plus and minus signs denote the unordered and ordered phases, respectively.

In a finite-size system the behavior of the specific heat in the vicinity of the critical point depends on the characteristic size L of the system. When the correlation length ξ_c becomes commensurate with the size of the system, the anomalous behavior of the specific heat is suppressed, and instead of increasing without bound it has a maximum

$$C_m \sim (L/\xi_0)^{\alpha/\nu}$$

at

$$t_m(L) = (T_m(L) - T_c)/T_c.$$

The increase in the correlation length also has a bound, and it reaches a maximum value $\xi_{\max} \approx \xi_0 t_m(L)^{-\nu}$.

According to finite-size scaling theory, the specific heat can be represented in the form

$$C_p(t, L) = A^{\pm} |t|^{-\alpha} F(t/t^*) + C_{\text{reg}}, \tag{2}$$

where $F(t/t^*)$ is a universal function of the dimensionless parameter t/t^*. Here the deviation of the temperature of the system t is defined relative to the temperature of the specific-heat maximum $T_m(L)$ [$t = (T - T_m(L))/T_m(L)$], which differs from the usual definition of t relative to the bulk critical temperature T_c. However, in an experiment $T_m(L)$ is the only special temperature in the behavior of the specific heat. In the limit for a bulk system ($L \to \infty$), $T_m(\infty) = T_c$. The characteristic temperature

$$t^* \sim (L/\xi_0)^{-1/\nu}$$

has the meaning of the temperature at which the behavior of the specific heat in a system of size L begins to deviate from the bulk behavior.

The explicit form of the universal function $F(t/t^*)$ is not predicted within finite-size scaling, but it is not difficult to determine its asymptotic behavior. In fact, for $(t/t^*) \gg 1$ the specific heat should behave as in a bulk system. In this case

$$F(t/t^*) = \text{const} \approx 1.$$

In the limit $(t/t^*)\ll 1$, the specific-heat anomaly is suppressed at the characteristic temperature $t^*(L)$, i.e., $A^{\pm}|t|^{-\alpha}F(t/t^*)=A^{\pm}|t^*|^{-\alpha}$, and, therefore,

$$F(t/t^*)\approx(t/t^*)^{\alpha}.$$

Thus, the universal function must have the following asymptotes:

$$F(t/t^*)=\frac{C_p(t,L)-C_{\rm reg}}{A^{\pm}|t|^{-\alpha}}\approx\begin{cases}\text{const}, & (t/t^*)\gg 1,\\ (t/t^*)^{\alpha}, & (t/t^*)\ll 1.\end{cases} \quad (3)$$

The present work is devoted to experimental testing of the predictions of finite-size scaling for the case of specific-heat measurements in the vicinity of the lower critical mixing point of a 2,6-lutidine–water mixture placed in porous media with different values of the characteristic pore size L (the porous media were regarded as finite-size systems). Since a porous medium has a complicated geometry, the following questions arise. Does the form of the universal function $F(t/t^*)$ depend on the geometry, and to what extent can a porous medium serve as a model of a finite-size system? For this reason we verified the existence of a single function $F(t/t^*)$ defined by (3) on the basis of experimental specific-heat data and compared it with the analogous dependence obtained from the results of a numerical calculation of the specific heat of the finite three-dimensional Ising model, which is an ideal finite-size system, for various values of the number of spins L^3.

2. EXPERIMENT

The specific heat of a 2,6-lutidine–water mixture placed in a porous medium, was measured in an adiabatic microcalorimeter[11] with a cell volume equal to 0.29 cm^3. The measurements were performed with scanning at the rate of 0.3–1 K/h. The temperature in the cell was measured by a platinum resistance thermometer (the nominal resistance at the triple point of water was 100 Ω), which was positioned on an internal heat shield. The temperature of the internal shield followed the temperature of the cell and was monitored by a thermocouple.

Porous glasses prepared in the Institute of Precision Mechanics and Optics (St. Petersburg) served as the porous media. The porous glasses were obtained from a sodium borosilicate glass by leaching the sodium and boron oxides after cooling below the phase-separation temperature and the glass transition and have the form of a network of randomly connected channels. We used porous glasses with characteristic pore sizes of 1000 Å and 100 Å and specific surfaces of ≈20 m^2/cm^3 and ≈100 m^2/cm^3, respectively. Cylindrical matrices with a diameter of 5 mm and a length of 14 mm were made from the glasses so that they could be inserted tightly into the calorimetric cell. Besides the glass matrices, a porous matrix with approximately the same pore sizes was prepared from a spherical nickel powder with a narrow size distribution, which was compacted under a high pressure and had a characteristic pore size of ≈2500 Å.[1)] The volume of the pores relative to the total volume of the matrices was 67% for the 1000 Å glass matrix, 26% for the 100 Å glass matrix, and 35% for the 2500 Å nickel matrix.

TABLE I. Values of the parameters in Eq. (1) obtained from the treatment of the experimental data on the specific heat of a 2,6-lutidine–water mixture in the bulk (first row) and data from the numerical calculation of the finite three-dimensional Ising model[16] extrapolated to an infinite system.

A^+	A^-	α	D^+	D^-	B^+	B^-	E
4.67	8.20	0.111	–	−17.17	88.59	91.10	−8.83
±1.5	±2.6	±0.018	–	±2.0	±0.64	±1.0	±1.65
1.49	2.15	(0.11)	0.41	−1.89	−1.81	−0.69	–
±0.1	±0.04	–	±0.2	±0.1	±0.02	±0.08	

Before being saturated with the mixture, the matrices were washed with the solvent and dried at ≈200 °C for several hours. The porous matrices were placed in a vessel with a prepared mixture of known concentration in the single-phase state ($T<T_c$) and were impregnated by the mixture under the action of capillary forces. Specially performed measurements on the 1000 Å glass matrix showed that the matrices are saturated with a mean concentration equal to the concentration in the vessel. To perform these measurements, a mixture with a lutidine weight fraction $x\approx 0.15$ was prepared in an amount slightly greater than twice the pore volume of the matrix. The matrix was saturated with this mixture, and its phase-separation temperature was determined from the specific-heat maximum. Then the matrix was dried and saturated again with the remaining portion of the volume. In both cases the measured phase-separation temperatures coincided. This is evidence that the concentration of the mixture in the vessel remained unchanged. The 1000 Å glass matrix was usually saturated with ≈0.18 g of the mixture, and the 100 Å glass matrix and the nickel matrix were saturated with about 0.06 g.

3. RESULTS

We previously[11] measured the bulk coexistence curve of a 2,6-lutidine–water mixture and the specific heat for the critical lutidine weight fraction $x_c=0.29$ in the vicinity of the critical mixing point. The specific-heat data were approximated by Eq. (1). The critical temperature $T_c=(33.6339\pm 0.0002)$ °C was determined from the characteristic kink on the temperature dependence of the enthalpy $H(T)$ with slow scanning at a rate of $\approx 5\times 10^{-3}$ K/h in a narrow vicinity of T_c. The results of the treatment are presented in the first row in Table I. The Wegner correction in the unordered phase was insignificant. The specific-heat critical exponent coincided with the theoretical value $\alpha=0.11$.[12] The universal ratio between the coefficients in front of the asymptotic term obtained in the experiment, $A^-/A^+=1.77$, also agrees with the theoretical value of 1.82 (Ref. 13) to within the error of its determination.

Coexistence curves were measured to determine the critical concentrations of the 2,6-lutidine–water mixture in the porous media. The phase-separation temperature $T_p(x)$ for different concentrations was determined from the specific-heat maximum. In the porous glass with a pore size of 100 Å the specific-heat maximum is expressed very weakly, and the error in the determination of the phase-separation temperature is large. Therefore, we shall not

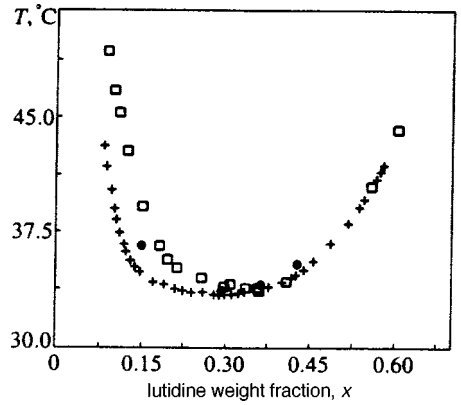

FIG. 1. Coexistence curve of a 2,6-lutidine–water mixture in the bulk (+), in the porous glass with a characteristic pore size of 1000 Å (□), and in the porous nickel with a pore size of 2500 Å (●).

present these data for quantitative comparison. Figure 1 presents the coexistence curves of a 2,6-lutidine–water mixture measured in the porous glass with a pore size of 1000 Å, in the porous nickel, and in the bulk. Two effects associated with the porous media are observed. The coexistence curves are shifted toward higher concentrations relative to the coexistence curve in the bulk. This indicates that a bound layer of lutidine forms on the surfaces of the porous matrices,[14] reducing the concentration within the pores. The formation of a lutidine layer on the surfaces of porous matrices made from Vycor glass was also observed in Ref. 4. The other effect is an increase in the temperature of the minimum on the coexistence curves in the porous media in comparison with T_c in the bulk [displacement of the temperature of the specific-heat maximum $T_m(L)$]. The following critical parameters of the mixture in the porous media were determined from the minima on the coexistence curves: the critical lutidine weight fraction $x_c \approx 0.32$ and $T_m = 33.93$ °C in the porous glass with a pore size of 1000 Å, and $x_c \approx 0.30$ and $T_m = 33.705$ °C in the porous nickel with a pore size of 2500 Å.

Figure 2 shows the behavior of the specific heat in the bulk and in the porous media in the vicinity of the critical point. As the pore size decreases, the specific-heat anomaly decreases sharply. To illustrate this, the figure shows the specific heat in the glass matrix with a characteristic pore size of ≈ 100 Å for a lutidine weight fraction $x = 0.57 \approx x_c$. The specific-heat anomaly is completely suppressed and appears as a very weak (on the level of the noise) maximum at $T \approx 45$°C (see the inset in Fig. 2), which is spread over a broad temperature range. The region of suppression of the specific-heat anomaly as a function of the pore size is clearly seen in Fig. 3, where the same data are presented as a function of $|t|^{-\alpha}$ [$\alpha = 0.11$, $t = (T - T_m(L))/T_m(L)$]. Far from the temperature of the maximum, T_m, the specific heat in the porous medium exhibits the same behavior as in the bulk and begins to deviate from the bulk behavior at a certain characteristic temperature $t^*(L)$.

For comparison, Fig. 4 presents data on the specific heat per spin,[15] which were obtained in numerical calculations of the finite three-dimensional Ising model for different values of the number of spins L^3 or, equivalently, for different values of the size of the periodic lattice L. It is seen that the behavior of the specific heat is qualitatively the same as that which we observed experimentally.

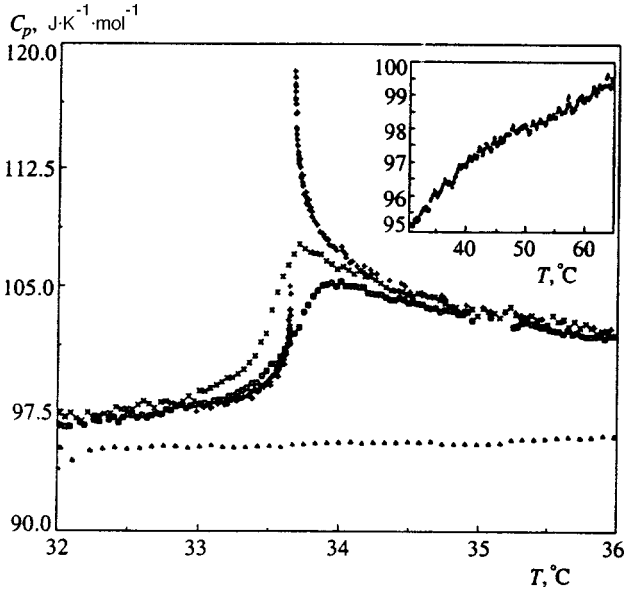

FIG. 2. Specific heat of a 2,6-lutidine–water mixture at the critical concentrations: + — in the bulk, × — in the 2500 Å porous nickel, □ — in the porous glass with a pore size of 1000 Å, and ▲ — in the porous glass with a pore size of 100 Å. The inset presents the specific heat in the porous glass with a pore size of 100 Å over a broad temperature range.

To construct the universal function $F(t/t^*)$, in accordance with Eq. (3), the regular part of the specific heat and the nonasymptotic Wegner correction obtained from treating the specific heat in the bulk (see Table I) were subtracted from the experimental values of the specific heat in the porous medium. It was assumed here that the regular part of the specific heat does not depend on the characteristic pore size.

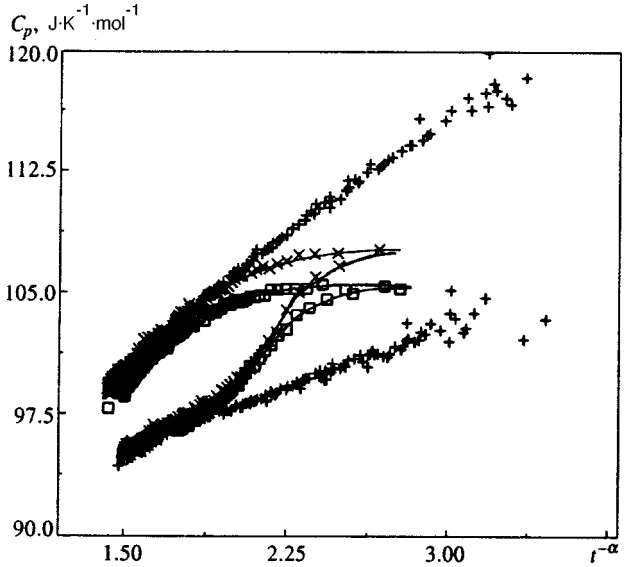

FIG. 3. Specific heat of a 2,6-lutidine–water mixture at the critical concentrations in the bulk (+), in the porous nickel with a pore size of 2500 Å (×), and in the porous glass with a pore size of 1000 Å (□) as a function of $t^{-\alpha}$ with $\alpha = 0.11$. The solid lines were drawn for better visualization.

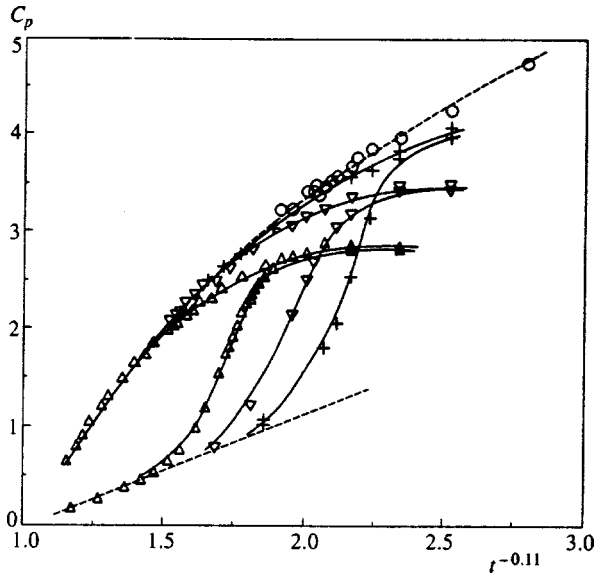

FIG. 4. Numerical calculation of the specific heat per spin in the finite three-dimensional Ising model for various values of the lattice size L (Ref. 15) as a function of $t^{-\alpha}$ ($\alpha=0.11$): $L=32$ (△), 64 (▽), 128 (+), and 256 (○). The solid lines were drawn for better visual perception.

The values of the coefficients $A^{\pm}$ and the critical exponent α were also taken from the treatment of the specific heat in the bulk. Figure 5 presents the plots of F thus obtained for the unordered and ordered phases in logarithmic coordinates. The characteristic temperature t^* served as a fitting parameter. In addition, the values of the specific heat in the porous medium were shifted relative to the specific heat in the bulk by a constant value due to the difference between the critical concentrations in the porous medium and in the bulk (the different molecular weights).

The figures also present plots of $F(t/t^*)$ constructed by a similar method from the results of a numerical calculation of the specific heat for the finite three-dimensional Ising model.[15] As in the case of the treatment of the experimental data, this required determination of the regular part of the specific heat and the coefficients $A^{\pm}$ and $D^{\pm}$. The dashed curves in Fig. 4, which are the envelopes of the specific-heat curves for different L, represent the behavior of the specific heat in the ordered (upper line) and unordered (lower line) phases in the three-dimensional system. The points lying on the envelopes were approximated by Eq. (1) without the linear term in the regular part. The critical temperature $T_c=4.511526$ obtained from the calculation and the critical exponent $\alpha=0.11$ were fixed. The results of this treatment are presented in the second row of Table I. Unfortunately, there were few data for the treatment in the unordered phase; therefore, these values of the coefficients should be regarded as approximate.

It is seen from Fig. 5 that $F(t/t^*)$ has different forms in the unordered and ordered phases. In our opinion, this is due to the asymmetric character of the behavior of the specific heat relative to the critical temperature T_c. The amplitude of the fluctuational part of the specific heat in the ordered phase A^- is approximately two times greater than the amplitude of the specific heat in the unordered phase A^+. Therefore, the specific heat reaches different values in the ordered and unordered phases at equal distances from T_c, while remaining finite. This results in the formation of a transition region from a small value of the specific heat to a large value with a specific-heat maximum at $t_m(L)$, which is always shifted in the direction of the ordered phase, and an additional contribution to the specific heat of the unordered phase (see Fig. 2). The additional contribution is manifested in the form of a ''hump'' on the plot of $F(t/t^*)$.

4. DISCUSSION

It follows from the treatment of the experimental data that there is a single dependence of

$$\frac{C_p(t,L)-C_{\rm reg}}{A^{\pm}|t|^{-\alpha}}=F(t/t^*)$$

on the dimensionless parameter (t/t^*), which describes the behavior of the specific heat in finite systems. In addition, the analogous function constructed on the basis of the data on the specific heat of the finite three-dimensional Ising model with different values for the number of spins L^3 is consistent with the experimental dependence. This allows us to state that $F(t/t^*)$ is a universal function and does not depend on the medium modeling a finite system.

The asymptotic behavior of the universal function obtained corresponds to the predictions of finite-size scaling theory [Eq. (3)]. In fact, as follows from Fig. 5, for $t/t^* \gg 1$ the function $F(t/t^*)$ reaches a constant value $F(t/t^*)$

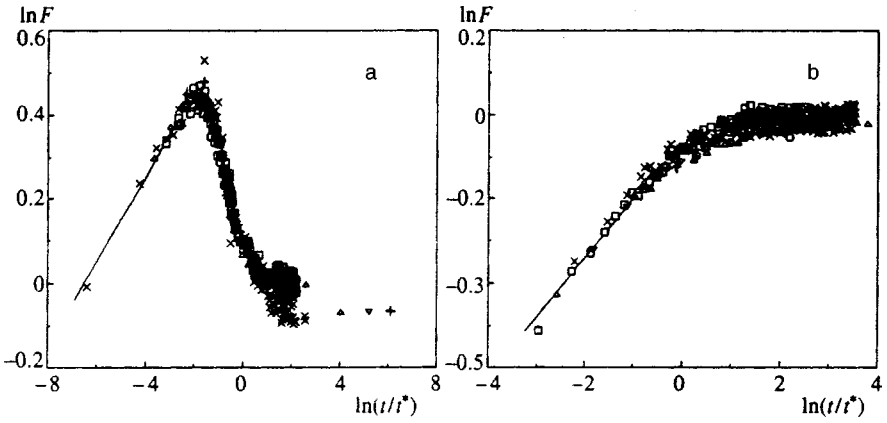

FIG. 5. The universal function $F(t/t^*)$ in unordered (a) and ordered (b) phases in logarithmic coordinates. The function was obtained from the experimental data on the specific heat of a 2,6-lutidine–water mixture in the 2500 Å porous nickel (×) and the 1000 Å porous glass (□) and on the basis of a numerical calculation of the specific heat for the finite three-dimensional Ising model with the lattice sizes $L=32$ (△), 64 (▽), 128 (+), and 256 (○).

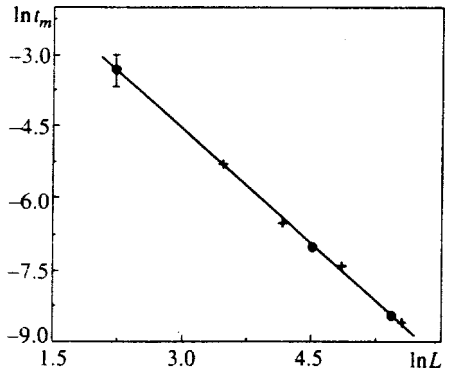

FIG. 6. Logarithmic dependence of the temperature of the specific-heat maximum in a finite system (relative to the critical temperature in the bulk) $t_m=|(T_c-T_m(L))/T_c|$ on the system size: ● — experimental data, + — data from the numerical calculation of the three-dimensional Ising model. To facilitate comparisons, the experimental data have been displaced along the x axis by a certain constant value. The slope of the straight line corresponds to $\nu=0.64\pm0.02$.

≈ 1, which corresponds to the behavior of the specific heat in the bulk. In the other limit $t/t^*\ll 1$, $F(t/t^*)$ is a power function: $F(t/t^*)\propto(t/t^*)^{\alpha}$. In addition, the slope of the solid lines in Fig. 5 corresponds to $\alpha=0.11\pm0.02$.

A plot of the universal function $F(t/t^*)$ calculated in an ϵ expansion with $\epsilon=1$, which is visually similar to the universal function presented in Fig. 5, but in the opposite phases, was presented in Ref. 16. The character of the behavior of $F(t/t^*)$ in the ordered and unordered phases in Ref. 16 is similar to the behavior of the function $F(t/t^*)$ that we obtained in the unordered and ordered phases, respectively. Therefore, it was not possible to compare the universal function constructed on the basis of our experimental data and the data from the numerical calculation of the finite three-dimensional Ising model with the results of the calculation in Ref. 16.

Finite-size scaling theory also predicts that the displacement of the temperature of the specific-heat maximum $T_m(L)$ relative to the critical temperature in the bulk T_c is described by a power function of the size L of the system:

$$t_m(L)=|T_c-T_m(L)|/T_c\propto L^{-1/\nu}. \tag{4}$$

Figure 6 presents the experimental data on $t_m(L)$ and the data obtained from the numerical calculation of the finite three-dimensional Ising model. To facilitate comparisons, the experimental data have been displaced along the x axis in the figure by a certain constant value. It is seen from the figure that in both cases the displacement of the temperature of the specific-heat maximum $t_m(L)$ obeys the same dependence of type (4), which corresponds to $\nu=0.64\pm0.02$.

5. CONCLUSIONS

Our experimental data on the specific heat of a 2,6-lutidine–water mixture in the bulk and in porous media with different pore sizes have been used to obtain a function of the dimensionless ratio t/t^*, which describes the behavior of the specific heat in a finite system. The analogous function constructed on the basis of the data from a numerical calculation of the specific heat of the finite three-dimensional Ising model[15] coincides with the experimental function. This provides a basis to state that $F(t/t^*)$ is a universal function and does not depend on the geometry of the medium modeling a finite system. The asymptotic behavior of the universal function for $t/t^*\gg 1$ and $t/t^*\ll 1$ is consistent with the prediction of finite-size scaling theory.

The displacement of the temperature of the specific-heat maximum $T_m(L)$ in finite systems relative to the critical temperature T_c in the bulk as a function of the size L of the system has been obtained and is in good agreement with the dependence predicted by finite-size scaling.

In conclusion, we wish to thank A. Talapov for supplying us with the preliminary data from the numerical calculation of the specific heat of the finite three-dimensional Ising model. We also thank E. Gorodetskiĭ for some very useful discussions, as well as V. Kulikov, A. Muratov, and V. Podnek for some valuable comments.

This work was performed with financial support from the Russian Fund for Fundamental Research (Grant No. 96-02-18235).

[1)]The matrices of the spherical nickel powder compacted under high pressure were prepared in the Institute of High-Pressure Physics (Troitsk, Moscow Region) according to a procedure developed by V. A. Sidorov and O. B. Tsiok.

[1] J. V. Maher, W. I. Goldburg, D. W. Pohl, and M. Lanz, Phys. Rev. Lett. **53**, 60 (1984).
[2] M. C. Goh, W. I. Goldburg, and C. M. Knobler, Phys. Rev. Lett. **58**, 1008 (1987).
[3] S. B. Dierker and P. Wiltzius, Phys. Rev. Lett. **58**, 1865 (1987).
[4] P. Wiltzius, S. B. Dierker, and B. S. Dennis, Phys. Rev. Lett. **62**, 804 (1989).
[5] A. P. Y. Wong and M. H. W. Chan, Phys. Rev. Lett. **65**, 2567 (1990).
[6] A. P. Y. Wong, S. B. Kim, W. I. Goldburg, and M. H. W. Chan, Phys. Rev. Lett. **70**, 954 (1993).
[7] M. Thommes and G. H. Findenegg, Langmuir **10**, 4270 (1994).
[8] Tar-pin Chen and F. M. Gasparini, Phys. Rev. Lett. **40**, 331 (1978).
[9] M. E. Fisher and A. E. Ferdinand, Phys. Rev. Lett. **19**, 169 (1967).
[10] M. E. Fisher and M. N. Barber, Phys. Rev. Lett. **28**, 1516 (1972).
[11] L. V. Entov, V. A. Levchenko, and V. P. Voronov, Int. J. Thermophys. **14**, 221 (1993).
[12] G. A. Barker, B. G. Nickel, and D. I. Meiron, Phys. Rev. B **17**, 1365 (1978).
[13] C. Bervillier and C. Godreche, Phys. Rev. B **21**, 5427 (1980).
[14] V. M. Buleĭko, V. P. Voronov, L. V. Entov, A. R. Ramazanov, JETP Lett. **61**, 34 (1995)].
[15] A. Talapov, Private communication (1995).
[16] J. Rudnick, H. Guo, and D. Jasnow, J. Stat. Phys. **41**, 353 (1985).

Translated by P. Shelnitz

Orientational states of C_{60} molecule in crystals

V. L. Aksenov and V. S. Shakhmatov*)

Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia

Yu. A. Osip'yan

Institute for Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 2 October 1997)
Zh. Éksp. Teor. Fiz. **113**, 1081–1093 (March 1998)

Local symmetry of orientational states of the C_{60} molecule in crystals has been investigated. It was shown that the various orientational phase transitions in different crystals are related to different orientational orbits. The model of orientational phase transitions based on a sequence of orientational states with different symmetry properties has been suggested. We have found that both the local symmetry of C_{60} molecule and the symmetry of its internal vibrations become higher after a reduction of crystal spatial symmetry at the phase transition. This effect is fairly common and can be observed in the orientational order–disorder phase transitions with wave vectors at the Brillouin zone boundary. Feasible manifestations of the predicted effect in various experiments are discussed. © *1998 American Institute of Physics.* [S1063-7761(98)02503-7]

1. INTRODUCTION

It is well known that phase transitions in fullerene crystals lead to changes in orientational states of the C_{60} molecule. It follows from experimental data on pristine C_{60} (fullerite) crystals[1] that in the high-symmetry $Fm3m$ phase the C_{60} molecule experiences anisotropic retarded rotation, and this rotation is stopped after the $Fm3m \rightarrow Pa3$ transition, when an orientational ordering of C_{60} molecules takes place. In AC_{60} crystals (fullerides), where A=K, Rb, a more complex $Fm3m \rightarrow Pnnm$ phase transition comes about, when the ordering of molecular orientations in the low-symmetry $Pnnm$ phase is accompanied by a molecule strain and, as a result of these two effects, covalent bonds between neighboring C_{60} molecules are formed.[2] In both cases, the order parameter of these phase transitions should describe the relation between the averaged orientational states of the C_{60} molecule in the high- and low-symmetry phases and the orientational ordering of molecules.

Presently there is no universal theory of changes in orientational states of the C_{60} molecule in crystals. The theoretical description of orientational phase transitions in fullerene crystals is based on two techniques, namely continuous and discrete, each of which provides a good description of C_{60} molecule orientational states in only one of two phases. For example, the continuous technique[1] has been applied to C_{60} rotating molecules in the $Fm3m$ phase, when the averaged anisotropic distribution of C_{60} molecular orientation is expanded in terms of symmetrized spherical functions transformed according to irreducible representations of the icosahedral (Y_h) or cubic (O_h) point groups. In this approximation, however, it is difficult to describe both the specific orientation of a C_{60} molecule in the low-symmetry phase and strain of the C_{60} molecule due to the phase transition in AC_{60} fullerides.

The discrete technique applied to an orientational phase transition in fullerites is based on discrete orientational states of a rotating C_{60} molecule[3] (see also Ref. 1). In this case, the orientational state of a C_{60} molecule in the low-symmetry phase is described exactly, whereas the description of a rotating molecule in the high-symmetry phase is approximate, namely, its state is presented in the form of a set of several discrete orientational states. Since the discrete technique exactly describes the orientational states of C_{60} molecules in the low-symmetry phase of the crystal, it is clear that all physical properties of the low-symmetry phase affected by an orientational phase transition can be also described in terms of this method. For example, a generalization of the discrete technique[4] was applied to describe the strain of the C_{60} molecule at the phase transition in the fulleride AC_{60} crystal.

In this paper, we demonstrate that some physical properties of the high-symmetry phase can be also described in terms of the discrete method because the various discrete orientational states of the C_{60} molecule have differing intrinsic local symmetries. Therefore physical phenomena in the high-symmetry phase related to specific discrete orientational states can be classified in accordance with symmetry properties of orientational states.

Section 2 describes all orientational states of the C_{60} molecule in the $Fm3m$ phase with differing local symmetry. The phase transition will be discussed in Sec. 3 in terms of orientational states, and distributions of the atomic density of a rotating C_{60} molecule will be constructed in Sec. 4. The local symmetry and the symmetry of internal vibrations of the C_{60} molecule in a crystal will be discussed in Sec. 5. In Sec. 6 we will discuss possible experimental tests of pre-

TABLE I. Symmetry elements of free C_{60} molecule and its position in crystal lattice of the $Fm3m$ cubic phase.

Symmetry elements	C_5	C_4	C_3	C_2	I
C_{60} molecule	6	–	10	15	1
Position with O_h symmetry in lattice	–	3	4	6	1

dicted types of orientational states using X-ray and neutron diffraction, along with manifestations of increasing symmetry of C_{60} molecule internal vibrations at the phase transition in light-scattering experiments.

2. SYMMETRY OF C_{60} MOLECULE ORIENTATIONAL STATES

In order to determine the types of orientational states, let us consider various embeddings of the C_{60} molecule, which has icosahedral symmetry, in a cubic crystal lattice.

The C_{60} molecule has the point symmetry group Y_h, which contains six five-fold rotation axes (C_5), ten triple rotation axes (C_3), fifteen double rotation axes (C_2), and inversion (I), all listed in Table I. The C_{60} molecule occupies in the crystal lattice the position with local symmetry O_h. The point group O_h has three four-fold rotation axes (C_4), four triple rotation axes, six double axes, and inversion.

By comparing the sets of symmetry axes contained in O_h and Y_h symmetry groups, one can infer that in order to have only one four-fold symmetry axis, which is the case in cubic crystals, the C_{60} molecule should occupy either two discrete orientational states if the double rotation axis of the C_{60} molecule is aligned with one of $\langle 100\rangle$ axes of the crystal lattice, or four discrete states in the case of a different orientation of the molecule. If the double rotation axis of the C_{60} molecule is aligned with a $\langle 100\rangle$ crystal axis, it can be rotated around its double axis so that its triple axes coincide with $\langle 111\rangle$ crystal axes. This is the case of the highest local symmetry of the C_{60} molecule embedding in the cubic crystal lattice (Fig. 1a). This embedding generates two so-called standard orientations, which can be transformed to one another by rotating the molecule around any $\langle 100\rangle$ crystal axes through an angle of 90°. A C_{60} molecule occupying a position with the O_h symmetry in the standard orientation in the lattice of the $Fm3m$ phase has the local symmetry $T_h = O_h \cap Y_h$.[5] This is the case in the A_3C_{60} crystal, where C_{60} molecules are randomly distributed between the two standard orientations.[6]

Now let us consider the situation when only one of the triple axes of C_{60} coincides with one of $\langle 111\rangle$ axes of the cubic crystal lattice (Fig. 1b). With due account of the inversion symmetry, one can see that the six symmetry elements of the point group $S_6 = (E, C_3, C_3^2) \times (E, I)$ are common elements of both the C_{60} molecule and point group O_h, where E is the identity element of the symmetry group. The local symmetry group in this case is S_6. Since the point group O_h has 48 different symmetry elements, the number of differing orientational states of the C_{60} molecule in this embedding is $8 = 48/6$. These eight orientational states were used to describe the phase transition in the C_{60} fullerite.[3]

If a double axis of the C_{60} molecule is aligned with the [110] axis in the cubic lattice (Fig. 1c), the common symmetry elements are $(E, C_2) \times (E, I)$, so the number of different orientational states of the C_{60} molecule is $12 = 48/4$. These twelve orientational states of the C_{60} molecule were used in a description of the phase transition in the AC_{60} fulleride.[4] The local symmetry group for the C_{60} molecule is C_{2h}. Note that, if a double axis of the molecule is aligned with the [100] crystal direction and no triple axis of the molecule is aligned with the $\langle 111\rangle$ crystal axes, a different twelve-fold orientational basis is generated (Fig. 1d). Below these two orientational bases are denoted as 12(a) and 12(b), respectively.

In the case when none of the C_{60} symmetry axes coincides with the corresponding symmetry elements of the crystal lattice, the only common symmetry elements are (E, I); hence the number of different orientational states of the C_{60} molecule is $24 = 48/2$, and the local symmetry group of the molecule is I (Fig. 1e).

Orientational states of one basis have the same continuous degrees of freedom. For example, the rotation angle around one triple axis in the case of the eight-fold basis is arbitrary. The situation is similar in the case of a twelve-fold basis, but here an arbitrary rotation is allowed around a double axis. In the case of the 24-fold basis two independent rotations are possible. Note that at some rotation angles an additional degeneracy of orientational states of the C_{60} molecule may occur because some symmetry axes of the C_{60} molecule can coincide with symmetrical directions of the lattice. For example, in the eight-fold basis double symmetry axes can coincide with $\langle 100\rangle$ directions in the lattice at some rotation angle around the triple axis. This brings about a

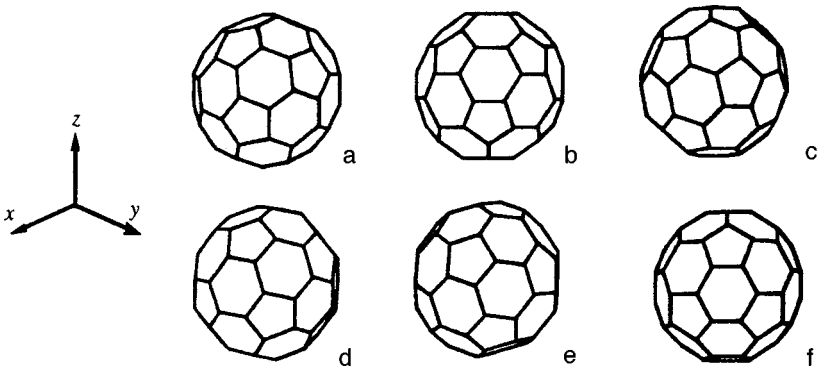

FIG. 1. Initial orientations of C_{60} molecule in various orientational bases: (a), (b), (c), (d), (e), or (f) is one of the orientational states of the C_{60} molecule in 2-, 8-, 12(a)-, 12(b)-, 24-, or 6-fold bases.

TABLE II. Different types of orientational bases for C_{60} molecule in the $Fm3m$ phase.

Multiplicity of orient. basis	Local symmetry of orient. state	Crystal	Note
2	T_h	A_3C_{60}	Static disorder
6	D_{2h}		
8	S_6	C_{60}	Dynamic disorder
12(a)	C_{2h}	AC_{60}	Dynamic disorder
12(b)	C_{2h}		
24	I		

TABLE III. Orientational states of C_{60} molecules in different phases of C_{60} and AC_{60} crystals.

Crystal	High-symmetry phase $T \gg T_c$	High-symmetry phase $T > T_c$	Low-symmetry phase $T < T_c$
C_{60} fullerite	{2,6,8,...}	{8}	8
AC_{60} fulleride	{2,6,8,...}	{12}	12

degeneracy of orientational basis functions, and we have only two standard orientations instead of eight.

In the two twelve-fold bases the number of independent orientational functions reduces to six when the double symmetry axis of the C_{60} molecule coincides with the $\langle 100 \rangle$ crystal direction in the 12(a) basis and with the $\langle 110 \rangle$ crystal direction in the 12(b) basis (Fig. 1f). In this case, the local symmetry of orientational states is determined by point group D_{2h}. It is clear that all of the above special positions can be also derived from the 24-fold basis. Note also that the orientational states of the two- and six-fold bases are most symmetrical and, unlike orientational states of other bases, have no rotational degrees of freedom.

Table II lists different types of orientational bases and materials (fullerite C_{60} and fulleride AC_{60}) whose orientational phase transitions are described in terms of these states. Note that no orientational phase transition occurs in A_3C_{60}, and static disorder in the distribution of C_{60} molecules between two standard orientations was experimentally detected in the $Fm3m$ phase.[6]

Functions of orientational states referred to a single basis, which transform into one another under rotational symmetry elements of group $Fm3m$, form an orbit (in terms of group theory[7]) — in this specific case, an orientational orbit.

3. ORIENTATIONAL PHASE TRANSITION

Diffraction studies[1] of fullerite C_{60} in phase $Fm3m$ indicate that the orientation of the rotating C_{60} molecule in the crystal has an anisotropic distribution. Consequently, different orientational states of the C_{60} molecule should be occupied with different probabilities. It was shown in the previous section that orientational states can be combined into orientational orbits with different local symmetry properties (Table II). Since the orientational states of one orbit are physically equivalent, these states are occupied by C_{60} molecule with equal probabilities, so the anisotropy of the molecule orientation distribution implies that states of different orbits are occupied by a rotating C_{60} molecule with different probabilities given by

$$P_n = \exp(-V_n/kT) / \sum_m \exp(-V_m/kT), \quad (1)$$

where k is Boltzmann's constant, V_n is the potential energy of state n of the orientational orbit, and the summation is performed over all orientational states of all orientational orbits.

Thus, we can suggest the following model of the orientational phase transition in fullerene crystals. At high temperatures ($T \gg T_c$, where T_c is the phase transition temperature), a rotating C_{60} molecule occupies all orientational states of all orientational orbits. At temperatures close to the phase transition ($T > T_c$) the rotating C_{60} molecule is in the states of one orientational orbit most of the time, namely on the orbit corresponding to a specific orientational crystalline potential of a C_{60} molecule in a specific crystal. At the transition point the orbit is ''frozen out,'' and the molecule occupies one of the orientational states of that orbit.

By drawing the analogy between the orientational phase transition and structural phase transition associated with displacements, we come to the following conclusions. The set of all states belonging to all orientational orbits is an analog of all possible displacements of all atoms in a crystal cell. Each orientational orbit can be interpreted as an analog of displacements affecting only the atoms associated with the soft mode at the phase transition. Thus, different orientational orbits are microscopic realizations of order parameters related to different phase transitions.

Table III lists orientational orbits and orientational states necessary for describing orientational phase transitions in fullerite C_{60} and fulleride AC_{60}.

Table III clearly illustrates how the orientational transition proceeds. At higher temperatures a rotating C_{60} molecule occupies in a C_{60} (AC_{60}) crystal all orientational states of all orientational orbits (the twofold orbit, sixfold orbit, etc.). Near the phase transition point ($T > T_c$) the molecule is in states of the eightfold(twelvefold) orbit most of the time, and in the low-symmetry phase the molecules are orientationally ordered in one of the states of the eightfold(twelvefold) orbit.

4. DISTRIBUTION OF ATOMIC DENSITY OF A ROTATING C_{60} MOLECULE

It has been shown in the previous section that a rotating C_{60} molecule is in orientational states of one orbit most of the time in the $Fm3m$ phase near the phase transition point, therefore the distribution of C_{60} molecule orientations in a crystal is determined by states of only one orbit. Let us select a certain orientational state; then, in this state ψ_1, the atomic density on the surface of molecule C_{60} at point $\mathbf{\Omega}$ defined by polar angles θ and φ in the coordinate system shown in Fig. 1 has the form

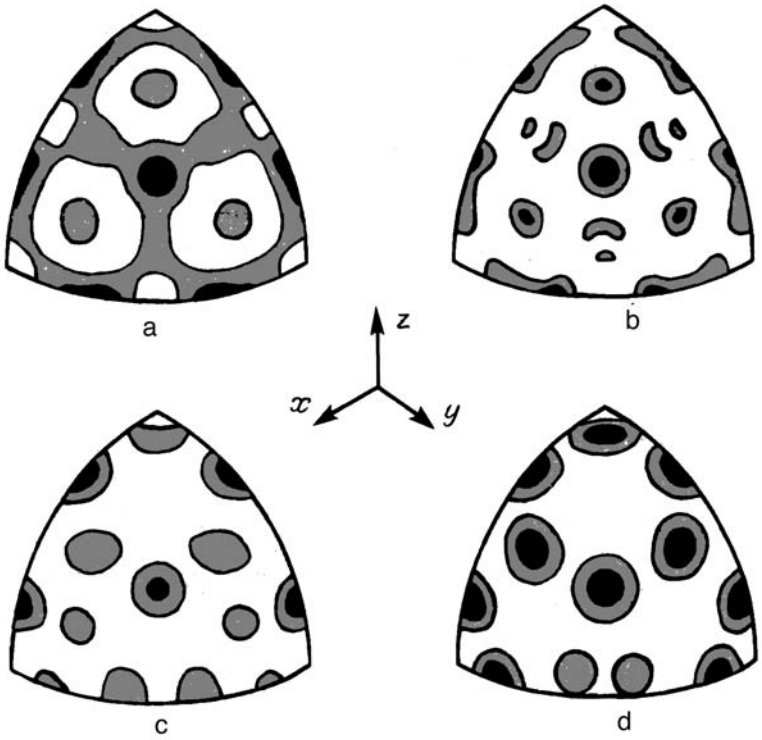

FIG. 2. a) A fragment (one-eighth of a molecule) of atomic density distribution of a rotating C_{60} molecule in $Fm3m$ phase. Dark regions correspond to the maximum densities (from 100 to 60%), lighter areas correspond to densities ranging between 66 and 33%, and the lightest areas to the range of 33 to 0%. Orientational states have a Gaussian distribution with a standard deviation of 5°. b, c, d) the same as (a), but with a standard deviation of 3°. Moreover, in (c) and (d) the crystals are deformed: (c) compression strain $e_{xx}=e_{yy}=-2e_{zz}\neq 0$; d) the same compression as (c) plus additional shear strain $e_{xy}\neq 0$.

$$\rho_1(\mathbf{\Omega})=\sum_n \delta_1(\mathbf{\Omega}-\mathbf{\Omega}_n), \tag{2}$$

where δ is the Dirac delta function and summation is performed over the 60 carbon atoms in the molecule. By taking the rest of the orientational states into account and modeling thermal motion of the C_{60} molecule using a normal distribution with standard deviation σ, we can express the mean atomic density $\langle\rho(\mathbf{\Omega})\rangle$ in the form

$$\langle\rho(\mathbf{\Omega})\rangle = C\int dS'\ \exp[-(\mathbf{\Omega}-\mathbf{\Omega}')^2/2\sigma^2]\sum_k \rho_k(\mathbf{\Omega}')$$

$$= \int dS'\ \exp[-(\mathbf{\Omega}-\mathbf{\Omega}')^2/2\sigma^2]$$

$$\times\sum_{k,n}\delta_k(\mathbf{\Omega}'-\mathbf{\Omega}_n), \tag{3}$$

where C is a numerical constant, the integration runs over the C_{60} molecule surface, $|\mathbf{\Omega}-\mathbf{\Omega}'|$ is the distance between points $\mathbf{\Omega}$ and $\mathbf{\Omega}'$ on the molecule surface and is expressed in degrees, and summation is performed over orientational states of the orbit (k) and the 60 positions of carbon atoms in the C_{60} molecule (n).

Figure 2 shows distributions of atomic density of a rotating C_{60} molecule in an AC_{60} crystal in the phase $Fm3m$. The functions were calculated by Eq. (3) with orientational states of the twelve-fold orbit. The initial orientational state ψ_1 is that shown in Fig. 1c. The other orientational states of the molecule were obtained by applying the following rotations: $\psi_2=C_4\psi_1$, $\psi_3=C_4^2\psi_1$, and $\psi_4=C_4^3\psi_1$. The rest of the orientational states were obtained by applying rotations C_3 and C_3^2 in accordance with the following symbolic scheme: $(\psi_5,\psi_6,\psi_7,\psi_8)=C_3(\psi_1,\psi_2,\psi_3,\psi_4)$ and $(\psi_9,\psi_{10},\psi_{11},\psi_{12})=C_3^2(\psi_1,\psi_2,\psi_3,\psi_4)$, where C_4 denotes a rotation around the z-axis through an angle of 90° and C_3 is a rotation around the [111] axis through an angle of 120°.

The symmetry analysis of the phase transition in fulleride AC_{60} indicates[8] that the low-symmetry phase is a result of condensation of one component of the sixfold order parameter; hence, we should expect emergence of six domains in the low-symmetry phase. The distribution of atomic density (Fig. 2a and 2b) is constructed on the basis of all twelve orientational states and therefore is connected with all six domains. In a deformed crystal, however, the twelve orientational states are no longer physically equivalent, so the probabilities of detecting a molecule in different states are unequal. The symmetry analysis of the phase transition[4,8] allows one to make out which of the twelve orientational states refer to a specific deformation pattern. The atomic density distribution due to external stress $\sigma_{xx}=\sigma_{yy}=-2\sigma_{zz}\neq 0$ is given in Fig. 2c. It is based on orientational states ψ_1, ψ_2, ψ_3, and ψ_4. By applying an additional shear stress $\sigma_{xy}\neq 0$, one can further limit orientational degrees of freedom of the C_{60} molecule. The atomic density distribution constructed from the two orientational states ψ_1 and ψ_3 for this case is given in Fig. 2d.

5. LOCAL SYMMETRY AND SYMMETRY OF INTERNAL VIBRATIONS OF THE C_{60} MOLECULE IN CRYSTAL LATTICE

It was shown in the previous sections that a rotating C_{60} molecule occupies all orientational states of all orbits, but near the phase transition point (in phase $Fm3m$) the molecule is in states of only one orbit most of the time. As a result, the atomic density distribution is anisotropic. The local symmetry of a rotating C_{60} molecule in a crystal lattice is determined by the state of lowest symmetry occupied by the

TABLE IV. Local symmetry of C_{60} molecule in different phases of C_{60} fullerite and AC_{60} fulleride.

Crystal	$Fm3m$ phase	$Pa3$ phase	$Pnnm$ phase
C_{60} fullerite	I	S_6	–
AC_{60} fulleride	I	–	C_{2h}

molecule. The local symmetry of the C_{60} molecule in the high-symmetry phase $Fm3m$ is therefore described by the inversion group I. On the other hand, orientations of C_{60} molecules in the low-symmetry phase are ordered, and they occupy one state of a specific orientational orbit. For example, C_{60} molecules in the fullerite in the low-symmetry phase $Pa3$ occupy the state with local symmetry S_6, and in fulleride AC_{60} in low-symmetry phase $Pnnm$ the state with local symmetry C_{2h}. Thus, as a result of a phase transition which lowers the crystal spatial symmetry, the local symmetry of a C_{60} molecule becomes higher (Table IV).

Further, let us consider the phonon symmetry in phase $Fm3m$. There are two possibilities, depending on the phonon energy. If the phonon frequency is higher than the characteristic rotation frequency of the C_{60} molecule, the phonon moves through the crystal with different orientations of C_{60} molecules and, as a result of slow rotation of molecules, the symmetry for the phonon is considerably lower than one can expect for the $Fm3m$ spatial group. On the other hand, the symmetry for a phonon with a frequency considerably lower than the characteristic rotation frequency of C_{60} molecule should correspond to the spatial group $Fm3m$.

Let us consider as an example internal vibrations of the C_{60} molecule in C_{60} and AC_{60} crystals, which can be investigated in Raman scattering experiments. It is well known that in IR absorption and Raman experiments phonons with wave vectors about zero are detected, and their symmetry is determined by the wave-vector group with $\mathbf{k}=0$, i.e., $G_{\mathbf{k}=0}$, or in other words, by the point symmetry group (local, point and other symmetry groups in crystals are described in Ref. 9). Primitive cells of C_{60} and AC_{60} crystals in the $Fm3m$ phase contain only one rotating C_{60} molecule, which occupies the site with local symmetry O_h. Therefore it seems feasible that internal vibrations of the C_{60} molecule in the $Fm3m$ phase can be analyzed on the base of icosahedral group Y_h (Refs. 10 and 11) or point group O_h (Refs. 10 and 11) if splitting of internal vibrations by the crystal field must be taken into account. But, since the C_{60} molecule has no fourfold axis (see Sec. 2), the highest local symmetry group of the C_{60} molecule in the crystal lattice can be only $T_h = O_h \cap Y_h$. This is one of the reasons why the symmetry analysis of internal vibrations of the C_{60} molecule in the $Fm3m$ phase was initially based on irreducible representations of T_h group.[12]

The analysis of orientational states of the C_{60} molecule described above indicates that internal vibrations of the molecule can be classified in accordance with irreducible representations of the group T_h only when the C_{60} molecule occupies states of the double orientational orbit. In C_{60} and AC_{60} crystals the orientational crystal potential is such that eight- and twelvefold orbits, respectively, must be considered, whose local symmetry is much lower, namely S_6 and C_{2h}. Since the frequencies of internal vibrations are considerably higher (by a factor of 10 to 100) than the characteristic rotation frequency of the C_{60} molecule, the corresponding phonons travel through the crystal, in which the C_{60} molecules appear statically disordered. The point symmetry for such phonons will be even lower. Thus, the symmetry of internal vibrations of the C_{60} molecules in the $Fm3m$ phase will be no higher than S_6 and C_{2h} for C_{60} and AC_{60} crystals, respectively. Since the point groups of low-symmetry phases for these crystals are T_h and D_{2h}, respectively, we come to the conclusion that the symmetry of internal vibrations of the C_{60} molecule becomes higher when the spatial symmetry decreases as a result of the orientational phase transition.

6. RESULTS AND DISCUSSION

Let us briefly summarize the results of this work. The model of the orientational phase transition in fullerene crystals based on the discrete approach has been suggested. All orientational states of a rotating C_{60} molecule have been classified in accordance with their symmetry properties. All orientational states have been attributed to different orientational orbits with different local symmetry properties. Specific orientational orbits are related to specific orientational phase transitions; moreover, orientational states of one orbit are analogs of atomic displacements that give rise to soft modes at structural phase transitions.

The distribution of atomic density of a rotating C_{60} molecule in the $Fm3m$ phase is controlled by orientational states of one orbit. The atomic density distribution is related to the development of the orientational phase transition in accordance with the scheme described in Sec. 3.

The analysis of the symmetry of orientational states of C_{60} molecules has led us to conclude that the local symmetry of the C_{60} molecule and the symmetry of internal vibrations of the molecule in a crystal lattice are higher after a phase transition to a state with lower spatial symmetry.

Now let us discuss possible experimental tests of our conclusions. The distributions of atomic density of a rotating C_{60} molecule calculated in the previous sections (Fig. 2) are directly related to diffraction spectra recorded in experiments. For example,[13] the intensities of Bragg peaks are determined by specific configurations of atoms in the crystal cell. In the case under consideration, one primitive cell contains one rotating C_{60} molecule. The intensities of Bragg peaks are determined by the structural factor squared:

$$F(\mathbf{Q}) = b_c \Bigg\{ 60 c_{\text{con}} \exp[-W_1(\mathbf{Q})] J(QR) + c_{\text{disc}} \exp[-W_2(\mathbf{Q})] \int d\mathbf{r}' \left\langle \sum_{k,n} \delta_k(\mathbf{r}' - \mathbf{r}_n) \right\rangle \times \exp(-i\mathbf{Q}\mathbf{r}') \Bigg\}, \tag{4}$$

where b_c is the neutron coherent scattering length on the carbon nucleus (or the atomic form factor in the case of

X-rays), the continuous and discrete weight factors, c_{con} and c_{disc}, are fit parameters satisfying the condition $c_{con}+c_{disc}=1$; in the case of fullerite C_{60} we have $c_{con}=0.31$ and $c_{disc}=0.69$[1] $J(QR)=\sin(QR)/(QR)$ is the Bessel function, R is the radius of the C_{60} molecule, and $W_1(\mathbf{Q})$ and $W_2(\mathbf{Q})$ are the Debye–Waller factors for the carbon atom in C_{60}. As a result of various averaging procedures, the factors $W_1(\mathbf{Q})$ and $W_2(\mathbf{Q})$ can be different. Equation (4) clearly shows that the structural factor contains $\langle\Sigma_{k,n}\delta_k(\mathbf{r}'-\mathbf{r}_n)\rangle$, which describes the anisotropy of the mean atomic density of C_{60}, and which was used in Sec. 4.

Note also that a rotating C_{60} molecule occupies orientational states of all orbits, but only one orbit is associated with an orientational phase transition. Therefore summation over k in $\langle\Sigma_{k,n}\delta_k(\mathbf{r}'-\mathbf{r}_n)\rangle$ is performed over the states of this orbit, and the effect of orientational states of other orbits is described by the isotropic Bessel function $j(QR)$.

Thus, the specific distributions given in Fig. 2 will determine the intensities of the Bragg peaks in the $Fm3m$ phase; hence, neutron and x-ray diffraction patterns can provide experimental tests of the theoretical atomic density distributions of a rotating C_{60} molecule.

It is known that the symmetry of crystal lattice vibrations can be determined using Raman scattering. Horoyski *et al.*[14] recorded Raman spectra of fullerite C_{60}. These spectra indicate that the lines recorded at $T=259$ K in the low-symmetry $Pa3$ phase have more clearly defined features than the same lines recorded at a temperature only two degrees higher, in the high-symmetry $Fm3m$ phase.

The large width of lines attributed to internal vibrations in C_{60} molecules in the $Fm3m$ phase can be ascribed to various physical causes. In pure C_{60} crystals, however, molecules are neutral, and this fact limits the range of possible causes. (For example, a charged molecule C_{60}^{-1} (or C_{60}^{-3}) can transfer as a result of the Jahn–Teller effect to a deformed state, which leads to splitting of internal molecular vibrations. Moreover, charged C_{60} molecules can generate macroscopic electric fields in a dielectric crystal, which leads to splitting of transverse and longitudinal optical phonons.) Therefore the large line widths in fullerite C_{60} can be attributed to either anharmonic interactions between internal vibrations of C_{60} molecules and other phonons, or the low local symmetry of the C_{60} molecule in the crystal lattice (splitting by the crystal field) and/or interaction between randomly aligned molecules (as in so-called Davydov splitting[15]).

Unlike the anharmonic interaction, the latter two effects determine the point group, which can be checked out in light scattering experiments with polarization analysis.

The symmetry analysis[10] indicates that Raman-active internal vibrations in the C_{60} molecule have symmetry $2A_g\oplus 8H_g$. Therefore Raman spectra should contain ten different lines, two of which have symmetry A_g (nondegenerate vibrations) and eight (fivefold degenerate) symmetry H_g. A crystal field of cubic symmetry lifts the degeneracy of vibrations with symmetry $H_g=E_g\oplus F_g$, where the vibration with symmetry F_g is triply degenerate and the vibration with symmetry E_g is doubly degenerate. Experimental results demonstrate[14] that such splitting really occurs for the vibrations $H_g(1)$ and $H_g(2)$ ($T=261$ K). Similar splitting probably takes place in the modes $H_g(6)$ and $H_g(7)$.[14] Measurements of polarized light scattering[14] indicate that vibration $H_g(1)$ with a lower frequency of 266.2 cm^{-1} has symmetry E_g, and the vibration with a frequency of 272.4 cm^{-1} has symmetry F_g. The situation is opposite in the case of the $H_g(2)$ vibration: the mode with a frequency of 430.3 cm^{-1} has symmetry F_g and that with a frequency of 434.3 cm^{-1} has symmetry E_g.

Since vibrations with symmetry A_g are nondegenerate, the large width of their lines can be due to interaction between neighboring molecules with different orientations. In this case, the strongest interactions between neighboring C_{60} molecules with different orientations in the $Fm3m$ phase (the first coordination sphere contains 12 molecules) can be modeled by defining larger lattice cells containing more than one C_{60} molecule. All such lattices with larger cells can be constructed from the $Fm3m$ phase using the symmetry analysis of all possible phase transitions that give rise to larger primitive cells. (In the cluster approximation, if interaction between neighboring C_{60} molecules is taken into account, wide Raman lines of symmetry A_g can be interpreted as sets of separate lines (one to thirteen) of different intensities, which are proportional, generally speaking, to the length of time for which C_{60} molecules are in the corresponding orientational states.)

Thus, the large widths of A_g lines can be attributed to interaction between neighboring C_{60} molecules with different orientations, i.e., the Davydov splitting,[15] given that the point group of the crystal with larger primitive cells can have a lower symmetry owing to uncorrelated orientations of neighboring molecules.

Now let us discuss the width of the F_g and E_g lines. The symmetry analysis of orientational states of C_{60} molecules in C_{60} and AC_{60} crystals given above indicates that most of the time, the C_{60} molecule is in states with symmetry S_6 and C_{2h}, respectively. In this case, owing to the low symmetry of the crystal field, the degeneracy of triply degenerate F_g vibrations is lifted. Moreover, as in the case of A_g lines, the Davydov splitting caused by interaction between neighboring molecules with different orientations takes place.

Thus, the symmetry of internal vibrations of C_{60} molecules in C_{60} and AC_{60} crystals cannot be higher than S_6 and C_{2h} because of the low local symmetry of orientational states and interaction between randomly oriented neighboring molecules. This means that vibrations of F_g symmetry, which could have been detected in the $Fm3m$ phase only in the case of orthogonal polarizations of incident and scattered light, should be also detected in the case of parallel polarizations; likewise, the vibrations of E_g symmetry should be observed in the case of perpendicular polarizations. Therefore, it would be of interest to conduct an accurate polarization analysis of F_g and E_g vibrations in fullerite C_{60} ($Fm3m$ phase) in measurements of Raman scattering.

This work is a part of the *Fullerenes and Atomic Clusters* Russian State program, grant No. 97032, and was supported by the INTAS–RFBR grant No. 95-0639.

*)E-mail: shank@nf.jinr.dubna.ru

[1] J. D. Axe, S. C. Moss, D. A. Neumann, in *Solid State Physics: Advances in Research and Applications*, H. E. Ehrenreich and F. Spaepen, (eds.), Academic Press, New York (1994), Vol. 48, p. 149.
[2] P. W. Stephens, G. Bortel, G. Faigel *et al.*, Nature (London) **370**, 636 (1994).
[3] K. Rapcewicz and J. Przystawa, Phys. Rev. B **49**, 13193 (1994).
[4] V. L. Aksenov, Yu. A. Ossipyan, and V. S. Shakhmatov, JETP Lett. **64**, 120 (1996).
[5] R. A. Dilanyan, O. G. Rybchenko, and V. Sh. Shekhtman, Kristallografiya **40**, 604 (1995) [Crystallogr. Rep. **40**, 553 (1995)].
[6] S. Teslic, T. Egami, and J. E. Fisher, Phys. Rev. B **51**, 5973 (1995).
[7] Yu. A. Izyumov and V. N. Syromyatnikov, *Phase Transitions of Symmetry of Crystals* [in Russian], Nauka, Moscow (1984).
[8] V. L. Aksenov, Yu. A. Ossipyan, and V. S. Shakhmatov, JETP Lett. **62**, 428 (1995).
[9] O. V. Kovalev, *Irreducible and Induced Representations and Corepresentations of Fedorov Groups* [in Russian], Nauka, Moscow (1986), p. 23.
[10] G. Dresselhaus, M. S. Dresselhaus, and P. C. Eklund, Phys. Rev. B **45**, 6923 (1992).
[11] V. N. Denisov, B. N. Mavrin, J. Ruani *et al.*, Zh. Éksp. Teor. Fiz. **102**, 300 (1992) [Sov. Phys. JETP **75**, 158 (1992)].
[12] Yu. S. Grushko, Yu. V. Ganzha, M. F. Kovalev *et al.*, Fiz. Tverd. Tela **35**, 980 (1993) [Phys. Solid State **35**, 503 (1993)].
[13] V. L. Aksenov and A. M. Balagurov, Usp. Fiz. Nauk **166**, 955 (1996).
[14] P. J. Horoyski, M. L. W. Thewalt, and T. R. Anthony, Phys. Rev. B **54**, 920 (1996).
[15] A. S. Davydov, *Solid State Theory* [in Russian], Nauka, Moscow (1976), p. 341.

Translation provided by the Russian Editorial office

Is superdense fluid hydrogen a molecular metal?

A. A. Likalter

Center of Applied Problems of Electrodynamics, Institute for High Temperatures, Russian Academy of Sciences, 127412 Moscow, Russia
(Submitted 18 July 1997)
Zh. Éksp. Teor. Fiz. **113**, 1094–1100 (March 1998)

Recent experiments on the compression of liquid hydrogen in reverberating shock waves, which indicate the transition into a metallic state at about nine times the liquid H_2 density [S. T. Weir, A. C. Mitchell, and W. J. Nellis, Phys. Rev. Lett. **76**, 1860 (1996)], have been interpreted by a microscopic percolation in a virtual molecular structure with a continuous spectrum of the electron excitations. The scaling dependence of the electron mobility on the energy above the percolation threshold has been used to qualitatively describe the electrical conductivity of fluid molecular hydrogen in the vicinity of the insulator-metal transition point. © *1998 American Institute of Physics.* [S1063-7761(98)02603-1]

Metallic hydrogen, the simplest element in the periodic system and the most abundant in the Universe, has been extensively studied as the prototype of the insulator-metal transition for over half a century.[1] The current research is mainly motivated by the significance of metallic hydrogen in astrophysics, especially for the magnetic-dynamo models of Jupiter and Saturn.[2] Although in this field the disordered phases are actually interesting, the insulator-metal transition is traditionally considered in solids, and most experiments have been done in diamond anvil cells at very low temperatures. At ninefold compression a phase transition indicated by the appearance of a strong infrared absorption band has been recovered, but the onset of metallization because of the band gap closure in solid molecular hydrogen is still a subject of controversy.[3] Recently, Weir, Mitchell, and Nellis[4] have reached almost the same densities in a fluid phase at much higher but still moderate temperatures by strongly compressing liquid hydrogen in reverberating shock waves. Under these conditions, a qualitative change of the electronic structure is revealed by strongly changing electronic properties, despite the thermal excitations which play a masking role. In the density range $(1.7–2.1)\times 10^{23}$ cm^{-3}, at temperatures in the range 2000–4000 K the electrical conductivity of shock-compressed fluid hydrogen increases by more than three orders of magnitude to 2000 $\Omega^{-1}\cdot$cm^{-1}, a value characteristic of metals. The activation energy, which occurs at lower densities, goes to zero at approximately 1.9×10^{23} cm^{-3}, marking the onset of metallization.

If even solid hydrogen is nonmetallic, the liquid can be metallic as in the case of silicon. Therefore, analysis of the insulator-metal transition in fluid hydrogen is based on the theory of disordered electron systems[5,6] rather than on the theory of crystalline solids. Recently, Ross[7] has proposed the metallization of H-atom subsystem in partially dissociated fluid hydrogen. We show that more naturally molecular hydrogen is metallized wholly.

Strong electron-ion interaction makes the structure near the insulator-metal transition in fluid hydrogen closely resemble the neutral molecular fluid. Therefore, analysis can be based on a microscopic percolation model,[8] which has been applied to expanded fluid mercury,[9] sodium-ammonia solutions,[10] and doped semiconductors Si:P.[11] The underlying idea is a virtual atomic-like structure which is retained in such systems in the transition region where the atoms are in mixed states described by the density matrix. Since classically accessible spheres of valence electrons of the neighboring atoms overlap, screening leads to admixing of free-like electron states. In the microscopic percolation model a sharp change of the electrical conductivity below the insulator-metal transition is governed by a high coupling parameter, namely, the ionization potential of atoms in expanded metals, or admixture states in ammonia solutions and semiconductors to the temperature. Applying this model to hydrogen, which consists of strongly bound, two-atomic molecules, we assume a virtual molecular structure. Otherwise, near the insulator-metal transition point we consider fluid hydrogen a molecular metal with partially free electrons in bonding quasimolecular orbitals admixed with free electron states (in contrast, nearly free metallic electrons could not bond the molecules). Two electrons of H_2 quasimolecules in a virtual molecular structure are equivalent and contribute equally to the conductivity. We show that the percolation model of the insulator-metal transition is capable to qualitatively describing the electronic properties of such a virtual molecular structure. The model of a virtual molecular structure can also shed light on a plasma phase transition, whose existence in hydrogen is still strongly questionable.[12]

Presumably, a virtual molecular structure exists at temperatures much lower than the ionization or dissociation energy of molecules and at moderately high densities, at which the molecular orbitals are not strongly overlapped. In such a structure, the transition point can be identified with a percolation threshold of the classically accessible regions of electrons in the ground-level molecular orbitals. Since electrons are mainly localized within the classically accessible region, the effective one-electron potential near the boundary of this region is close to the molecular ion potential. Thus, the boundary surface is determined by the equation

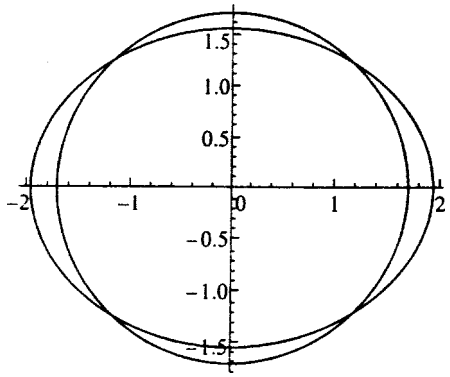

FIG. 1. Classically accessible region of the molecular orbital of hydrogen on the atomic unit scale. The mean sphere of the same volume is also shown.

$$-\frac{e^2}{2r_1}-\frac{e^2}{2r_2}=-J, \tag{1}$$

where e is the electron charge, r_1 and r_2 are the distances of an electron from the nuclei, and $J\approx 16$ eV is the vertical ionization potential of the molecules. This surface, which is close to the prolate ellipsoid with half-axes $1.55a_0$ and $1.95a_0$, where a_0 is the Bohr radius (Fig. 1), bounds almost the same volume as a mean sphere of radius

$$R_m=\frac{e^2}{J}. \tag{2}$$

The upper limit for the percolation threshold corresponds to ellipsoids with the parallel rotational axes, which coincide with those of spheres of the same volume. Furthermore, for the strongly correlated fluids like hard spheres with very thin overlapping shells,[13] the percolation threshold corresponds to the random close packing fraction,

$$\frac{4\pi}{3}R_m^3 n_m=\zeta_{\rm th}, \tag{3}$$

where n_m is the molecule number density, and $\zeta_{\rm th}\approx 0.64$. We note that Eq. (3) can be rewritten as the Edwards–Sienko correlation for the insulator-metal transitions in doped semiconductors:[14]

$$\frac{e^2 n_m^{1/3}}{J}=0.534.$$

From Eqs. (2) and (3) we obtain the number density of hydrogen molecules at the insulator-metal transition point:

$$n_m\approx 2.1\times 10^{23}\ \ {\rm cm}^{-3},$$

which corresponds to the mass density 0.7 g/cm^3. This value agrees within the experimental uncertainty with the above-mentioned estimate from the disappearance of the activation energy.[4]

The Coulomb interaction in a strongly coupled plasma above the insulator-metal transition point is characterized by the coupling parameter

$$\Gamma=\frac{z^2e^2}{R_sT}=\zeta_{\rm th}^{1/3}z^2\,\frac{J}{T}, \tag{4}$$

where $z=2$ is the total charge number of tightly bound protons, and $R_s=(4\pi n_m/3)^{-1/3}$ is the Wigner–Seitz radius of the molecular cell. Although the classical Coulomb parameter is higher than 10^2, it does not necessarily strongly influence the structure of compressed fluid determined by the repulsion between quasimolecules. Evidently, strongly repulsive exchange interaction between quasimolecules in the dense hydrogen plasma precludes a plasma phase transition, which is induced otherwise by the Coulomb attraction.

The overlap of the classically accessible regions of electrons in neighboring molecules qualitatively changes the electronic states. The number of electrons screening the nuclei can fluctuate; i.e., the nuclei are virtually screened by electrons of neighboring molecules. Therefore, in mixed quasimolecular states electrons are partially free, and the spectrum of excitations is continuous. According to the variational principle of quantum mechanics, an internal energy spectrum of a quasimolecule (i.e., without the energy of the intermolecular interactions) is bound from below by the ground energy level of the free molecule. Using molecular orbitals, we write the one-electron energy spectrum in the form

$$E_p=-J+\varepsilon_p,\quad \varepsilon_p=p^2/2m, \tag{5}$$

where ε_p is the electron excitation energy, p is the momentum of an electron far from the virtually screened nuclei, and m is the electron mass. On the other hand, the mean internal energy of the quasimolecule in a mixed state is

$$E_p=-a_{00}J+a_{pp}\varepsilon_p, \tag{6}$$

where a_{pp} and a_{00} are the diagonal elements of the density matrix normalized by the condition $a_{pp}+a_{00}=1$. From Eqs. (5) and (6), it follows that the admixture of free motion in the quasimolecular orbitals, which is determined by the ratio of the matrix elements, is

$$a_{pp}/a_{00}=\varepsilon_p/J. \tag{7}$$

Therefore, low-lying quasimolecular states, which are preferably occupied, contain only a small admixture of free motion. The lifetime and the extension of the quasimolecular states, which in the neighboring molecules can overlap, are bound by transitions of electrons between molecules.[8]

In the vicinity of the insulator-metal transition point the electric current is carried by electrons which transfer between overlapping, classically accessible spheres in a virtual molecular structure. Although below the transition point the electrons in the ground-level state are localized in finite clusters, classically accessible regions of electrons excited over a mobility gap could form an infinite cluster. At still higher energies practically the entire volume becomes classically accessible, and the electron mobility increases to the minimal gas-kinetic value. By definition, the absolute mobility gap Δ_1 is determined by the appearance of the mobility, and the soft mobility gap Δ_2 is determined by its increase to the gas-kinetic value. This two-parameter gap is

$$\Delta_k=J-e^2(4\pi n_m/3\zeta_k)^{1/3},\quad k=1,2, \tag{8}$$

where $\zeta_1\approx 0.64$ is the percolation threshold equal to the random close packing fraction, and ζ_2 $(>\zeta_1)$ is an accessible

volume fraction of the same order of magnitude. The latter parameter is of less importance since it does not lead to the exponential dependence of the conductivity. One can use the regular close packing fraction $\zeta_2=0.74$. In the scaling theory, the mobility is a power function of the distance from the threshold:[15]

$$\mu(\varepsilon_p)=\frac{e\tau}{m}\left(\frac{\varepsilon_p-\Delta_1}{\Delta_2-\Delta_1}\right)^{\nu}, \quad \varepsilon_p<\Delta_2, \tag{9}$$

where $\nu\approx 0.9$ is the critical exponent, $\tau=l/v_T$ is the relaxation time, $l\approx R_s$ is the minimal free path length, and $v_T=\sqrt{8T/\pi m}$ is the mean thermal velocity. Below and at the transition point the Boltzmann statistics of the quasiatomic excitations has been shown to apply even at very low temperatures.[11] By averaging with the density of states, which corresponds to free motion in a major part of the quasimolecular-state volume, we obtain the electrical conductivity

$$\sigma=\frac{ze^2n_m\tau\vartheta}{m}, \tag{10}$$

where the factor $\vartheta<1$ describes the partial localization of the electrons. This factor is expressed by the following combination of the incomplete gamma-functions $\Gamma(m,x)$:

$$F_k=\frac{2}{\sqrt{\pi}}\left[\Gamma\left(\frac{5}{2},\frac{\Delta_k}{T}\right)-\frac{\Delta_k}{T}\,\Gamma\left(\frac{3}{2},\frac{\Delta_k}{T}\right)\right].$$

Below the insulator–metal transition point the localization factor is

$$\vartheta=\frac{T(F_1-F_2)}{\Delta_2-\Delta_1}\approx\frac{2}{\sqrt{\pi}}\,\frac{\sqrt{\Delta_1 T}}{(\Delta_2-\Delta_1)}\exp\left(-\frac{\Delta_1}{T}\right),$$

$$\Delta_2>\Delta_1\gg T, \tag{11}$$

which is exponentially small. On the metallic side

$$\vartheta=\frac{3T/2-\Delta_1-TF_2}{\Delta_2-\Delta_1}\approx\frac{3T/2-\Delta_1}{\Delta_2-\Delta_1},$$

$$\Delta_1\leqslant 0, \quad \Delta_2\gg T. \tag{12}$$

Taking into account the electron degeneracy above the insulator-metal transition point, the thermal energy $3T/2$ in the last formula must be substituted by a renormalized Fermi-energy

$$\varepsilon_F'=\hbar^2k_F^2\vartheta^2/2m, \tag{13}$$

where the prime indicates the difference from the electron gas, and $k_F=(6\pi^2n_m)^{1/3}$ is the Fermi wave-vector. In practice, the localization factor in this case goes to unity. Moreover, this factor drops out in the expression of conductivity. Indeed, for the degenerated electrons the relaxation time is

$$\tau=l/v_F', \quad v_F'=v_F\vartheta, \tag{14}$$

where v_F' is a renormalized Fermi velocity, and v_F is the Fermi velocity of the electron gas. Substituting Eq. (14) into Eq. (10), we reduce the localization factor. Therefore, we obtain the minimal metallic conductivity

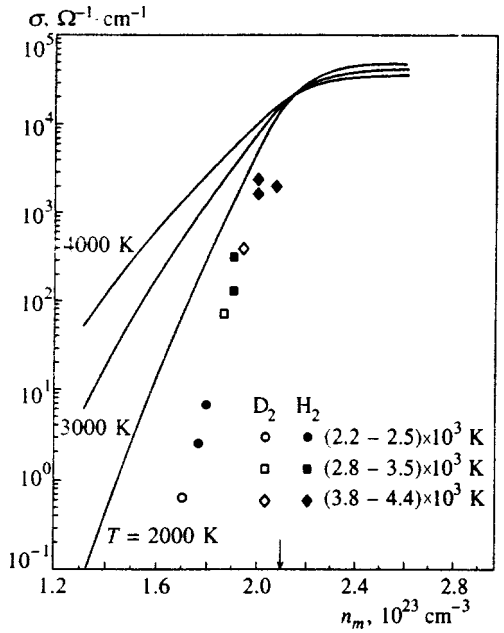

FIG. 2. Electric conductivity of dense fluid hydrogen in the insulator-metal transition range. Experimental dots from Ref. 4 and solid lines by the microscopic percolation theory are shown. The estimated insulator-metal transition point is shown by the γ arrow. The γ parameter is 0.1.

$$\sigma=\frac{e^2n_el}{mv_F}. \tag{15}$$

In fact, for the degenerated electrons the electric conductivity can be described in gas-kinetic terms. We use the Ioffe–Regel criterion for the minimal free path length

$$l\sim\hbar/\Delta p, \tag{16}$$

where Δp is the quantum uncertainty of the electron momentum. At high temperatures the momentum uncertainty can be substituted by the thermal momentum mv_T. The minimal free path is then the thermal wave length. From Eq. (10) we then obtain

$$\sigma\approx\frac{e^2n_e}{k_Fmv_T}\approx\frac{e^2n_eR_s}{mv_T}. \tag{17}$$

Except for a numerical coefficient of the order of unity, this expression is the same as it would be in the case of Boltzmann statistics if the free path length were equal to R_s. With such an accuracy, Eqs. (10)–(12) for the percolation conductivity can also be directly extrapolated for the degenerate electrons on the metallic side of the transition.

In the case of strong degeneracy the momentum uncertainty can be expressed only by the Fermi momentum

$$\Delta p=\gamma mv_F,$$

where γ is a coefficient, and the free path length is

$$l\sim 1/\gamma k_F.$$

Matching it to the extrapolated Boltzmann formula on the metallic side of the transition (at the point at which $v_T=\gamma v_F$), we obtain

$$l=R_s/\gamma, \tag{18}$$

i.e., the parameter γ is the inverse free path length of the intermolecular spacing. For a rough estimate of the parameter γ, applying the $v_T = \gamma v$ equality to the transition point, we obtain $\gamma \sim 0.1$. Of course, the uncertainty of this parameter does not influence the Boltzmann conductivity below the insulator-metal transition point.

The percolation conductivity of dense molecular hydrogen in the insulator-metal transition region is shown in Fig. 2 to qualitatively agree with the experiment. Thus, an idea of a microscopic percolation can be actually instructive for understanding of the insulator-metal transition in the strongly compressed dielectric liquid. Obviously, the theory based on a virtual molecular structure is limited to the vicinity of the transition point, in particular, because hydrogen molecules dissociate deep into the metallic state. While the metallic state is reasonably well understood,[16] the transition is still a subject of study.

In conclusion, we apply a microscopic percolation theory for the description of the insulator-metal transition in dense fluid hydrogen with a virtual molecular structure and show that molecular bonding can still be consistent with the metallic properties.

This work was partially supported by the Russian Fund for Fundamental Research (Grant No. 97-02-16222). A part of this work has been done during my visit to the Institut für Niedertemperatur-Plasmaphysik, Greifswald (Deutschland) with a partial financial support of Deutscher Akademischer Austauschdients. I am grateful to H. Schneidenbach for help in programming and to H. Hess for help and hospitality.

[1] N. F. Mott, *Metal-Insulator Transition*, London, Taylor, & Francis (1991).
[2] D. Saumon, W. B. Hubbard, G. Chabrier, and H. M. van Horn, Astrophys. J. **391**, 827 (1992).
[3] H. K. Mao and R. J. Hemley, Rev. Mod. Phys. **66**, 671 (1994).
[4] S. T. Weir, A. C. Mitchell, and W. J. Nellis, Phys. Rev. Lett. **76**, 1860 (1996).
[5] J. M. Ziman, *Models of Disorder*, Cambridge University Press (1979).
[6] B. I. Shklovsky and A. L. Efros, *Electronic Properties of Doped Semiconductors*. Springer-Verlag, New York (1989) [Russ. original, Nauka, Moscow, 1979)].
[7] M. Ross, Phys. Rev. B **54**, R9589 (1996).
[8] A. A. Likalter, Usp. Fiz. Nauk. **162**, 119 (1992) [Sov. Phys. Usp. **357**, 591 (1992)].
[9] A. A. Likalter, Zh. Eksp. Teor. Fiz. **94**, 157 (1988) [Sov. Phys. JETP **67**, 2478 (1988)].
[10] A. A. Likalter, Zh. Eksp. Teor. Fiz. **111**, 938 (1997) [JETP **84**, 516 (1997)].
[11] A. A. Likalter, Zh. Eksp. Teor. Fiz. **107**, 1996 (1995) [JETP **80**, 1105 (1995)].
[12] H. Reinholz, R. Redmer, and S. Nagel, Phys. Rev. E **52**, 5368 (1995).
[13] A. L. Bug, S. A. Safran, and G. S. Grest, Phys. Rev. Lett. **55**, 1896 (1985).
[14] P. P. Edwards and M. J. Sienko, Phys. Rev. B **17**, 2575 (1978).
[15] P. A. Lee and T. V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985).
[16] A. A. Abrikosov, Astron. Zh. **31**, 112 (1954).

Published in English in the original Russian journal. Reproduced here with stylistic changes by the Translation Editor.

Ferromagnetism of manganese compounds

R. O. Zaĭtsev

Russian Research Center ''Kurchatov Institute,'' 123182 Moscow, Russia

Yu. V. Mikhaĭlova

State Research Center ''NII Teplopribor,'' 129085 Moscow, Russia

(Submitted 13 October 1997)
Zh. Éksp. Teor. Fiz. **113**, 1101–1111 (March 1998)

The possibility of ferromagnetic instability in a system with hopping between manganese cations and oxygen anions has been investigated on the basis of the concept of strong electron–electron interaction in one unit cell. A phase diagram for ferromagnetic ordering as a function of the filling factors of the $2p^6$-shell of oxygen (n_p) and $3t_g^6$-shell of manganese (n_t) has been constructed. © *1998 American Institute of Physics.* [S1063-7761(98)02703-6]

1. INTRODUCTION

Strong electron–electron interaction in the same unit cell is the cause of a considerable increase in the spin part of the magnetic susceptibility.[1] It turned out, however, that the electron–electron scattering amplitude at the Fermi surface decreases with an increase in electron density,[2] which makes ferromagnetic ordering impossible throughout the region which can be studied in the gas approximation. These statements are confirmed by cluster calculations and in fact apply only to s-states of electrons, which have only two internal spin degress of freedom.

As follows from experimental data, the $4s$-shell of pure metallic manganese contains 0.6 electron per cell. Therefore the rest of the $3d$-electrons completely fill the six-electron t_{2g}-shell, while the e_g^4-shell, with four electron, contains only 0.4 of one e_g-electron. It can be shown that the system cannot be ferromagnetic at such a low concentration.

In compounds like MnO, a bivalent Mn^{2+} cation has a partially filled t_{2g}-electron shell with five electrons that corresponds to a one-hole state if $2p^6$-shell of oxygen is completely filled. Since the partially filled t_{2g}-shell contains an odd number of electrons and the manganese oxide is in a dielectric state at $T>T_N$, it is natural to suppose[3] that the Hubbard energy for t_{2g}-electrons is higher than the hopping energy between the manganese cation and oxygen anion.

In the antiferromagnetic state of $La^{3+}Mn^{3+}O_3^{2-}$ the manganese cation is in a trivalent state, so that the manganese has two holes in the six-electron t_{2g}-shell.

In the $Ca^{2+}Mn^{4+}O_3^{2-}$ antiferromagnetic compound the manganese cation is in the tetravalent state, in agreement with the suggestion about the presence of three holes in the t_{2g}^6-shell.[4]

If we denote by n_t the number of holes in the t_{2g}^6-shell of the manganese and by n_p the number of holes in the $2p^6$-shell of O^{2-} anions, the electroneutrality condition for $La_{(1-x)}^{3+}Me_x^{2+}MnO_3$ is expressed as follows:

$$n_t+3n_p=x+2, \tag{1}$$

where $x<1$ is the concentration of bivalent substitutional metal ions (Me) of the second group (Ca, Sr, Ba).

As the number of oxygen holes n_p changes from zero to unity, the number of holes in cations, n_t, changes from three to zero.

2. PROBLEM STATEMENT AND GENERAL RELATIONSHIPS

In compounds with the perovskite structure, the most probable hopping is between the transitional metal cation, which is at the center of a simple cubic unit cell, and the nearest oxygen anions at face centers.

Let us assume that the hopping does not change the spin projection so that the tunneling Hamiltonian is only determined by one hopping integral $t_{(r)}$:

$$\hat{H}=\sum_{\mathbf{r},\mathbf{r}',\sigma,\eta,\nu} t(\mathbf{r}-\mathbf{r}')[\hat{d}^+_{\mathbf{r},\sigma,\eta}\hat{p}_{\mathbf{r}',\sigma,\nu}+\text{H.c.}]+\sum_{\mathbf{r},\sigma,\eta}[\epsilon_d - \sigma H]\hat{d}^+_{\mathbf{r},\sigma,\eta}\hat{d}_{\mathbf{r},\sigma,\eta}+\sum_{\mathbf{r},\sigma,\nu}[\epsilon_p-\sigma H]\hat{p}^+_{\mathbf{r},\sigma,\nu}\hat{p}_{\mathbf{r},\sigma,\nu}. \tag{2}$$

Here μ is the chemical potential, $\sigma=\pm$ is the spin index, and H is the applied magnetic field. The crystal index λ takes three values ($\eta=xy,yz,zx$) when the t_{2g}-shell is filled and three values ($\nu=x,y,z$) when the p-shell is filled.

In the (Ln, Me)MnO_3 compounds, triply degenerate states of manganese cations overlap with three states of oxygen anions, each of which is triply degenerate. Thus the matrix of hopping integrals has in the general case the form of a 3×9 rectangle. If only hopping between nearest neighbors is taken into account, each t_{2g}-state overlaps with only the two p-states of one of the two oxygen anions on the same straight line. Introducing for convenience the notation $\phi_k=t[1-\exp(ip_k)]$, we present the matrix of tunneling transitions in Table I.

Thus, in the simplest approximation used in this work, each state of manganese cations overlaps independently with the p-states of the four nearest oxygen cations. For this rea-

TABLE I.

Atomic states	$p_x^{(1)}$	$p_y^{(1)}$	$p_z^{(1)}$	$p_x^{(2)}$	$p_y^{(2)}$	$p_z^{(2)}$	$p_x^{(3)}$	$p_y^{(3)}$	$p_z^{(3)}$
Mn—xy	0	ϕ_x	0	ϕ_y	0	0	0	0	0
Mn—yz	0	0	0	0	0	ϕ_y	0	ϕ_z	0
Mn—zx	0	0	ϕ_x	0	0	0	ϕ_z	0	0

son, the equations that determine the one-particle Green's function split into three independent equations.

As noted above, the Hubbard energies for both manganese and oxygen ions are the largest energy parameters; we therefore assume that they are infinite to simplify our calculations.

Given this condition, solutions must be sought independently for each integer change in n_t. At the same time, we assume n_p is no greater than unity.

3. INTERVAL OF n_t BETWEEN ZERO AND UNITY

In the limiting case of infinite Hubbard energy, it is convenient to express creation and annihilation operators in terms of $\hat{X}$, i.e., operators of transition between empty (0) and one-particle ($|\sigma,\lambda\rangle$) states:

$$\hat{d}_{\mathbf{r}\sigma,\lambda}=\hat{X}_{\mathbf{r}}^{(0|\sigma,\lambda)}, \quad \hat{d}_{\mathbf{r}\sigma,\lambda}^{+}=\hat{X}_{\mathbf{r}}^{(\sigma,\lambda|0)}, \tag{3'}$$

$$\hat{p}_{\mathbf{r}\sigma,\nu}=\hat{Y}_{\mathbf{r}}^{(0|\sigma,\nu)}; \quad \hat{p}_{\mathbf{r}\sigma,\nu}^{+}=\hat{Y}_{\mathbf{r}}^{(\sigma,\nu|0)}. \tag{3''}$$

In calculating mean occupation numbers $n_{p,d}^{(\sigma)}$, we use the simplest approximation[2] with the excitation energy determined by an averaged self-energy part, which is expressed in terms of so-called terminal factors $f_{p,d}^{(\sigma)}$:

$$n_d^{(\sigma)}=3f_d^{(\sigma)}\sum_{\mathbf{p}\lambda=\pm} a_{\mathbf{p}}^{(-\lambda)}n_F(\xi_{\mathbf{p}}^{(\sigma,\lambda)}), \tag{4'}$$

$$n_p^{(\sigma)}=f_p^{(\sigma)}\left[2n_F(\epsilon_p)+\sum_{\mathbf{p}\lambda=\pm} a_{\mathbf{p}}^{(\lambda)}n_F(\xi_{\mathbf{p}}^{(\sigma,\lambda)})\right], \tag{4''}$$

where $n_F(\epsilon)$ is the Fermi distribution function,

$$\xi_{\mathbf{p}}^{(\sigma,\pm)}=\pm\sqrt{\left(\frac{r}{2}\right)^2+f_p^{(\sigma)}f_d^{(\sigma)}t_{\mathbf{p}}^2}-\sigma H-\mu,$$

$$a_{\mathbf{p}}^{(\pm)}=\frac{1}{2}\left[1\pm\frac{r}{\xi_{\mathbf{p}}^{(\sigma,+)}-\xi_{\mathbf{p}}^{(\sigma,-)}}\right],$$

$$t_{\mathbf{p}}^2=t^2[2-\cos(p_x)-\cos(p_y)], \quad \mu=-(\epsilon_p+\epsilon_d)/2,$$

$$r=\epsilon_p-\epsilon_d.$$

Taking into account the cubic symmetry, the terminal factors can be expressed in terms of mean occupation numbers of one-particle states, $n^{(\sigma)}=\sum_s^3 n_s^{(\sigma)}$:

$$f_\lambda^{(\sigma)}=1-\sum_s^3 n_s^{(\bar{\sigma})}-\sum_{s\neq\lambda}^3 n_s^{(\sigma)}=1-n^{(\bar{\sigma})}-\frac{2n^{(\sigma)}}{3}. \tag{5}$$

Hereinafter $\tilde{\sigma}=-\sigma=\pm$.

From these equations, we easily derive the spin magnetic moment $M_{p,d}=n^{(\sigma)}-n^{(\tilde{\sigma})}$ for both p- and d-states. The possibility of ferromagnetic instability is determined by the existence of a singularity in the magnetic susceptibility:

$$\det\begin{pmatrix}1-(K_0+L_d), & -f_dL_d/f_p\\ -f_pL_p/f_d, & 3-P_0-L_p\end{pmatrix}=0. \tag{6}$$

Here we have introduced the following functions:

$$K_0=\sum_{\mathbf{p},\kappa=\pm} a_{\mathbf{p}}^{(-\kappa)}n_F(\xi_{\mathbf{p}}^{(\lambda)})=\frac{n_t}{6-5n_t},$$

$$\xi_{\mathbf{p}}^{(\pm)}=\pm\sqrt{(r/2)^2+f_tf_pt_{\mathbf{p}}^2}-\mu, \tag{7}$$

$$P_0=2n_F(\epsilon_p)+\sum_{\mathbf{p},\kappa=\pm} a_{\mathbf{p}}^{(\kappa)}n_F(\xi_{\mathbf{p}}^{(\lambda)})=3\frac{n_p}{6-5n_p},$$

$$L_{p,d}=\sum_{\mathbf{p},\lambda=\pm}\left\{\frac{\delta}{\delta t_{\mathbf{p}}^2}[t_{\mathbf{p}}^2n_F(\xi_{\mathbf{p}}^{(\lambda)})a_{\mathbf{p}}^{(\pm\lambda)}]-n_F(\xi_{\mathbf{p}}^{(\lambda)})a_{\mathbf{p}}^{(\pm\lambda)}\right\}. \tag{8}$$

All the functions depend on one parameter that is invariant under transformations of the quadratic symmetry group:

$$t_{\mathbf{p}}^2=2t^2(2-\cos p_x-\cos p_y). \tag{9}$$

It is convenient to introduce the density of states for an ''empty'' quadratic lattice:

$$\rho_0(\epsilon)=\sum_{p_x,p_y}\delta(\epsilon-\cos p_x-\cos p_y),$$

after which all coefficients can be expressed in terms of integrals of the density of states and predetermined functions of the dimensionless variable ϵ.

4. INTERVAL OF n_t BETWEEN UNIT AND TWO

Consider the situation when t_{2g}-states resonate between one- and two-hole states, whereas the number of $2p$-holes in the $2p^6$-shell of oxygen is still less than one ($n_p<1$).

One-particle states $\hat{a}_\sigma^+|0\rangle$, $\hat{b}_\sigma^+|0\rangle$, and $\hat{c}_\sigma^+|0\rangle$ have spin 1/2. The lowest two-particle states $^3A_2$ have spin $S=1$:

$$\hat{a}_\sigma^+\hat{b}_\sigma^+|0\rangle \quad (S_z=\sigma), \quad \frac{\hat{a}_\uparrow^+\hat{b}_\downarrow^++\hat{a}_\downarrow^+\hat{b}_\uparrow^+}{\sqrt{2}}|0\rangle \quad (S_z=0). \tag{10}$$

The remaining six two-particle states are obtained by cyclic permutation $a\to b\to c\to a$.

States ^1E and $^1A_1$ with higher energies are not taken into account.

At finite magnitudes of applied magnetic field the variations of the terminal factors depend on variations of both one-particle ($n_{\rm I}$) and two-particle ($n_{\rm II}$) occupation numbers. Taking into account the system symmetry under permutations of a-, b-, and c-states, we calculate variations of terminal factors:

$$f_1^{(\sigma)}=n_{\rm II}^{(\sigma)}+n_{\rm I}^{(\sigma)},\quad \delta f_1^{(\sigma)}=\delta n_{\rm II}^{(\sigma)}+\delta n_{\rm I}^{(\sigma)},$$
$$f_2^{(\sigma)}=n_{\rm II}^{(0)}+n_{\rm I}^{(\bar\sigma)},\quad \delta f_2^{(\sigma)}=\delta n_{\rm I}^{(\bar\sigma)}=-\delta n_{\rm I}^{(\sigma)}. \tag{11}$$

Thus, unlike the previous case of one-particle states, we need two independent equations for calculating one-particle and two-particle variations.

In order to derive these equations, let us multiply the dominant part of the annihilation operator expansion

$$\hat a_{\mathbf{r}\sigma}=\hat Z_{\mathbf{r}}=g_1\hat X_{\mathbf{r}}^{(0,\sigma,0|\sigma,\sigma,0)}+g_2\hat X_{\mathbf{r}}^{(0,\bar\sigma,0|\sigma,\bar\sigma,0)}+g_3\hat X_{\mathbf{r}}^{(0,0,\sigma|\sigma,0,\sigma)}+g_4\hat X_{\mathbf{r}}^{(0,0,\bar\sigma|\sigma,0,\bar\sigma)}, \tag{12$'$}$$

where $g_1=-g_3=1$, $g_2=-g_4=1/\sqrt{2}$, by an arbitrary linear combination of conjugated X-operators:

$$\hat Y_{\mathbf{r}}=\beta_1\hat X_{\mathbf{r}}^{(\sigma,\sigma,0|0,\sigma,0)}+\beta_2\hat X_{\mathbf{r}}^{(\sigma,\bar\sigma,0|0,\bar\sigma,0)}+\beta_3\hat X_{\mathbf{r}}^{(\sigma,0,\sigma|0,0,\sigma)}+\beta_4\hat X_{\mathbf{r}}^{(\sigma,0,\bar\sigma|0,0,\bar\sigma)}. \tag{12$''$}$$

By averaging separate T-products over states with a given temperature and chemical potential in the one-loop approximation, we obtain an equation relating the two-particle occupation numbers $n_{\rm II}$, Fourier components of the virtual one-particle Green's function $\hat G_\omega(\mathbf{p})$, and terminal factors $f_k(\sigma)$, where $k=1,2,3,4$, and also $f_p(\sigma)$.

The matrix inverse of the one-particle Green's function, in turn, can be expressed in terms of terminal factors f_k, energy difference $r=\epsilon_p-\epsilon_d$, and the sum $\epsilon_p+\epsilon_d=-2(\mu+\sigma H)$:

$$\hat G_\omega^{-1}(\mathbf{p})=\begin{pmatrix}(i\omega-\epsilon_d 1)\delta_{k,n}, & -g_k f_k\nu_x, & -g_k f_k\nu_y\\ -f_p\nu_x^* g_n, & i\omega-\epsilon_p, & 0\\ -f_p\nu_y^* g_n, & 0, & i\omega-\epsilon_p\end{pmatrix}. \tag{13}$$

Let us calculate T-products in the zero-loop approximation ''Hubbard I'':

$$\begin{aligned}-\langle\hat T(\hat Z_{\mathbf{r}}(\tau)\hat Y_{\mathbf{r}}(\tau+0))\rangle&=(g_1\beta_1+g_3\beta_3)\langle X^{(\sigma,\sigma|\sigma,\sigma)}\rangle+(g_2\beta_2+g_4\beta_4)\langle X^{(A_0|A_0)}\rangle\\&=(g_1\beta_1+g_3\beta_3)n_{\rm II}^{(\sigma)}+(g_2\beta_2+g_4\beta_4)n_{\rm II}^{(0)}\\&=T\sum_{\omega,\mathbf{p};1\leqslant k,s\leqslant 4}g_k G_\omega^{(k,s)}(\mathbf{p})\beta_s f_s e^{i\omega\delta}.\end{aligned} \tag{14}$$

Here δ is a small positive number and f_s is the terminal factor.

By varying Eq. (14) with respect to magnetic field, we obtain two equations. Under the condition

$$\sum_{1\leqslant k\leqslant 4}g_k\beta_k=0$$

we obtain an equation independent of the magnitude of applied magnetic field:

$$\delta n_{\rm II}^{(\sigma)}(1-K_0)-2K_0\delta n_{\rm I}^{(\sigma)}=\delta f_1^{(\sigma)}(1-K_0)+\delta f_2^{(\sigma)}(1+K_0)=0. \tag{15}$$

Here the electron (or hole) density lies in the interval $1<n_t<2$ and the coefficients can be calculated at zero external magnetic field:

$$K_0=\sum_{\mathbf{p},\lambda=\pm}a_{\mathbf{p}}^{(-\lambda)}n_F(\xi_{\mathbf{p}}^{(\lambda)})=2\frac{n_t-1}{4-n_t},$$
$$\xi_{\mathbf{p}}^{(\pm)}=\pm\sqrt{(r/2)^2+g^2f_tf_pt_{\mathbf{p}}^2}-\mu, \tag{16}$$
$$f_t=\frac{4-n_t}{18},\quad f_p=\frac{6-5n_p}{6},\quad g^2=3. \tag{17}$$

Under the condition $g_k=\beta_k$, we have an equation for the susceptibility:

$$\delta n_{\rm II}^{(\sigma)}=\delta f_2^{(\sigma)}+\delta f_1^{(\sigma)}=[K_0+L_d]\sum_{k=1,2}g_k^2\delta f_k^{(\sigma)}+g^2\frac{f_d}{2f_p}L_d\delta f_p-g^2fD_0\sigma\delta H, \tag{18}$$

where the coefficient L_d is given by the following general formula:

$$\begin{aligned}L_d&=\sum_{\mathbf{p},\lambda=\pm}\left\{\frac{\delta}{\delta t_{\mathbf{p}}^2}[t_{\mathbf{p}}^2n_F(\xi_{\mathbf{p}}^{(\lambda)})a_{\mathbf{p}}^{(-\lambda)}]-n_F(\xi_{\mathbf{p}}^{(\lambda)})a_{\mathbf{p}}^{(-\lambda)}\right\}\\&=\sum_{\mathbf{p},\lambda=\pm}\left\{\frac{\delta}{\delta t_{\mathbf{p}}^2}[t_{\mathbf{p}}^2n_F(\xi_{\mathbf{p}}^{(\lambda)})a_{\mathbf{p}}^{(-\lambda)}]\right\}-K_0,\end{aligned} \tag{19}$$

$$D_0=\sum_{\mathbf{p},\kappa=\pm}a_{\mathbf{p}}^{(-\kappa)}n_F'(\xi_{\mathbf{p}}^{(\kappa)}),$$
$$a_{\mathbf{p}}^{(\pm)}=\frac{1}{2}\left[1\pm\frac{r}{\sqrt{r^2+4g^2f_tf_p\tau_{\mathbf{p}}^2}}\right]. \tag{20}$$

The equation for the variations of oxygen p-states is derived from the equation for occupation numbers n_p similar to Eq. (4$''$):

$$n_p^{(\sigma)}=f_p^{(\sigma)}\left[2n_F(\epsilon_p)+\sum_{\mathbf{p}\kappa=\pm}a_{\mathbf{p}}^{(-\kappa)}n_F(\xi_{\mathbf{p}}^{(\sigma,\kappa)})\right]. \tag{21}$$

The relation between the variations of the occupation numbers and the variations of the terminal factors $f_p^{(\sigma)}$ is derived from the general equation (5):

$$\delta n_p^{(\sigma)}=3\delta f_p^{(\sigma)}=2L_p\frac{f_p}{g^2f_d}\sum_{k=1,2}g_k^2\delta f_k^{(\sigma)}+[P_0+L_p]\delta f_p-f_pR_0\sigma\delta H. \tag{22}$$

Its coefficients are calculated at zero magnetic field:

$$P_0=\frac{n_p}{f_p}=3\frac{n_p}{6-5n_p},\quad R_0=2n_F'(\epsilon_p)+\sum_{\mathbf{p},\lambda=\pm}a_{\mathbf{p}}^{(\lambda)}n_F'(\xi_{\mathbf{p}}^{\lambda}), \tag{23}$$

$$L_p=\sum_{\mathbf{p},\lambda=\pm}\left\{\frac{\delta}{\delta t_{\mathbf{p}}^2}[t_{\mathbf{p}}^2 n_F(\xi_{\mathbf{p}}^{(\lambda)})a_{\mathbf{p}}^{(\lambda)}]-n_F(\xi_{\mathbf{p}}^{(\lambda)})a_{\mathbf{p}}^{(\lambda)}\right\}$$

$$=\sum_{\mathbf{p},\lambda}\left\{\frac{\delta}{\delta t_{\mathbf{p}}^2}[t_{\mathbf{p}}^2 n_F(\xi_{\mathbf{p}}^{(+)})a_{\mathbf{p}}^{(\lambda)}]\right\}-P_0+2n_F(\epsilon_p). \quad (24)$$

Thus the system of three equations, (15), (21), and (22), determines changes in the three terminal factors, in terms of which one can express the variations of all occupation numbers.

The solvability condition for the respective homogeneous equation system is the condition of ferromagnetic instability. As a result, the condition of a phase transition to the ferromagnetic state is expressed in a simple form:

$$(3-P_0)[3K_0(1-K_0)-L_d(1+3K_0)]$$
$$-3L_pK_0(1-K_0)=0. \quad (25)$$

In this equation

$$K_0=2\frac{n_t-1}{4-n_t}. \quad (26)$$

Functions P_0 and $L_{p,d}$ are the same as in Eqs. (7) and (8), but with genealogical coefficients and terminal factors defined in a different manner:

$$g^2=3,\quad g_1^2=1,\quad g_2^2=1/2,\quad f_d=(4-n_t)/18,$$
$$\xi_{\mathbf{p}}^{(\pm)}=\pm\sqrt{r^2/4+g^2f_pf_dt_{\mathbf{p}}^2}-\mu. \quad (27)$$

The equation of state for n_p has the same form as Eq. (2). The equation for n_t is (see Eq. (3))

$$n_t=1+9f_d\sum_{\mathbf{p},\lambda=\pm}a_{\mathbf{p}}^{(-\lambda)}n_F(\xi_{\mathbf{p}}^{(\lambda)}). \quad (28)$$

The normal coordinates and excitation spectrum are determined by the same equations as in the previous section, but with a factor f_t and genealogical coefficients g_k defined in a different manner:

$$g^2=3,\quad f_t=\frac{4-n_t}{18},\quad f_p=\frac{6-5n_p}{6}. \quad (29)$$

This equation relates energy parameters ϵ_p and ϵ_d. By eliminating these variables using equations of state (12) and (17) in zero field, we obtain an equation for the magnetic phase diagram in variables n_t and n_p:

$$\sum_{\mathbf{p},\lambda=\pm}a_{\mathbf{p}}^{(-\lambda)}n_F(\xi_{\mathbf{p}}^{(\lambda)})=2\frac{n_t-1}{4-n_t}=K_0,$$
$$n_p=2f_p\left[2n_F(\epsilon_p)+\sum_{\mathbf{p},\lambda=\pm}a_{\mathbf{p}}^{(\lambda)}n_F(\xi_{\mathbf{p}}^{(\lambda)})\right]. \quad (30)$$

5. INTERVAL OF CONCENTRATIONS n_t BETWEEN 2 AND 3

Experimental data indicate that in (La, Ca)MnO_3 compounds, the manganese cations have a positive charge ranging between $+3$ and $+4$. Hence we conclude that in this material the d electron states of manganese resonate between $3d^3$- and $3d^4$-states. According to the electroneutrality condition (1), these states correspond to the interval $2<n_t<3$ in the hole representation, for which we have resonance between two- and three-particle t_{2g}-states, whereas the occupation number n_p is still less than 1.

The lowest three-particle state has $S=3/2$ and is fourfold degenerate in the spin projection:

$$\hat{a}_\sigma^+\hat{b}_\sigma^+\hat{c}_\sigma^+|0\rangle,\quad S_z=3\sigma/2;$$
$$\frac{1}{\sqrt{3}}(\hat{a}_{\bar{\sigma}}^+\hat{b}_\sigma^+\hat{c}_\sigma^+|0\rangle+\hat{a}_\sigma^+\hat{b}_{\bar{\sigma}}^+\hat{c}_\sigma^+|0\rangle+\hat{a}_\sigma^+\hat{b}_\sigma^+\hat{c}_{\bar{\sigma}}^+|0\rangle),$$
$$S_z=\frac{\sigma}{2}. \quad (31)$$

The three lowest triplet states with spin 1 are constructed from three different products of creation operator pairs (see definition (10)).

The expansion in terms of X-operators describing transitions between two- and three-particle states with the lowest energies is determined by three genealogical coefficients:

$$\hat{a}_{\mathbf{r}\sigma}=\hat{X}_{\mathbf{r}}^{(0,\sigma,\sigma|3\sigma/2)}+\sqrt{\frac{2}{3}}\,\hat{X}_{\mathbf{r}}^{(A(yz,xz)|\sigma/2)}+\frac{1}{\sqrt{3}}\,\hat{X}_{\mathbf{r}}^{(0,\bar{\sigma},\sigma|\bar{\sigma}/2)}. \quad (32)$$

The expansion of two other annihilation operators is derived from Eq. (32) by using the operation of cyclic permutation.

At zero field, all mean occupation numbers and terminal factors can be expressed in terms of n_t, i.e., the mean number of holes per cell. Taking degeneracy into account, we have

$$9n_{\mathrm{II}}+4n_{\mathrm{III}}=1,\quad 18n_{\mathrm{II}}+12n_{\mathrm{III}}=n_t,$$
$$f_t=\frac{5n_t-6}{36}. \quad (33)$$

To obtain the equation of state for $H=0$, we express occupation numbers of three-particle states in terms of the one-particle Green's function at coincidence points.

After summation with respect to the spin index, we obtain the equation of state

$$n_t=2+4fK_0,\quad K_0=\sum_{\mathbf{p},\lambda=\pm}a_{\mathbf{p}}^{(-\lambda)}n_F(\xi_{\mathbf{p}}^{(\lambda)}),$$
$$\xi_{\mathbf{p}}^{(\pm)}=\pm\sqrt{(r/2)^2+2f_pf_tt_{\mathbf{p}}^2}-\mu. \quad (34)$$

The equations for variations of three-particle occupation numbers,

$$\delta n_{\mathrm{III}}^{(3\sigma/2)},\quad \delta n_{\mathrm{III}}^{(\sigma/2)}=-\delta n_{\mathrm{III}}^{(-\sigma/2)},$$

can be derived from the general equation for the mean value of T-products of annihilation operator (32) for a linear combination of three conjugate operators with arbitrary coefficients γ_s:

$$g_5\beta_5n_{\mathrm{III}}^{(3\sigma/2)}+g_6\beta_6n_{\mathrm{III}}^{(\sigma/2)}+g_7\beta_7n_{\mathrm{III}}^{(-\sigma/2)}$$
$$=T\sum_{5\leqslant k,n\leqslant 7}\sum_{\omega\mathbf{p}}g_kG_\omega^{k,n}(\mathbf{p})\gamma_nf_n. \quad (35)$$

By calculating matrix elements of the one-particle Green's function using the general definition (13) for the inverse Green's function, we obtain the right-hand side of Eq. (35):

$$g_5\beta_5\delta n_{\rm III}^{(3/2)}+g_6\beta_6\delta n_{\rm III}^{(1/2)}+g_7\beta_7\delta n_{\rm III}^{(-1/2)}=K_0[g_5\beta_5\delta f_5+g_6\beta_6\delta f_6+g_7\beta_7\delta f_7]+(\mathbf{b}\cdot\boldsymbol{\beta})f\delta G. \tag{36}$$

If vector $\boldsymbol{\beta}$ is assumed to be orthogonal to $\mathbf{b}$, i.e., $\sum_{5\leqslant k\leqslant 7} g_k\beta_k=0$, we can obtain two equations that do not contain the magnetic field variation in explicit form.

The first equation is obtained under the conditions $g_7\beta_7=g_5\beta_5$ and $b_6\beta_6=-2b_5\beta_5$:

$$\delta n_{\rm III}^{(3/2)}=3\,\delta n_{\rm III}^{(1/2)}. \tag{37}$$

Assuming that $\beta_6=0$ and $g_7\beta_7=-g_5\beta_5$, we obtain the second equation:

$$(1-K_0)(\delta n_{\rm III}^{(3/2)}+\delta n_{\rm III}^{(1/2)})-2K_0\delta n_{\rm II}=0. \tag{38}$$

Variations of terminal factors can be expressed in terms of variations of occupation numbers:

$$\delta f_5=\delta n_{\rm III}^{(3/2)}+\delta n_{\rm II},\quad \delta f_6=\delta n_{\rm III}^{(1/2)},\quad \delta f_7=\delta n_{\rm III}^{(-1/2)}-\delta n_{\rm II}. \tag{39}$$

Using the additional condition $\delta n_{\rm III}^{(-1/2)}=-\delta n_{\rm III}^{(-1/2)}$, we obtain the reciprocal relations

$$\delta n_{\rm III}^{(3/2)}=\delta f_5+\delta f_6+\delta f_7,\quad \delta n_{\rm III}^{(1/2)}=\delta f_6,$$
$$\delta n_{\rm II}=-\delta f_6-\delta f_7. \tag{40}$$

The variation of the virtual Green's function δG contains three types of terms:

$$g_5^2\delta n_{\rm III}^{(3\sigma/2)}+g_6^2\delta n_{\rm III}^{(\sigma/2)}+g_7^2\delta n_{\rm III}^{(-\sigma/2)}=[K_0+L_d]\times\sum_{k=5,6,7} g_k^2\delta f_k^{(\sigma)}+g^2\frac{f_d}{f_p}L_d\delta f_p-g^2f_dD_0\sigma\delta H, \tag{41}$$

where the coefficients $L_{p,d}$ are determined by general equation (8), and

$$D_0=\sum_{\mathbf{p},\lambda=\pm} a_{\mathbf{p}}^{(-\lambda)}n_F'(\xi_{\mathbf{p}}^{(\lambda)}),$$
$$a_{\mathbf{p}}^{(\pm)}=\frac{1}{2}\left[1\pm\frac{r}{\sqrt{r^2+4g^2f_tf_p\tau_{\mathbf{p}}^2}}\right]. \tag{42}$$

Thus, we have obtained the same equations as Eqs. (14)–(16) but with different definitions of the quantities:

$$g^2=3,\quad f_t=\frac{5n_t-6}{36},\quad g_5^2=1,\quad g_6^2=\frac{2}{3},$$
$$g_7^2=\frac{1}{3},\quad K_0=9\,\frac{n_t-2}{5n_t-6}. \tag{43}$$

The equation defining the variation of n_p is similar to Eqs. (18) and (19):

$$\delta n_p^{(\sigma)}=3\,\delta f_p^{(\sigma)}=L_p\frac{f_p}{g^2f_d}\sum_{k=5,6,7} g_k^2\delta f_k^{(\sigma)}+[P_0+L_p]\delta f_p-f_pR_0\sigma\delta H. \tag{44}$$

All the coefficients can be calculated at zero magnetic field:

$$P_0=\frac{n_p}{f_p}=3\,\frac{n_p}{6-5n_p},$$
$$R_0=2n_F'(\epsilon_p)+\sum_{\mathbf{p},\kappa=\pm} a_{\mathbf{p}}^{(\kappa)}n_F'(\xi_{\mathbf{p}}^{(\kappa)}). \tag{45}$$

Thus, we have a system of four equations that yields the criterion for onset of ferromagnetic ordering:

$$\det\begin{pmatrix} 1-g_5^2(K_0+L_d) & 4/3-g_6^2(K_0+L_d) & 1-g_7^2(K_0+L_d) & -g^2f_dL_d/f_p \\ -1 & +2 & -1 & 0 \\ 1-K_0 & 2 & 1+K_0 & 0 \\ -g_5^2f_pL_p/g^2f_d & -g_6^2f_dL_p/g^2f_d & -g_7^2f_p/g^2f_d & 3-P_0-L_p \end{pmatrix}=0. \tag{46}$$

Calculating the determinant yields

$$(3-P_0)[3K_0(1-K_0)-L_d(2+3K_0)]-3L_pK_0(1-K_0)=0, \tag{47}$$

where

$$K_0=9\,\frac{n_t-2}{5n_t-6},\quad f_t=\frac{5n_t-6}{36},\quad g^2=2,$$
$$n_t=2+4f_t\sum_{\mathbf{p},\lambda=\pm} a_{\mathbf{p}}^{(-\lambda)}n_F(\xi_{\mathbf{p}}^{(\lambda)}). \tag{48}$$

6. DISCUSSION AND CONCLUSIONS

An important difference between the situations at $n_t<1$ (Eq. (6)) and at $3>n_t>1$ (Eqs. (25) and (47)) is that at $n_t<1$ condition (6) of ferromagnetic ordering for small numbers of particles, when $K_0\to 0$, can be satisfied only when L_d or L_p are of order unity.

In the other two cases ($2>n_t>1$ and $3>n_t>2$) ferromagnetic ordering will necessarily occur even at small numbers of quasiparticles, i.e., when $K_0\ll 1$. The ferromagnetic instability vanishes at $K_0\approx L_t$, as follows from Eqs. (25) and (47) (Fig. 1).

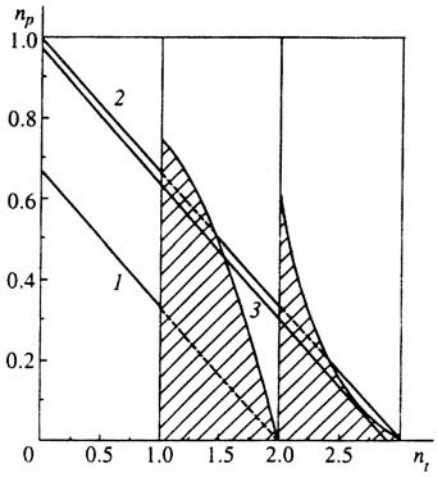

FIG. 1. Magnetic phase diagram for the $p-d$-electron subsystem at $T=0$ calculated in the flat-band approximation. The regions of ferromagnetic ordering are hatched. The lines of electroneutrality correspond to the following materials: 1) pure $LaMnO_3$; 2) pure $CaMnO_3$; 3) $La_{0.1}Ca_{0.9}MnO_3$, i.e., at critical concentration $x=x^*\approx 0.9$.

It is convenient to express coefficients $L_{p,d}$, which are proportional to the scattering amplitudes of p- and d-excitations with opposite spins, in terms of K_0 and P_0:

$$L_d=w_d-K_0, \quad L_p=w_p-P_0+2n_F(\epsilon_p). \tag{49}$$

The parameters $w_{p,d}$ can be represented as a sum of total derivatives

$$w_{p,d}=\sum_{\mathbf{p},\lambda=\pm}\frac{\delta}{\delta t_{\mathbf{p}}^2}[t_{\mathbf{p}}^2 n_F(\xi_{\mathbf{p}}^{(\lambda)})a_{\mathbf{p}}^{(\pm\lambda)}]. \tag{50}$$

They can easily be calculated for the flat-band model, when the density of states for $t_{\mathbf{p}}^2$ is constant. For example, when the lowest subband is filled, $\min\{\xi_{\mathbf{p}}^{(-)}\}<0$, but $\xi_{\mathbf{p}}^{(+)}>0$ for all momenta $\mathbf{p}$, we have in the limit $T=0$

$$w_{p,d}=\frac{1}{2}\left\{1\mp\frac{r}{\sqrt{r^2+16g^2t^2f_pf_t}}\right\}. \tag{51}$$

Hence, at all finite values of the dimensionless parameter $|r/t|$, $w_{p,d}$ has a finite positive value no greater then unity.

We therefore conclude that the two scattering amplitudes L_p and L_d decrease from the positive number $w_{p,d}$ to negative $[-1+w_{p,d}]$ as the chemical potential increases.

In addition to this conclusion, note also that at small K_0 the left-hand sides of Eqs. (25) and (47) are negative. Therefore ferromagnetic instability is inevitable when the concentration n_t deviates slightly from roughly 1 or 2 to the high side, since in these conditions the system resonates between magnetic (Hund) states (Fig. 1).

With increasing concentrations n_p and n_d, the signs of scattering amplitudes L_p and L_d change, which corresponds to a transition across the boundary of ferromagnetic instability.

This phenomenon was observed when trivalent lanthanum was replaced with a bivalent calcium cation in $La_{1-x}Ca_xMnO_3$. As the charge of the MnO_3 complex increased, i.e., the concentration x rose, the ferromagnetic ordering temperature passed through a broad peak shifted towards the region of concentrations about $x\approx 1/4$. At higher concentrations, $x>1/4$, the Curie temperature decreased rapidly, and vanished by $x>1/2$.

These effects are in qualitative agreement with the phase diagram in Fig. 1.

When crossing the electroneutrality lines 1 to higher concentrations of bivalent cations, we approach the boundary of ferromagnetic instability. In the long run, at $x\geqslant x^*$ the electroneutrality line 1 turns out to lie partially outside the ferromagnetic instability region, and twice crosses the ferromagnetic ordering boundary.

As can be seen in Fig. 1, the critical concentration $x^*\approx 0.9$ is higher than the measured value $x^*\approx 0.5$.

[1] E. C. Stoner, Proc. R. Soc. London **165**, 372 (1938).
[2] J. Hubbard and K. P. Jain, J. Phys. C **2**, 1650 (1968).
[3] J. C. Slater, Phys. Rev. **82**, 538 (1951).
[4] J. Goodenough, Phys. Rev. **100**, 564 (1955).
[5] R. O. Zaitsev, Zh. Eksp. Teor. Fiz. **70**, 1100 (1976) [Sov. Phys. JETP **43**, 574 (1976)].

Translation provided by the Russian Editorial office

Dynamic model of a double chain with hydrogen bonds

A. A. Samoletov and A. É. Filippov*)

Donetsk Physicotechnical Institute, Ukrainian National Academy of Sciences, 340114 Donetsk, Ukraine
(Submitted 19 March 1997)
Zh. Éksp. Teor. Fiz. **113**, 1112–1121 (March 1998)

The evolution of a system of two interacting molecular chains joined by the collective field of an adiabatically rapidly moving subsystem of light (hydrogen) atoms is considered. The motion of all three subsystems is simulated directly on the basis of the equation of the dynamics of an open system. An ordered structure, as well as its collective fluctuations, are obtained as a result of self-organization of the system in the presence of noise and relaxation. The possibilities of developing such a description for the direct simulation of DNA molecules are discussed.
 [S1063-7761(98)02803-0]

1. INTRODUCTION

According to current ideas, the static or averaged structure of biological molecules by itself does not completely determine their biological function. In particular, the process of transcription is a typical example, in which the dynamics of the deoxyribonucleic acid (DNA) molecule have a significant bearing on its function, since the double helix must have local openings in order to expose the coding bases to the chemical reaction.

This is, of course, a complex phenomenon involving RNA-polymerase, which may utilize an energy-localization mechanism as it moves along the DNA molecule and the local openings in the double helix.[1] The DNA denaturation process, which is simpler in design and has been studied extensively by experimental means, exhibits some similarity to transcription and is a convenient and important process for simulating the dynamics of such complex molecules as DNA.[2,3]

The dynamic simulation of the evolution of DNA molecules has aroused increasingly intense interest in recent years. In this context two principal areas of research can be identified:

a) descriptions based on relatively simple phenomenological models, which can be reproduced numerically and investigated analytically;[3–11]

b) direct numerical simulations which take into account the motion of all the atoms comprising the DNA molecule.[12–19]

In case (a) the so-called simple models are employed (the term ''simple model'' was introduced in the review in Ref. 20 and has become the standard term in the literature), in which some details of the dynamics in the usual mechanistic sense of the word are lost. A considerable part of the structural information regarding the DNA molecule is also lost. However, this is the price for being able to simulate the surrounding physiological conditions characteristic of the molecule instead of treating it as an isolated structure.

This is an extremely important point. First, it corresponds to the fundamentally biological essence of the problem (although it is investigated by physical methods). Small changes, for example, in the chemical composition of the medium, lead to radical reorganization of the dynamics of the biomolecule and its structure. Second, from the physical standpoint, the molecule becomes an open system and requires a description by the methods of nonequilibrium statistical physics. This is already a matter of kinetics, rather than dynamics per se.

In case (b) the macromolecule is investigated by molecular dynamics with consideration of its complete atomic composition. Since one base pair in a DNA molecule contains about 60 atoms, a simulation can be performed only for small fragments of it and very short intervals of real physical time (less than 10^{-9} s), which are significantly shorter than the characteristic times of such processes as replication, transcription,[1] and denaturation.[2] However, some details of the molecular motion manage to be manifested even within such short time intervals. They provide information on the intramolecular dynamic cooperative movements of the atoms, which cannot be obtained by any other method and which underlies slow and large-scale biological processes.[12,18]

One recent promising direction in the development of the numerical simulation of DNA is the retention of as few degrees of freedom in the molecular-dynamics equations as possible to achieve a compromise between maximum simplicity and preservation of the specific properties of the system.[21] It should be supplemented by a compromise for taking into account the features of the surrounding physiological medium.[22,23]

The specific properties of DNA include the presence of hydrogen bonds which join the two molecular chains to one another. Many dynamic properties of the system (primarily the fluctuational properties associated with the transport of nonlinear excitations along it) are determined by hydrogen bonds. Qualitatively, the existence of such bonds is a consequence of the rapid motion of the light hydrogen subsystem superposed on the adiabatically slow movements of the heavy molecular complexes making up the chains.

In a first approximation the heavy molecules move in a certain collective (time-averaged) potential created by the hydrogen atoms. This conception permitted the construction

of a self-consistent phonon theory of such motion.[18,19] The approximation is valid when the system is in a state close to one of its stationary states.

However, the most important problem in the theory is to describe the significantly nonlinear dynamics of DNA during the cleavage of bonds and the local opening of fragments of the double helix. In these extreme configurations the motion of the hydrogen subsystem cannot be considered ergodic, and the time-averaged potentials poorly reflect the instantaneous distribution of the forces in the dissociated fragment.

The use of time-averaged potentials is based on the analytical tradition, within which finding a rough description involved transforming the problem, which originated as a many-particle dynamic problem of the motion of an ensemble of interacting subsystems, into a particular continuum description on the basis of collective fields (densities). The possibilities currently offered by computer technology make it possible to numerically simulate fairly complicated systems with a long-range interaction and different characteristic rates for the processes occurring in the subsystems comprising them.

The continuum densities and the effective potentials associated with them are reconstructed in such an *a posteriori* approach. This makes it possible, in principle, to verify the results of the continuum theories and to reveal the structural features of the system that were totally lost in the rough description. The use of simple bare potentials even provides a definite gain both in the speed of the calculations and in the volume of the system accessible to simulation.

A DNA molecule contains two types of hydrogen bonds involving two or three hydrogen atoms. The most convenient objects for the theory, however, are artificially created polymers, for example, poly-TATA, or the corresponding fragments of natural molecules.[24,25] In a numerical simulation the complication associated with an arbitrary sequence of bases is insignificant, but it leads to a large number of dynamic scenarios. To demonstrate the most general features of the model, in this paper we present the results of a simulation of the poly-TATA system.

The purpose of the work is to obtain a working model of a double chain stabilized by the rapid motion of the light subsystem which qualitatively reproduces the most general features of a real DNA molecule and to carry out a preliminary investigation of it. The formulation of the model is given in Sec. 2.

The main result of this work is a demonstration that direct dynamic stabilization of the double helix by the rapid motion of the light subsystem not only is possible, in principle, but also provides for a speed in the model that is sufficient for simulating large molecules (with 10^3 or more base pairs) over long time intervals. The model thus devised ensures the global dynamic stability of the system under investigation under one set of conditions and opens up the possibility for a model description of thermal and chemical denaturation under another set of conditions. The results are discussed in Sec. 3.

2. FORMULATION OF THE MODEL

Let us consider a model system consisting of a sequence of identical groups of particles, each of which contains particles of two kinds (two particles of each kind) having different masses: $m_2 = 10^2 m_1$. We write the equations of motion in the form of a system of stochastic differential equations:

$$d^2\mathbf{r}_i/dt^2 = \Big(-\gamma d\mathbf{r}_i/dt - \sum_j \partial \mathbf{U}(\mathbf{r}_i - \mathbf{r}_j)/\partial \mathbf{r}_i + \xi(\mathbf{r}_i, t) \Big)/m_i, \tag{1}$$

where the final temperature of the system is modeled by the simultaneous addition of an additive noise, i.e., a generalized delta-correlated Gaussian random field[26,27] (white noise)

$$\langle \xi(\mathbf{r},t) \rangle = 0, \quad \langle \xi(\mathbf{r},t)\xi(\mathbf{r}',t) \rangle = 2D\delta(\mathbf{r}-\mathbf{r}')\delta(t-t') \tag{2}$$

and the relaxation terms $\gamma\, d\mathbf{r}_i/dt$. The ratio of the noise intensity D to the relaxation rate γ can be varied by assigning the effective temperature $T \sim D$ of the system.

This is the simplest form of the equations of the dynamics of an open system. As we have already noted, the dynamics are important for understanding the biological function of the molecule, for which, in turn, the surrounding physiological medium (its composition and temperature) is important. A limiting description in terms of the system of stochastic differential equations (1) and (2) is obtained, if there is a clear-cut hierarchy of characteristic time scales for the intramolecular processes and the interaction with the medium.

The real values of γ, U, and ξ within the molecular system are different. In the general case the concrete structure of the noise depends on the external environment of the molecule. In particular, a DNA molecule is surrounded by water molecules (where the decisive structural role also belongs to hydrogen bonds), which transform the external white noise acting on the DNA molecule into noise with finite memory times. To obtain equations of types (1) and (2) that take into account these effects, the system must be expanded by including the corresponding molecules in it.

Strictly speaking, the potentials U are not potentials in the mechanical sense, but are nonequilibrium thermodynamic potentials. Some aspects of this question were discussed in Refs. 22 and 23 in connection with the kinetics of DNA. It was shown that an approach based on stochastic equations like (1) and (2) is applicable for describing the kinetics of a DNA molecule under not excessively restrictive conditions. The potentials U then incorporate not only the direct interaction between the atoms, but also the influence of the surrounding physiological medium, the partial screening, etc. These nonequilibrium potentials were obtained in an explicit form within the model in Ref. 3.

The solution of the system of stochastic equations (1) and (2) has its own special features, which have a bearing both on the purely mathematical structure of the problem[26–29] and, in part, on the investigation of the model

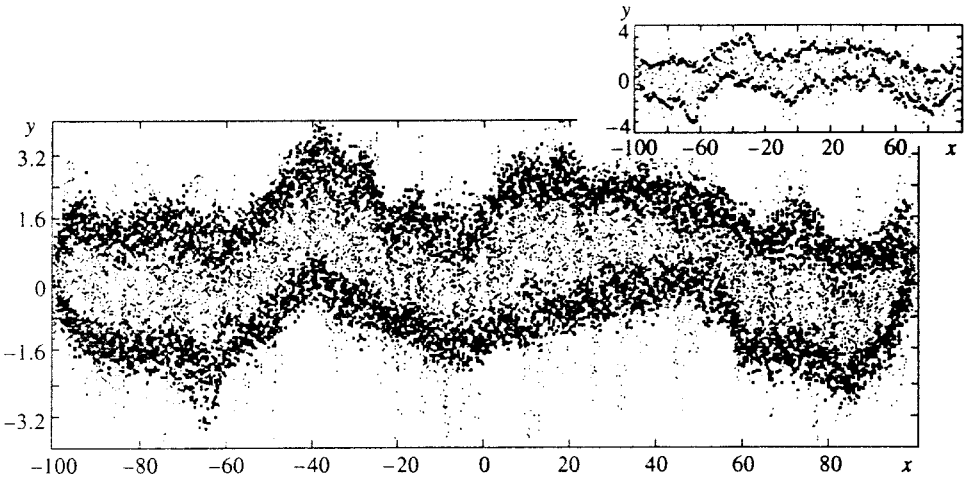

FIG. 1. Set of successive positions of the molecules comprising a fragment of 200 elementary units of a double chain (the ordinary and large points correspond to the light and heavy molecules) accumulated over the course of the time interval $t_{\min}<t<t_{\max}$ described in the text. The inset shows the current positions of the molecules of the same fragment of the structure at the time $t=t_{\max}$.

in numerical experiments. We, therefore, dwell briefly on a qualitative description of the variation of the expected solutions with time.

Introducing the Hamiltonian

$$\mathscr{H}=\sum_{j<i} U(\mathbf{r}_i-\mathbf{r}_j)+\sum_i p_i^2/2m_i,$$

we write system (1) in the form

$$d\mathbf{r}_i/dt=\partial\mathscr{H}/\partial\mathbf{p}_i, \quad d\mathbf{p}_i/dt=-\gamma\mathbf{p}_i/m_i-\partial\mathscr{H}/\partial\mathbf{r}_i+\xi. \tag{3}$$

The corresponding Fokker–Planck equation has the form

$$\partial\rho/\partial t+\{\mathscr{H},\rho\}=\sum_i \gamma[\partial(\mathbf{p}_i\rho)/\partial\mathbf{p}_i]/m_i + D\sum_i [\partial^2\rho/\partial p_i^2]. \tag{4}$$

Equation (4) has a stationary solution of the Gibbs type

$$\rho_0=\exp[-\beta\mathscr{H}], \tag{5}$$

if the Einstein relation holds, i.e., if

$$D=\gamma kT, \quad \beta=1/kT. \tag{6}$$

The distribution of ρ_0 can clearly be factored into velocity and coordinate distributions.

The evolution to this distribution proceeds in such a manner that at first a Maxwellian velocity distribution is established,[20,23] and then slower Smoluchowski diffusion occurs in the configuration space of the system. This process is described either by the Kramers equation for a probability density distribution[20] or by the system of stochastic equations

$$\gamma d\mathbf{r}_i/dt=-\partial\mathscr{H}/\partial\mathbf{r}_i+\xi. \tag{7}$$

The latter description is closer to the essence of the present work. The Boltzmann distribution corresponds to averaging of the noise ξ in all events, while in a numerical experiment (as well as in a physical experiment) we investigate a single event.

Over long times noise leads to a stationary sequence in the form of a Markov chain, whose states are stationary states of the deterministic part of system (7),[28] permitting observation of the corresponding transitions.

In the case under consideration such transitions take place between two intuitively expected states of hydrogen (see below), as well as between nonlinear excitations in the chains of heavy bases that are difficult to predict *a priori*. As the temperature (the noise intensity) increases, this picture is destroyed,[29] and the double chain dissociates. Information on a particular system is contained in the structure of the potentials $\mathbf{U}(\mathbf{r}_i-\mathbf{r}_j)$ and can vary.

In the context of the present work stability is important. When the noise is not excessively intense, the set of particles obeying the equations of the model spontaneously form the necessary configurations without alteration of any of the other degrees of freedom (the vibrational and rotational degrees of freedom, the possibility of denaturation, etc.). The latter was ensured by the following choice of the potentials $\mathbf{U}(\mathbf{r}_i-\mathbf{r}_j)$:

1) the Coulomb potentials regularized at $|\mathbf{r}_i-\mathbf{r}_j|\to 0$ were taken to describe the attraction of the hydrogen atoms to the two heavy molecules within one segment [$\mathbf{U}_1(\mathbf{r}_0-\mathbf{r}_1)$ and $\mathbf{U}_1(\mathbf{r}_0-\mathbf{r}_2)$, respectively] and the repulsion of the heavy molecules from one another $\mathbf{U}_2(\mathbf{r}_1-\mathbf{r}_2)$;

2) the interaction of the neighboring molecules in each of the two chains $\mathbf{U}_2(\mathbf{r}_{1,2}-\mathbf{r}'_{1,2})$ was chosen in the form of the isotropic potential

$$\mathbf{U}_2(\mathbf{r}_{1,2}-\mathbf{r}'_{1,2})=\mathbf{U}_2(|\mathbf{r}_{1,2}-\mathbf{r}'_{1,2}|) =-|\mathbf{r}_{1,2}-\mathbf{r}'_{1,2}|^2(a^2-|\mathbf{r}_{1,2}-\mathbf{r}'_{1,2}|^2/2),$$

which has a minimum at a certain distance $|\mathbf{r}_{1,2}-\mathbf{r}'_{1,2}|=a\neq 0$ between them (which essentially assigns the characteristic spatial scale of the problem $a\equiv 1$ along the chain);

3) the radius of the skeleton σ at which the attraction of the hydrogen atoms to the molecules in the chains gives way to repulsion, which assigns the second (in the transverse direction to the chains) spatial scale of the problem, was modeled by replacing the purely Coulomb potential by $\mathbf{U}_1(\mathbf{r}_0-\mathbf{r}_1)=\mathbf{U}_1(|\mathbf{r}_0-\mathbf{r}_1|)=(1-\sigma/|\mathbf{r}_0-\mathbf{r}_1|)/|\mathbf{r}_0-\mathbf{r}_1|$.

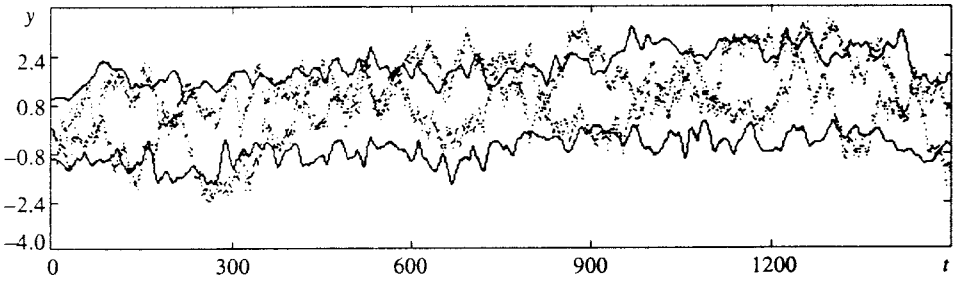

FIG. 2. Typical variations of the coordinates $y_j(t)$ for motion of the system that is steady with time. The motion of four molecules making up an arbitrarily selected internal segment in the structure. The points correspond to the movements of the light atoms.

3. DISCUSSION OF RESULTS

Some results of the investigation of the evolution of the model system are summarized in Figs. 1–5. The relative simplicity of the model permits direct observation of the concerted movements and fluctuations of the double chain of heavy molecules stabilized by the rapid motion of the light atoms on the computer screen.

Figure 1 gives a qualitative representation of the character of this motion. It shows a set of successive positions of the molecules comprising a fragment of 200 base pairs. The ordinary and large points correspond to the light and heavy molecules. This image was accumulated over the course of the time interval $t_{\min}<t<t_{\max}$, which was selected so that the initial time $t_{\min}$ would be significantly longer than the time needed to achieve a stationary state (in the sense discussed above) after ''activation'' of the system at the time $t=0$.

More specifically, Fig. 1 presents 20 events in the distribution $y_j(t)$ for $j=0$, 1, 2, which were obtained with an interval of 50 elementary steps of $\Delta t=0.02$ between two successive images for $t_{\min}=5\times10^4\cdot\Delta t$. The inset presents the current position of the molecule in the same fragment of the structure at the time $t=t_{\max}=t_{\min}+10^3\Delta t=t_{\min}+20$. The difference between the patterns of motion of the fast and slow subsystems is directly visible.

Figure 2 shows the motion of the system in significantly longer time intervals (of the order of $7.5\times10^4\Delta t$). Here typical variations of the coordinates $y_j(t)$ for motion of the system that is steady with time are shown for the case of four molecules (two light and two heavy) comprising an arbitrarily selected internal complementary pair in the structure.

The points in Fig. 2 correspond to the motion of the internal light atoms. It is easy to see that both the quasiperiodic fluctuations of the slow subsystem and the intermittency in the motion of each light atom gravitate first to one and then the other chain.

All these motions can be characterized quantitatively using the correlation functions

$$G(j,i;t)=\langle y_j(0)y_i(t)\rangle-\langle y_j(t)y_i(t)\rangle$$
$$=\int dt' y_j(t'+t)y_i(t')/T-\langle y_j(t)\rangle\langle y_i(t)\rangle, \quad (8)$$

calculated for both subsystems $j,i=0,1;2,3$.

The functions

$$G(0,0;t)=\langle y_0(0)y_0(t)\rangle-\langle y_0(t)\rangle^2,$$
$$G(2,2;t)=\langle y_2(0)y_2(t)\rangle-\langle y_2(t)\rangle^2.$$

are presented in Figs. 3a and 4a, respectively, as examples. The function $G(2,2;t)$ has a clearly expressed structure, which is associated with the existence of a characteristic time scale corresponding to fluctuations of the slow subsystem.

The existence of this scale is confirmed by the presence of characteristic extrema of the Fourier transform of this function

$$G(2,2;\omega)=\int dt\ \exp(i\omega t)G(2,2;t)/2\pi i, \quad (9)$$

which is shown in Fig. 4b.

The values of $\pm\omega$ in the figure are normalized to 100%, and the interval between the extrema in dimensionless units is ~0.03. Unlike $G(2,2;t)$, $G(0,0;t)$ is scarcely structured and, as is seen from Fig. 3b, where this function is represented in semilogarithmic coordinates, decays exponentially as t increases.

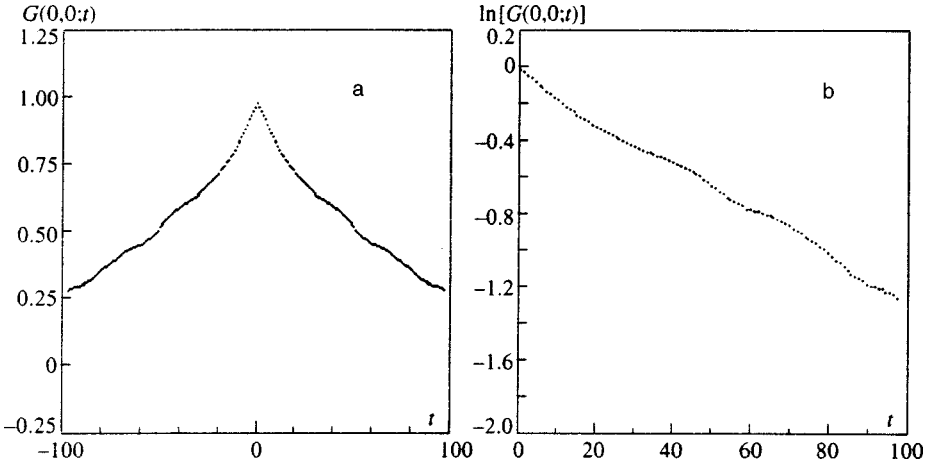

FIG. 3. The correlation function $G(0,0;t)=\langle y_0(0)y_0(t)\rangle-\langle y_0(t)\rangle^2=\int dt'\ y_0(t'+t)\times y_0(t')/T-\langle y_0(t)\rangle^2$ in ordinary (a) and semilogarithmic (b) coordinates.

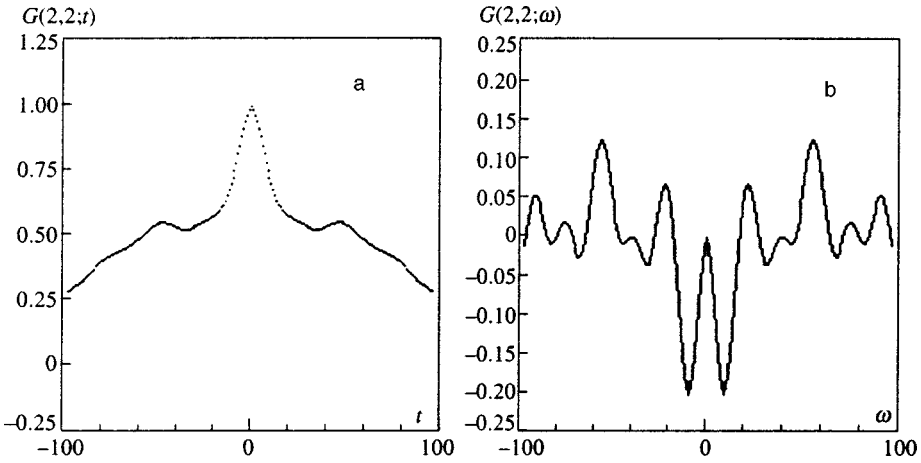

FIG. 4. The correlation function $G(2,2;t)=\langle y_2(0)y_2(t)\rangle-\langle y_2(t)\rangle^2$ plotted directly as a function of time (a) and using the Fourier transform $G(2,2;\omega)=\int dt\,\exp(i\omega t)\times G(2,2;t)/2\pi i$ (b).

The intermittency in the motion of each light atom corresponds to the intuitively expected mechanism of bonding between the chains, which is specified by the dynamics of the fast subsystem. The particles, i.e., the mediators of the interaction, alternately spend a relatively long time near each of the chains and thereby create the desired charge-mediated bonding on the average. This, in turn, means that a state with a displaced position for a light atom should be energetically preferable. The effective potential for it is determined by the mutual adjustment of all the forces of the problem, including the balance between the noise and relaxation.

If the sum of the interaction potentials $\mathbf{U}_k(\mathbf{r}_i-\mathbf{r}_j;t)$ at the arbitrary moments in time t

$$\mathbf{U}_{\text{eff}}(\mathbf{r}_i-\mathbf{r}_j;t)=\mathbf{U}_1(\mathbf{r}_0-\mathbf{r}_i;t)+\mathbf{U}_1(\mathbf{r}_0-\mathbf{r}_2;t)+\mathbf{U}_2(\mathbf{r}_1-\mathbf{r}_2;t), \qquad (10)$$

is formally calculated, we find that $\mathbf{U}_{\text{eff}}(\mathbf{r}_i-\mathbf{r}_j;t)$ also has a very large spread because of the significant random component in the variable $\mathbf{r}_0(t)$.

Figure 5 presents a typical distribution of the interaction energy $\mathbf{U}_{\text{eff}}(\mathbf{r}_i-\mathbf{r}_j;t)$, which was calculated for an arbitrarily selected internal segment of the structure and plotted as a function of the instantaneous coordinate of one of the hydrogen atoms $y_0(t)$.

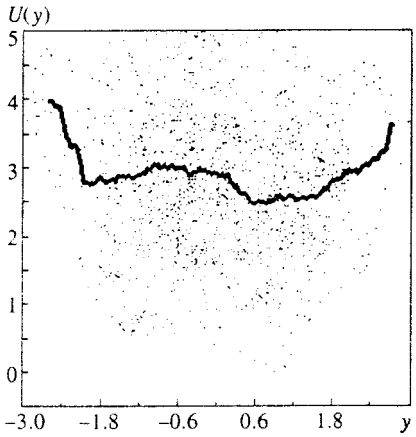

FIG. 5. Typical distribution of the interaction energy $\mathbf{U}_{\text{eff}}(\mathbf{r}_i-\mathbf{r}_j;t)$ for an arbitrarily selected internal segment of the structure as a function of the instantaneous coordinate of one of the hydrogen atoms $y_0(t)$. The heavy curve shows the effective double-well potential obtained as a result of averaging over a time interval $T=7.5\times10^4\cdot\Delta t$.

The significance of the role of the random component of $\mathbf{r}_0(t)$ in shaping $\mathbf{U}_{\text{eff}}(\mathbf{r}_i-\mathbf{r}_j;t)$ is revealed by the spread of the points in this figure.

The effective double-well potential sought is obtained only as a result of averaging over a sufficiently long time interval $T\gg\Delta t$. In particular, Fig. 5 shows a set of 1500 successive events with the result $\mathbf{U}_{\text{eff}}(\mathbf{r}_i-\mathbf{r}_j;t)$ calculated with an interval of 50 elementary steps of $\Delta t=0.02$ between one another. The result of the averaging over $T=7.5\times10^4\Delta t=1.5\times10^3$ is illustrated by the thick solid curve.

The results obtained provide evidence that the working model presented in this paper of a double chain stabilized by the rapid motion of the light subsystem reproduces both the most general features of a real DNA molecule and the averaged characteristics of the system expected on the basis of the conventional adiabatic approximation.

Direct dynamic stabilization of the double chain, which does not resort to such an approximation, provides for a speed in the model that is sufficient for investigating large molecules over long time intervals.

It imparts global dynamic stability to the system under investigation at a relatively small noise intensity (temperature) and opens up a possibility for describing thermal and chemical denaturation when the intensity of the external disturbances is increased.

One of us (A. S.) thanks M. Peyrard for bringing this problem to his attention, as well as for the hospitality provided during his stay in the Ecole Normale Supérieure de Lyon (France).

*)E-mail: filippov@host.dipt.donetsk.ua

[1] *Transcription: Mechanisms and Regulation*, R. C. Conaway and J. W. Conaway (eds.), Raven, New York (1994).
[2] R. M. Wartell and A. S. Benight, Phys. Rep. **126**, 67 (1985).
[3] M. Peyrard and A. R. Bishop, Phys. Rev. Lett. **62**, 2755 (1989).
[4] M. Techera, L. L. Daemen, and E. W. Prohofsky, Phys. Rev. A **40**, 6636 (1989).
[5] M. Techera, L. L. Daemen, and E. W. Prohofsky, Phys. Rev. A **41**, 4543 (1990).

[6] M. Techera, L. L. Daemen, and E. W. Prohofsky, Phys. Rev. A **42**, 1008 (1990).
[7] V. Muto, P. S. Lomdahl, and P. L. Christiansen, Phys. Rev. A **42**, 7452 (1990).
[8] T. Dauxois, M. Peyrard, and A. R. Bishop, Phys. Rev. E **47**, 684 (1993).
[9] P. L. Christiansen and V. Muto, Physica D **68**, 93 (1993).
[10] A. Yu. Grosberg and A. R. Khokhlov, *Statistical Physics of Macromolecules*, American Institute of Physics, New York (1994).
[11] S. N. Volkov, Phys. Lett. A **224**, 93 (1996).
[12] J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge (1987).
[13] U. C. Sing, S. J. Weiner, and P. Kollman, Proc. Natl. Acad. Sci. USA **82**, 755 (1985).
[14] W. F. Gunsteren *et al.*, Ann. (N.Y.) Acad. Sci. **482**, 287 (1986).
[15] S. N. Rao and P. Kollman, Biopolymers **29**, 517 (1990).
[16] S. Falsafi and N. O. Reich, Biopolymers **33**, 459 (1993).
[17] C. Prevost *et al.*, Biopolymers **33**, 335 (1993).
[18] F. Brinki, ''Etude des mouvements internes dans l'ADN par simulations de dynamique moléculaire newtonienne et stochastique,'' Thesis, Universite de Paris VI, Paris (1993).
[19] F. Brinki and D. Genest, J. Biomol. Struct. Dyn. **11**, 43 (1993).
[20] G. Gaeta, C. Reiss, M. Peyrard, and T. Dauxois, Riv. Nuovo Cimento Soc. Ital. Fis. **17**, 1 (1994).
[21] F. Zhang and M. A. Collins, Phys. Rev. E **52**, 4217 (1995).
[22] T. Soboleva, A. Samoletov, and M. Peyrard, Phys. Rev. B **38**, 341 (1997).
[23] A. Samoletov, T. Soboleva, and M. Peyrard, in *3rd European Conference on Mathematics Applied to Biology and Medicine (ECMBM)*, Heidelberg (1966), p. 62.
[24] Y. Kim, K. V. Devi-Prasad, and E. W. Prohofsky, Phys. Rev. B **32**, 5185 (1985).
[25] E. W. Prohofsky, Phys. Rev. A **38**, 1538 (1988).
[26] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag, Heidelberg (1985).
[27] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions, Vol. 4: Applications of Harmonic Analysis*, Academic Press, New York (1964).
[28] A. D. Ventsel' and M. I. Freĭdlin, Usp. Mat. Nauk **25**, 3 (1970).
[29] R. Z. Khas'minskiĭ (Hasminskii), *Stochastic Stability of Differential Equations*, Nauka, Moscow (1969); Noordhoff, Alphenaanden Rijn, Netherlands (1980).

Translated by P. Shelnitz

Hexagonal optical structures in photorefractive crystals with a feedback mirror

P. M. Lushnikov*)

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia
(Submitted 13 October 1997)
Zh. Éksp. Teor. Fiz. **113**, 1122–1146 (March 1998)

A nonlinear theory is presented for the formation of hexagonal optical structures in a photorefractive medium equipped with a feedback mirror. Oppositely directed beams in photorefractive crystals are unstable against the excitation of sideband waves. It is shown here that as this instability evolves to its nonlinear stage, the three-wave interaction between weak sideband beams does not stabilize it, but rather leads to explosive growth of the amplitudes of beams whose transverse wave vectors form angles that are multiples of $\pi/3$. As a result, sideband beams at these angles are found to be correlated. A range of parameters is found in which four-wave interactions saturate the explosive instability, which explains the appearance of stable hexagons in the experiment. Outside this region, nonlinearities of higher order saturate the explosive instability, and the process of hexagon generation must be studied numerically. Matrix elements are obtained for the three- and four-wave interactions as functions of the distance to the feedback mirror, and an equation for the time evolution of the sideband wave amplitudes is derived that describes the hexagon generation. A comparison is made with experimental results for the photorefractive crystals $KNbO_3$ and $BaTiO_3$. © *1998 American Institute of Physics.* [S1063-7761(98)02903-5]

1. INTRODUCTION

Oppositely directed optical beams passing through nonlinear media often exhibit transverse instability against the excitation of waves at small angles to the primary propagation direction and the generation of transverse hexagon-shaped optical structures.[1–5] This instability is caused by positive feedback between the counterpropagating beams, and is absolute in nature. In photorefractive crystals such phenomena have been especially well-studied,[6–9] due to the extreme ease with which the evolution of the transverse instability and formation of regular structures can be observed. Characteristic times for the creation of these structures range from tenths to tens of seconds. Typical nonlinear lengths within these crystals, over which the amplitude of the light beams changes appreciably, are several millimeters, and the intensities of the pump beams required to generate them are achievable using cw lasers.[10]

Until recently, theoretical studies of the transverse instability concentrated mostly on calculating threshold conditions for the generation of transverse optical structures. This was first done for Kerr media,[11] followed by threshold calculations for the photorefractive crystals $KNbO_3$ and $BaTiO_3$ equipped with feedback mirrors,[12] and the crystals $LiNbO_3$ and $LiTaO_3$ illuminated by oppositely directed pump beams but with no feedback mirror.[13] The crystals $KNbO_3$ and $BaTiO_3$ are the best ones to study from an experimental point of view. When these crystals are equipped with feedback mirrors, it is found that stationary hexagonal structures form, with an instability threshold that is in rather good agreement with theoretical predictions.[6,8,12] In Ref. 12, Honda and Banerjee found the threshold by assuming that the instability was aperiodic, i.e., the imaginary part of the instability growth rate equaled zero at the threshold of the instability (if this were not the case, moving optical structures would be seen. Experimentally, such a motion can only be produced by mismatching the directions of the oppositely directed beams slightly). In other words, the frequency detuning between pump and sideband waves is assumed to be zero at the instability threshold.

Above threshold, the instability leads to generation of weak optical beams at small angles $\theta=|\mathbf{k}_\perp|/k_0$ to the pump beams, where k_0 is the wave vector of the pump beams and $|\mathbf{k}_\perp|$ is the transverse component of the excitation beam wave vector. For pumping slightly above threshold, the only beams generated had wave vectors in a narrow layer around $|\mathbf{k}_\perp|\simeq k_{0\perp}$, where the instability growth rate is a maximum (here $k_{0\perp}$ is the value of the transverse wave vector corresponding to maximum gain). Thus, the initial stage of the evolution of the instability involves the creation of annular structures (in the plane perpendicular to the pump beam) with amplitudes that decay exponentially with time. The distribution of intensity along the beam at this stage is arbitrary and is determined by fluctuations in the medium. This theory is linear in the amplitudes of the weak beams, and therefore cannot describe the subsequent evolution of the instability which leads to the formation of regular hexagonal optical structures. These structures arise from nonlinear interactions between the weak optical beams. The task of this paper is to derive a theory of this interaction for photorefractive $KNbO_3$ and $BaTiO_3$ crystals equipped with a feedback mirror. The small parameter used in the theory is the amplitude of the sideband waves normalized by the square root of the intensity of the pump beams. Since the instability is aperiodic, three-wave interactions between optical beams whose transverse wave vectors make angles with each other that are

multiples of $\pi/3$ become important as the instability evolves toward its nonlinear stage. It is shown here that this interaction does not stabilize the instability, but rather leads to explosive growth in the amplitudes of the weak sideband waves, as a result of which hexagonal structures form. This can be understood from the following example. Let us assume that as a result of the evolution of the linear instability three weak beams are excited with wave vectors $\mathbf{q}_j=(k_0,\mathbf{k}_{\perp j})$, $j=1,2,3$, equal real amplitudes A, and transverse wave vectors $\mathbf{k}_{\perp j}$ that make angles of $\pi/3$ with each other. The linear instability theory of Ref. 12 implies that three other beams with the transverse wave vector $-k_{\perp j}$ will also be excited in the system at the same time, and six beams with $-k_0$: $\mathbf{q}=(-k_0,\pm\mathbf{k}_{\perp 1,2,3})$. The amplitudes of these waves all equal A.

It will be show below that the evolution of $A(t)$ is determined by the following equation:

$$\frac{\partial A}{\partial t}=\nu_0 A+UA^2, \tag{1}$$

where $\nu_0=\nu_{k_{0\perp}}$ is the instability growth rate at its maximum point $|k_\perp|=k_{0\perp}$ and U is the matrix element for the three-wave interaction. Setting $A|_{t=0}=A_0$ and integrating this equation gives

$$A=\frac{A_0\nu_0}{(\nu_0+UA_0)e^{-\nu_0 t}-UA_0}, \tag{2}$$

This expression reveals that there exists a time $t_{\rm cr}>0$ at which the amplitude A goes to infinity when the condition $A_0U>0$ holds. This is an example of a so-called explosive instability, for which the solution becomes singular at a finite time. Thus, a fundamental feature of this problem is that three-wave processes lead to correlations between perturbations whose transverse wave vectors $\mathbf{k}_\perp$ make angles with one another that are multiples of $\pi/3$. These correlated perturbations generate hexagons in the plane of transverse wave vectors $\mathbf{k}_\perp$. In this case there is no correlation between different hexagons, and the interaction between them is small. Four-wave and higher-order wave processes stabilize the instability, and also suppress the generation of other hexagons with smaller amplitudes (i.e., those hexagons which begin to grow after the primary one). Thus, hard excitation of hexagons takes place until amplitudes are reached at which the excitation is stabilized by nonlinearities of fourth and higher order. This hard excitation of hexagons is the analog of a first-order phase transition.

A more general equation, which describes the evolution of weak beams A_k with transverse wave vector $\mathbf{k}$, will be derived below (here and below the sign $\perp$ will be omitted), taking into account three- and four-wave interactions:

$$\frac{\partial A_k}{\partial t}=\nu_0 A_k+\frac{1}{2}\,U\sum_{\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}}A_{k_1}A_{k_2} - \frac{1}{3!}\sum_{\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3=\mathbf{k}}T_{-k\,k_1k_2k_3}A_{k_1}A_{k_2}A_{k_3}, \tag{3}$$

where U, $T_{-kk_1k_2k_3}$ are the matrix elements of the three- and four-wave processes evaluated on the surface $|\mathbf{k}|=k_{0\perp}$. Since linear instability theory predicts the growth of perturbations within the thin ring $|\mathbf{k}|\simeq k_{0\perp}$, in summing in Eq. (3) it is sufficient to retain only those transverse wave vectors whose magnitude corresponds to the maximum linear growth rate, i.e., $|\mathbf{k}|=k_{0\perp}$. Thus, Eq. (3) is essentially a Laudau expansion in the amplitude A_k of the growing linear modes (see, e.g., Ref. 14).

Note that the applicability of Eq. (3), i.e., the possibility of limiting the treatment to three- and four-wave processes, assumes that the nonlinearity is small. For this to be true in general, the matrix element U is required to be small independent of how far above threshold the system is. In the theory of phase transitions this corresponds to a first-order phase transition that is close to a second-order phase transition by virtue of the smallness of the order parameter discontinuity. In photorefractive crystals, the matrix element U is not particularly small; however, the matrix element $T_{-kk_1k_2k_3}$ contains a rather large numerical factor which justifies the existence of a range where Eq. (3) is applicable whenever the total contribution of four-wave processes can provide both saturation of the explosive instability and stability of the stationary hexagonal lattices. In what follows, a range of parameters will be found in which such saturation and stability are actually achieved, and the results of analytic four-wave theory will be compared with a numerical experiment that takes into account wave processes of higher order. As a result, it will be shown that the analytic theory describes the process of hexagon generation in a qualitatively correct manner, but that the stationary amplitudes it predicts differ by roughly a factor of 2 from those obtained by the numerical experiment. Thus, higher-order wave processes lead to an important renormalization of the hexagon amplitudes. The numerical calculations also show that when a range of parameters is deliberately chosen for which four-wave theory cannot saturate the explosive instability, the hexagons are stabilized at larger amplitudes, i.e., for stronger nonlinearity.

The plan of this paper is as follows: in Sec. 2 nonlinear equations are derived that describe the evolution of sideband beams in photorefractive media with a feedback mirror, and a general boundary value problem is formulated for solving these equations. In Sec. 3 the linear theory of transverse instability is investigated by linearizing this boundary value problem. As a result, the threshold condition is found for the instability, along with eigenvectors for the direct and Hermitian-conjugate linear boundary value problems. In Sec. 4 expressions are obtained for the matrix element U of the three-wave interaction. The fact that this matrix element will turn out to be nonzero is of fundamental importance. In Sec. 5 the overall amplitude equation (3) is derived, along with an expression for the four-wave interaction matrix element $T_{-kk_1k_2k_3}$. For photorefractive crystals like $KNbO_3$ and $BaTiO_3$, the matrix elements U and $T_{-kk_1k_2k_3}$ turn out to be purely real. In Sec. 6 an equation is found that describes the time evolution of the hexagon intensity, and the hexagon stability is investigated. This analysis makes it possible to determine the range of parameters in which the four-wave interaction is sufficient to stabilize the growth of the hexagons. In Sec. 7 the generation of hexagons is investigated

numerically, taking into account higher-order nonlinearities, and the theoretical results are compared with experimental results for the photorefractive crystals $KNbO_3$ and $BaTiO_3$. In the last section, all of these results are summarized.

2. FUNDAMENTAL EQUATIONS

Assume that a pump light wave $F_0 \exp[i(n_0k_0z-\omega_0t)]$ propagates along the z axis in a photorefractive crystal along with an oppositely directed wave $B_0 \exp[-i(n_0k_0z-\omega_0t)]$ arising from reflection by a feedback mirror. Here $F_0(z)$, $B_0(z)$ are complex wave amplitudes that vary slowly with z, k_0 is the wave vector of the light waves in vacuum, ω_0 is their frequency, and n_0 is the index of refraction of the crystal. For simplicity we will refer to both waves as pump waves. We denote the distance between the back face of the crystal and the feedback mirror by L and the length of the crystal along the z axis by l, and set the coordinate of the front face of the crystal $z=0$ (see also the experiment setup in Refs. 8 and 12). Consider perturbations of the pump beams in the form of weak sideband waves, and write the total amplitudes of the beams in the form

$$\begin{aligned} F\ \exp[i(n_0k_0z-\omega_0t)] &= F_0\ \exp[i(n_0k_0z-\omega_0t)] \\ &\times\left[1+\sum_{\mathbf{k}}\ \exp(i\mathbf{k}\cdot\mathbf{r}_\perp)F_k(z,t)\right], \\ B\ \exp[-i(n_0k_0z-\omega_0t)] &= B_0\ \exp[-i(n_0k_0z-\omega_0t)] \\ &\times\left[1+\sum_{\mathbf{k}}\ \exp(i\mathbf{k}\cdot\mathbf{r}_\perp)B_k(z,t)\right], \end{aligned} \tag{4}$$

where $\mathbf{k}=(k_x,k_y)$ is the transverse wave vector in the xy plane, $\mathbf{r}_\perp$ is the spatial coordinate in this plane, and F_k, B_k are amplitudes of the sideband waves normalized by the amplitudes of the pump waves. Assume that the polarizations of all the waves are the same, the amplitudes of the sideband waves are small, i.e., $|F_k|$, $|B_k|\ll 1$, and that $|\mathbf{k}|\ll n_0k_0$, i.e., the sideband waves propagate at small angles to the pump beams.

The wave beams in the photorefractive medium interact via the following mechanism. Under the action of the light, current carriers are excited and the crystal acquires a photoconductivity. The modulation of the light intensity caused by interference of the light beams leads to modulation of the photoconductivity and the appearance of a space-charge electric field. Modulation of this space-charge field E in turn leads to modulation of the dielectric constant of the crystal, $\epsilon=\epsilon_0+\delta\epsilon$ according to the linear electrooptic effect: $\delta\epsilon=-n_0^4rE$, where n_0 is the index of refraction of the crystal and r is the effective electrooptic coefficient.[10,15] In particular, if the z axis coincides with the crystallographic z axis, then $r=r_{13}$. Note that only the longitudinal z-component $E\equiv E_z$ of the space-charge field is included in the analysis that follows, since the other components give rise to negligibly small contributions. Thus, the interaction between the light beams arises from their diffraction by the refractive index modulations they induce in the crystal.[10,15]

In this paper we will assume that the wave interaction is mediated only by generation of reflecting refractive index gratings, whose wave vectors are close to $\pm 2n_0k_0$. Under the conditions of the experiment in which the hexagons were seen, interaction via transmission gratings is negligibly small.[8,12] Therefore, the space-charge field $E(\mathbf{r}_\perp,z,t)$ can be written in the form

$$\begin{aligned} E(\mathbf{r}_\perp,z,t) &= \exp(2in_0k_0z)E_{2k_0} \\ &\times\left(1+\sum_{\mathbf{k}}\ \exp(i\mathbf{k}\cdot\mathbf{r}_\perp)E_{2k_0,k}\right) \\ &+\exp(-2in_0k_0z)E_{-2k_0} \\ &\times\left(1+\sum_{\mathbf{k}}\ \exp(i\mathbf{k}\cdot\mathbf{r}_\perp)E_{-2k_0,k}\right), \end{aligned} \tag{5}$$

where $E_{\pm 2k_0}(z)$ are the space-charge field amplitudes arising from the pump beams and $E_{\pm 2k_0,k}(z,t)$ are the space-charge field amplitudes of the sideband beams normalized by the pump-beam field amplitudes. These amplitudes are written by the pump and sideband waves, and also by pairs of weak sideband waves. The reality of the quantity $E(\mathbf{r}_\perp,z,t)$ leads to the following relations between these amplitudes:

$$E_{-2k_0}=E_{2k_0}^*, \quad E_{-2k_0,-k}=E_{2k_0,k}^*.$$

The wave amplitudes F and B vary with time on a scale that is the same order as the characteristic relaxation time of the space-charge field E, which in photorefractive crystals can range from tenths of seconds to tens of seconds;[10] therefore, this dependence can be neglected in the wave equation.[10,15] The equations for the amplitudes F and B, which are slowly varying in z, take the following form in light of Eq. (5):

$$\begin{aligned} \left(\frac{d}{dz}-\frac{i}{2k_0n_0}\,\Delta_\perp\right)F &= -\frac{ik_0n_0^3r}{2}\,BE_{2k_0} \\ &\times\left(1+\sum_{\mathbf{k}}\ e^{i\mathbf{k}\cdot\mathbf{r}_\perp}E_{2k_0,k}\right), \\ \left(\frac{d}{dz}+\frac{i}{2k_0n_0}\,\Delta_\perp\right)B &= \frac{ik_0n_0^3r}{2}\,FE_{-2k_0} \\ &\times\left(1+\sum_{\mathbf{k}}\ e^{i\mathbf{k}\cdot\mathbf{r}_\perp}E_{-2k_0,k}\right), \end{aligned} \tag{6}$$

where $\Delta_\perp=\partial^2/\partial x^2+\partial^2/\partial y^2$.

When the amplitudes of the sideband waves can be neglected, the following expressions for the pump beams follow from Eqs. (6) and (4):

$$\begin{aligned} \frac{d}{dz}\,F_0 &= -\frac{ik_0n_0^3r}{2}\,B_0E_{2k_0}, \\ \frac{d}{dz}\,B_0 &= \frac{ik_0n_0^3r}{2}\,F_0E_{-2k_0}. \end{aligned} \tag{7}$$

Equations (6), (7) must be supplemented by a relation between the space-charge field E and the wave amplitudes. As shown above, the photorefractive nonlinearity is extremely slow; therefore, the generation of the space-charge field turns out to be affected only by the total intensity of the

optical beams averaged over the rapid oscillations with frequency ω_0, i.e., the standing-wave optical patterns formed by interference between the oppositely directed beams. Let us write the light intensity I averaged over the rapid temporal oscillations in the form

$$I=I_\Sigma[I_0+\exp(2in_0k_0z)I_{2k_0}+\exp(-2in_0k_0z)I_{-2k_0}], \tag{8}$$

where

$$\begin{aligned} I_0=&\frac{1}{I_\Sigma}\left\{|F_0|^2\left[1+\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_\perp}(F_k+F^*_{-k})\right.\right.\\ &\left.+\sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{r}_\perp}F_{k_1}F^*_{-k_2}\right]+|B_0|^2\left(1+\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_\perp}(B_k\right.\\ &\left.\left.+B^*_{-k})+\sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{r}_\perp}B_{k_1}B^*_{-k_2}\right)\right\},\\ I_{2k_0}=&\frac{F_0B_0^*}{I_\Sigma}\left[1+\sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}_\perp}(F_k+B^*_{-k})\right.\\ &\left.+\sum_{\mathbf{k}_1,\mathbf{k}_2} e^{i(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{r}_\perp}F_{k_1}B^*_{-k_2}\right], \end{aligned} \tag{9}$$

here $I_{-2k_0}=I^*_{2k_0}$, and $I_\Sigma=|F_0|^2+|B_0|^2$ is the sum of the pump beam intensities.

Assume that the photorefractive crystal contains a set of donor and acceptor levels with densities N_D and N_A, and densities of ionized donors and conduction-band electrons N_D^+ and n respectively.[16] The compensating acceptor levels are completely occupied by electrons and do not participate in any transitions, and $N_D>N_A$. Let us neglect thermal transitions of electrons from donor levels into the conduction band. Then the acceptor charge $-eN_A$ entirely compensates the charge due to ionized donors in the dark. Assume that phototransitions take electrons from donor levels to the conduction band with a probability $sI(N_D-N_D^+)$, and that electrons are trapped by ionized donors with a probability $g_0N_D^+n$, where s is the photoionization cross section and g_0 is the recombination coefficient. Then the density of ionized donors is given by the equation

$$\frac{\partial N_D^+}{\partial t}=sI(N_D-N_D^+)-g_0N_D^+n, \tag{10}$$

which must be supplemented by the Poisson equation

$$\operatorname{div} E=4\pi\frac{e}{\epsilon_\parallel}(N_D^+-N_A-n) \tag{11}$$

and the equation of continuity

$$\frac{\partial(N_D^+-n)}{\partial t}+\frac{1}{e}\operatorname{div} j=0, \tag{12}$$

Here $\epsilon_\parallel$ is the static dielectric constant along the z axis. Note that it is sufficient to take into account only the longitudinal dielectric constant, since only small-angle perturbations $|\mathbf{k}|\ll n_0k_0$ are treated in this paper. This in turn implies that it is sufficient to take into account only the longitudinal component $E=E_z$ of the space-charge field. The electric current density j is determined by drift and diffusion of electrons:

$$j=e\mu nE+eD\nabla n, \tag{13}$$

where μ is the carrier mobility, $D=\mu k_BT/e$ is the diffusion coefficient, T is the temperature, and k_B is the Boltzmann constant. In the majority of experiments on excitation of hexagons in photorefractive crystals, conditions are such that N_D, N_D^+, N_A, $|N_D-N_A|\gg n$;[10] furthermore, the characteristic recombination time $1/g_0N_A$ for carriers (electrons) is considerably smaller than the characteristic relaxation time $t_d=\epsilon_\parallel g_0N_A/4\pi e\mu sI_\Sigma(N_D-N_A)$ for the space-charge field. Therefore, the time derivative in Eq. (10) can be neglected, and n can be expressed as a function of the light intensity as follows:

$$n=\frac{sI(N_D-N_A)}{g_0N_A}. \tag{14}$$

From Eqs. (11)–(14) we obtain an equation for the space-charge field:[13,15]

$$I_\Sigma t_d\frac{\partial E}{\partial t}=-\widetilde{IE}-E_{sc}\tilde{I}, \tag{15}$$

where the tilde instructs us to separate out the spatially oscillating parts with wave vectors $\pm 2n_0k_0$. Here $E_{sc}=2in_0k_0D/\mu$ is a characteristic photoinduced electric field that depends only on the properties of the crystal,[15,17] and is caused by diffusion of photoelectrons. If the conductivity of the crystal is predominantly due to holes, we need only change the coefficients that multiply t_d and E_{sc}.[10]

Equations (15), (8), (9) imply the following expression for the amplitude $E_{\pm 2k_0}$ of the reflecting grating created by the space-charge field generated by the pump beams:

$$E_{2k_0}=-E_{sc}\frac{F_0B_0^*}{I_\Sigma},\quad E_{-2k_0}=E^*_{2k_0}. \tag{16}$$

The theory presented above is entirely suitable for treating the crystals $KNbO_3$ and $BaTiO_3$ under investigation in this paper. This is because diffusion of photoelectrons dominates in these crystals, and according to Eq. (16) the spatial modulation of the refractive index is shifted by $\pi/2$ relative to the modulation of the standing optical pattern, which corresponds to the so-called nonlocal photorefractive response.[10,17] Note that inclusion of either an applied external electric field or the photogalvanic effect leaves the form of Eq. (15) unchanged, only changing the constant E_{sc}.[15,17] In particular, the quantity E_{sc} is almost real in the crystals $LiNbO_3$ and $LiTaO_3$; therefore, there is no phase shift, which corresponds to the local response caused by drift of electrons in an external field or by the photogalvanic effect.[13,15] Therefore, in the interest of greater generality it is assumed below that E_{sc} is an arbitrary complex constant.

Let us assume that the feedback mirror has unit reflection coefficient, $R=1$, and that the reflection of the light beam from the crystal face is negligibly small. Then we obtain from Eqs. (7) and (16)

$$F_0(z)=F_0(0)e^{i\gamma z/2},\quad B_0(z)=B_0(0)e^{-i\gamma^* z/2},$$

$$|F_0(z)|^2=|B_0(z)|^2=|F_0(0)|^2e^{-\gamma_i z}, \tag{17}$$

where $\gamma\equiv\gamma_r+i\gamma_i=k_0n_0^3rE_{sc}/2$ is the coupling constant of the photorefractive crystal, which can easily be obtained from experiment by using Eqs. (17). Then the z-dependence of the characteristic relaxation time for the space-charge field has the form

$$t_d(z)=t_d(0)e^{\gamma_i z}. \tag{18}$$

Equations (4), (6), (8), (9), and (15)–(18), can be used to obtain a closed system of equations for the weak-beam amplitudes. For subsequent calculations it is convenient to write this system in matrix form:

$$\mathbf{J}\partial_t\mathbf{\Psi}_k=\mathscr{L}\mathbf{\Psi}_k+\sum_{\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}}\boldsymbol{\eta}(\mathbf{\Psi}_{k_1},\mathbf{\Psi}_{k_2})+\sum_{\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3=\mathbf{k}}\mathbf{\Theta}(\mathbf{\Psi}_{k_1},\mathbf{\Psi}_{k_2},\mathbf{\Psi}_{k_3}), \tag{19}$$

where

$$\mathbf{\Psi}_k=(F_k,F^*_{-k},B_k,B^*_{-k},E_{2k_0,k},E_{-2k_0,k}) \tag{20}$$

is a six-dimensional vector,

$$\mathbf{J}=e^{\gamma_i z}\begin{pmatrix}\mathbf{0}&&\\&\mathbf{0}&\\&&\mathbf{1}\end{pmatrix} \tag{21}$$

is a diagonal 6×6 matrix, and $\mathbf{0}$, $\mathbf{1}$ are respectively the zero and unit 2×2 matrices. The linear operator $\mathscr{L}$ has the form

$$\mathscr{L}=-\mathbf{N}i\partial_z-\mathbf{K}k_d+\frac{1}{2}\times\begin{pmatrix}\gamma&0&-\gamma&0&-\gamma&0\\0&-\gamma^*&0&\gamma^*&0&\gamma^*\\\gamma^*&0&-\gamma^*&0&0&\gamma^*\\0&-\gamma&0&\gamma&-\gamma&0\\1&-1&-1&1&-2&0\\-1&1&1&-1&0&-2\end{pmatrix}, \tag{22}$$

where

$$\mathbf{N}=\begin{pmatrix}\mathbf{1}&&\\&\mathbf{1}&\\&&\mathbf{0}\end{pmatrix},\quad \mathbf{K}=\begin{pmatrix}-\boldsymbol{\sigma}_3&&\\&\boldsymbol{\sigma}_3&\\&&\mathbf{0}\end{pmatrix}$$

are diagonal 6×6 matrices, σ_3 is a Pauli matrix, $k_d=k^2/2k_0n_0$, and the time t has been made dimensionless by dividing it by the characteristic relaxation time of the field E at the front face of the crystal: $t/t_d(0)\to t$.

The quadratic nonlinearity in Eq. (19) is written in the form of a vector $\boldsymbol{\eta}$ that depends on the two arguments $\mathbf{\Psi}_{k_1}$ and $\mathbf{\Psi}_{k_2}$:

$$\boldsymbol{\eta}(\mathbf{\Psi}_{k_1},\mathbf{\Psi}_{k_2})=\begin{pmatrix}-\frac{\gamma}{2}E_{2k_0,k_1}B_{k_2}\\\frac{\gamma^*}{2}E_{-2k_0,k_1}B^*_{-k_2}\\\frac{\gamma^*}{2}E_{-2k_0,k_1}F_{k_2}\\-\frac{\gamma}{2}E_{2k_0,k_1}F^*_{-k_2}\\-\frac{1}{2}F_{k_1}F^*_{-k_2}-\frac{1}{2}B_{k_1}B^*_{-k_2}+F_{k_1}B^*_{-k_2}-\frac{1}{2}E_{2k_0,k_1}(F_{k_2}+F^*_{-k_2}+B_{k_2}+B^*_{-k_2})\\-\frac{1}{2}F_{k_1}F^*_{-k_2}-\frac{1}{2}B_{k_1}B^*_{-k_2}+B_{k_1}F^*_{-k_2}-\frac{1}{2}E_{-2k_0,k_1}(F_{k_2}+F^*_{-k_2}+B_{k_2}+B^*_{-k_2})\end{pmatrix}, \tag{23}$$

while the cubic nonlinearity vector $\mathbf{\Theta}$ depends on three arguments:

$$\mathbf{\Theta}(\mathbf{\Psi}_{k_1},\mathbf{\Psi}_{k_2},\mathbf{\Psi}_{k_3})=\begin{pmatrix}0\\0\\0\\0\\-\frac{1}{2}E_{2k_0,k_1}(F_{k_2}F^*_{-k_3}+B_{k_2}B^*_{-k_3})\\-\frac{1}{2}E_{-2k_0,k_1}(F_{k_2}F^*_{-k_3}+B_{k_2}B^*_{-k_3})\end{pmatrix}, \tag{24}$$

Equations (19) must be supplemented by boundary conditions. In particular, at the back face of the crystal the sideband beams B_k acquire an additional phase due to reflection of the F_k beam from the feedback mirror. In order to obtain this phase shift, let us find the difference in the optical paths of two parallel rays 1 and 2 incident at an angle θ' on the back face AC of the crystal (see Fig. 1). The optical path difference $\Delta_{1,2}$ will then have the form

$$\Delta_{1,2}=l_{AB}+l_{BC}-n_0l_{CD}+\lambda/2,$$

where

$$l_{AB}=l_{BC}=L/\cos\,\theta,\quad l_{CD}=2L\,\tan\,\theta\,\sin\,\theta',$$

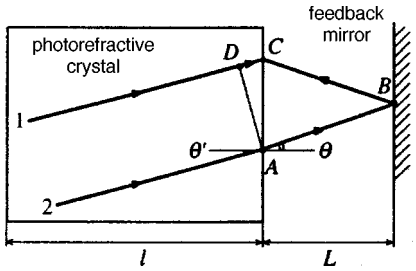

FIG. 1. Segment *DA* is perpendicular to the direction of propagation of rays *1* and *2* within the photorefractive crystal.

$$\sin\theta/\sin\theta' = n_0,$$

L is the distance between the feedback mirror and the back face of the crystal, and λ is the wavelength of the light in vacuum. Subtracting the path difference of the pump beams $2L+\lambda/2$ from the optical path difference $\Delta_{1,2}$, we obtain the required phase shift ϕ (in the small-angle approximation $\theta \ll 1$):

$$\phi = 2L(\cos\theta - 1)k_0 \simeq -2k_d n_0 L,$$

while the boundary conditions for the sideband waves have the form

$$\begin{aligned} &F_k(0) = F^*_{-k}(0) = 0,\\ &B_k(l) = F_k(l)\exp(-2ik_d n_0 L),\\ &B^*_{-k}(l) = F^*_{-k}(l)\exp(2ik_d n_0 L), \end{aligned} \tag{25}$$

where l is the length of the crystal along the z axis. Note that the signs of the exponents in (25) are reversed compared to Ref. 12. This discrepancy is obviously due to a typographical error in Ref. 12, since subsequent expressions in this paper are correct.

Thus, the problem of describing the evolution of sideband waves reduces to a boundary value problem for the system of equations (19) with boundary conditions (25). A characteristic feature of the photorefractive nonlinearity is the fact that the right side of Eq. (19), and consequently the stationary solutions of the boundary-value problem (19), (25), are independent of the pump intensity I_Σ. The pump intensity determines only the overall normalization of the optical intensities and the characteristic time t_d for setting up the stationary solutions.

3. LINEAR INSTABILITY THEORY

As a first step, let us study the evolution of sideband waves in the linear approximation. The system (19) is linearized by discarding the nonlinear (in $\mathbf{\Psi}_k$) terms $\boldsymbol{\eta}$, $\boldsymbol{\Theta}$. If we assume that the time dependence of $\mathbf{\Psi}_k$ has the form $\mathbf{\Psi}_k \propto \exp(\nu_k t)$, we obtain a linear boundary value problem for the complex eigenvalues ν_k:

$$\mathscr{L}\mathbf{\Psi}_k = \nu_k \mathbf{J}\mathbf{\Psi}_k. \tag{26}$$

Solution of this boundary value problem in the general case presents considerable difficulty, because Eq. (26) is a system of ordinary differential equations (in the coordinate z) with nonconstant coefficients. The linear boundary value problem can be solved in two special cases where the system reduces to a system of ordinary differential equations with constant coefficients. The first is the case where the coupling constant $\gamma = \gamma_r$ is real (a medium with a local photorefractive response). Then $|F_0(z)|^2 = |B_0(z)|^2 = \text{const}$. This case was investigated in Ref. 13 under the additional condition Re $\nu_k = 0$, i.e., the instability threshold was found. However, the coupling constant can be treated approximately as a real number only in photorefractive crystals like $LiNbO_3$ and $LiTaO_3$, and no one has experimentally observed the formation of hexagons in these crystals to date. In the crystals $KNbO_3$ and $BaTiO_3$ discussed in this paper, the constant is unequivocally complex; in fact, we have $\gamma \simeq i\gamma_i$ (indicating a nonlocal photorefractive response).[10,15] In the second case, the coupling constant γ is an arbitrary complex quantity, but it is assumed that the reflection coefficient from the feedback mirror is exactly equal to 1, i.e., $R=1$ (reflection from the crystal faces is neglected as before), and hence $|F_0(z)|^2 = |B_0(z)|^2$. It is also necessary to assume that $\nu_k = 0$. This implies that the boundary value problem is solved at the instability threshold Re $\nu_k = 0$, and that this instability is aperiodic, so that Im $\nu_k = 0$, i.e., the frequency detuning between pump beams and sideband beams vanishes. If it were the case that the detuning satisfied Im $\nu_k \neq 0$, a moving optical pattern would be observed in the experiment. However, the experiments of Refs. 8 and 12 showed that the optical pattern is motionless; therefore, the assumption Im $\nu_k = 0$ appears to be fully justified. The boundary value problem with $R=1$, $\nu_k = 0$ was solved in Ref. 12. In this case the system (26) takes the form

$$\begin{aligned} &\left(\frac{d}{dz} + ik_d\right)F_k = -\frac{i\gamma}{4}(F_k + F^*_{-k} - B_k - B^*_{-k}),\\ &\left(\frac{d}{dz} - ik_d\right)F^*_{-k} = \frac{i\gamma^*}{4}(F_k + F^*_{-k} - B_k - B^*_{-k}),\\ &\left(\frac{d}{dz} - ik_d\right)B_k = -\frac{i\gamma^*}{4}(F_k + F^*_{-k} - B_k - B^*_{-k}),\\ &\left(\frac{d}{dz} + ik_d\right)B^*_{-k} = \frac{i\gamma}{4}(F_k + F^*_{-k} - B_k - B^*_{-k}),\\ &E_{2k_0,k} = \frac{1}{2}(F_k - F^*_{-k} - B_k + B^*_{-k}),\\ &E_{-2k_0,k} = -\frac{1}{2}(F_k - F^*_{-k} - B_k + B^*_{-k}), \end{aligned} \tag{27}$$

where the space-charge field amplitudes were eliminated from the first four equations by using the last two equations of the system. The solution to the system (27) combined with the boundary conditions (25) leads to the threshold condition for appearance of the instability, which exactly coincides with the results of Ref. 12

$$\cos(wl)\cos(k_d l) + \frac{\gamma_i}{2w}\sin(wl)\cos[k_d(l + 2n_0 L)]$$

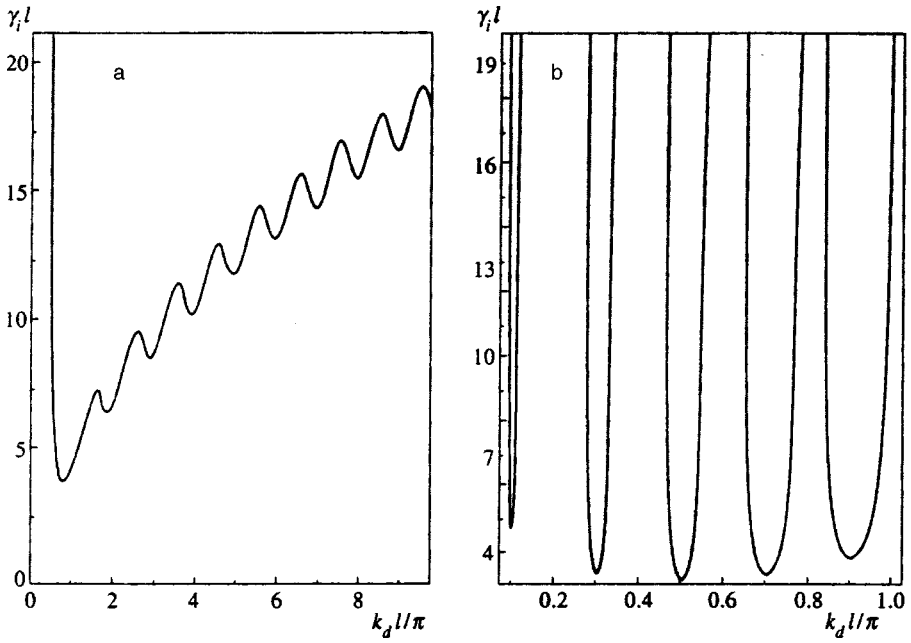

FIG. 2. Threshold dependence of $\gamma_i l$ on $k_d l$ for $L=0$ (a) and $n_0 L/l=4.44$ (b).

$$+\frac{\gamma_r+2k_d}{2w}\sin(wl)\sin(k_d l)-\frac{\gamma_r}{2w}\sin(wl)$$

$$\times\sin[k_d(l+2n_0L)]=0, \tag{28}$$

where $w=\sqrt{k_d^2+\gamma_r k_d-\gamma_i^2/4}$. As we have already noted, the coupling constant satisfies $\gamma\simeq i\gamma_i$ in the $KNbO_3$ and $BaTiO_3$ crystals to high accuracy, i.e., is pure imaginary. Therefore, in what follows we will set $\gamma=i\gamma_i$, and then obtain the threshold function $\gamma_i(k_d l)$ from Eq. (28) for each value of the distance to the feedback mirror L. This function consists of a sequence of minima. Figure 2 shows special cases of the threshold curves $\gamma_i(k_d l)$ for $L=0$ and $n_0 L/l=4.44$; in the second case, the dependence of $\gamma_i l$ on $k_d l$ is so steep that only those parts of the threshold curve near the minima could be plotted. The region of instability lies above the threshold curve $\gamma_i(k_d l)$. As L changes, the position of the minima changes; however, the number of minima and their relative spacings remain unchanged, which allows us to label them in order of increasing $k_d l$ as 1,2,3,$\cdots$, etc. In what follows a positive integer m is used to label each minimum. For $n_0 L/l<1.43\cdots$, the first minimum $m=1$ is the deepest, while as L increases the second minimum becomes deepest, then the third, etc., in succession. Figure 3 shows the L-dependence of $\gamma_i^{\min} l$ and $k_d^{\min} l$ corresponding to these minima. There is only one set of values of the parameter L for which $\gamma_i^{\min} l$ and $k_d^{\min} l$ can be found analytically, namely

$$n_0 L/l=-3/2+2m, \quad \gamma_i^{\min} l=\pi, \quad k_d^{\min} l=\pi/2, \tag{29}$$

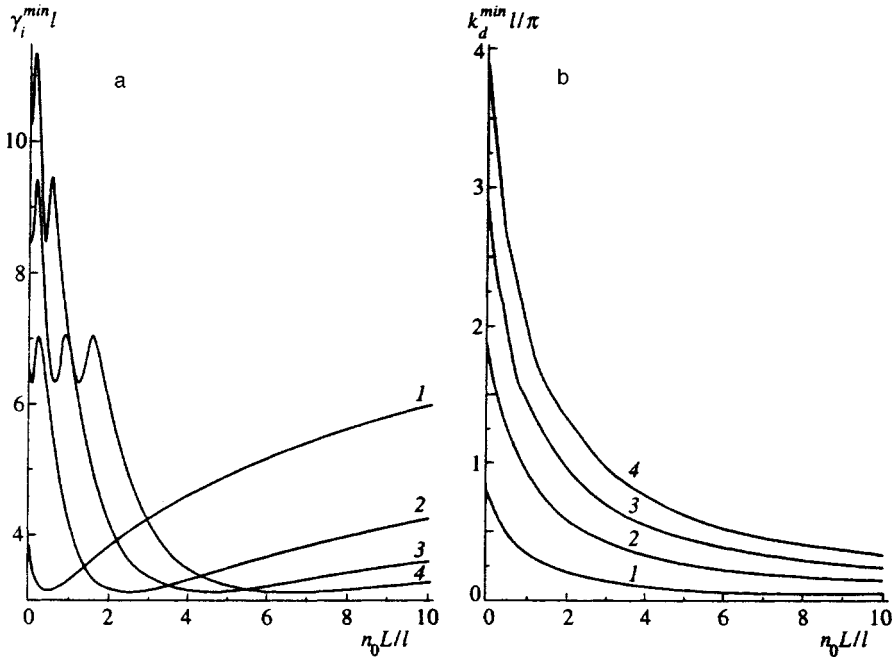

FIG. 3. Dependence of $\gamma_i^{\min} l$ (a) and $k_d^{\min} l$ (b) on the dimensionless distance to the feedback mirror $n_0 L/l$.

It is clear from Fig. 3a that at these values of L the minimum value $\gamma_i^{\min} l=\pi$ is reached as a function of L. It is important to note that for $n_0 L/l \gtrsim -3/2+2m$ the position of the mth minimum is given with high accuracy by the relation

$$k_d(l+2n_0L)=(2m-1)\pi \tag{30}$$

(which is an identity when Eq. (29) holds). The limiting case of this relation for $2n_0L/l \gg 1$, $m=1$ has been seen in experiments.[8] Actually, condition (30) implies that we are choosing that phase shift in Eq. (25) between sideband beams propagating in opposite directions along the z axis for which these beams interact most efficiently. In order to understand the physical meaning of Eq. (30), let us investigate the dependence of the space-charge field amplitude $E_{2k_0,k}$ on the z coordinate. From Eqs. (27) and (25) we find

$$\begin{aligned} F_k+B_{-k}^*&=\exp[-ik_d(z-l)][F_k(l)\\ &\quad+\exp(2ik_dn_0L)F_{-k}^*(l)],\\ F_{-k}^*+B_k&=\exp[ik_d(z-l)]\\ &\quad\times[\exp(-2ik_dn_0L)F_k(l)+F_{-k}^*(l)]. \end{aligned} \tag{31}$$

Then from Eq. (27) it follows that

$$\begin{aligned} E_{2k_0,k}&=i\,\sin[k_d(l-z+n_0L)][\exp(-ik_dn_0L)F_k(l)\\ &\quad+\exp(ik_dn_0L)F_{-k}^*(l)]. \end{aligned} \tag{32}$$

Physical considerations suggest that the lowest instability threshold corresponds to the most effective interaction between sideband waves mediated by the space-charge field $E_{2k_0,k}$. According to Eq. (32), the amplitude of this field is sinusoidal, and thus on the average the magnitude of the field amplitude will be a maximum within the crystal when the peak of the sinusoid is located at the center of the crystal $z=l/2$, from which we obtain $k_d(l/2+n_0L)=(m-1/2)\pi$, where m is a whole number, which exactly coincides with Eq. (30). Numerical calculations actually confirm that when $n_0L/l \gtrsim -3/2+2m$ holds the maximum of the amplitude $E_{2k_0,k}$ coincides to good accuracy with the center of the crystal, and that for values of k_d away from the threshold minimum $k_d=k_d^{\min}$ the maximum of this amplitude moves away from the crystal center. Nevertheless, the physical justification presented above is somewhat qualitative, because in addition to the interaction of sideband beams via the field $E_{2k_0,k}$ there is also a contribution associated with diffraction of the sideband beams by the space-charge field of the pump beams E_{2k_0}, as is apparent from the linear part of the system Eq. (19). This contribution does not allow such a simple interpretation; however, our success in explaining Eq. (30) is reason to hope that the overall physical justification is correct.

Let us now allow γ_i to exceed threshold somewhat: $0<h=(\gamma_i-\gamma_i^{\min})/\gamma_i^{\min} \ll 1$, where $\gamma_i^{\min}$ corresponds to the deepest minimum for a given value of L. Near threshold the instability growth rate can be written in the form

$$\nu_k=\nu_0-(k_d-k_d^{\min})^2 f,$$

where ν_0 is the maximum instability growth rate, $|k_d-k_d^{\min}|/k_d^{\min} \ll 1$, and $f>0$ is a constant. Because the linear boundary value problem cannot be solved analytically for $\nu_k \neq 0$, the values of ν_0 and f remain unknown. However, it will be clear from what follows that these values themselves are not important for the existence of an explosive three-wave instability. Furthermore, although ν_0 gives a correction to the hexagon amplitudes, near threshold we have $\nu_0 \to 0$ and so this correction is negligibly small. Since the value of ν_0 is positive above threshold, in the linear approximation the amplitudes F_k and B_k will grow exponentially with time until the three-wave nonlinearity becomes important. For small values of the ''supercriticality'' h, the gain of the instability is positive only within a narrow ring $|\mathbf{k}| \simeq \sqrt{2k_d^{\min}k_0n_0}$, so that when sideband waves are excited in this ring, the magnitudes of their transverse wave vectors $|\mathbf{k}|$ can be treated as practically constant.

In addition to the threshold condition (28), the linear boundary value problem Eqs. (27) and (25) allows us to find the six-dimensional eigenvector Eq. (20) $\boldsymbol{\psi}_k^{(0)}$ of this problem at the instability threshold. The zero superscript indicates that all quantities are calculated at the threshold point, and the lower case symbol $\boldsymbol{\psi}_k$ is used in place of the upper-case symbol $\boldsymbol{\Psi}_k$ to emphasize that the latter is a solution to the nonlinear boundary value problem, while the former is a solution to its linear portion only.

Let us briefly describe the procedure for finding $\boldsymbol{\psi}_k^{(0)}$: the general solution to the system of four ordinary differential equations (27) can be written in the form of a sum of four independent solutions with arbitrary coefficients $\mathbf{c}=(c_1,c_2,c_3,c_4)$. The values of these coefficients are determined from boundary conditions (25), which reduce to a homogenous system of linear equations for $\mathbf{c}$. The condition that the system be solvable is that its determinant vanish, which leads to the threshold condition (28), from which we find a solution $\mathbf{c}$ of the homogenous system of linear equations defined up to an arbitrary factor. Therefore, the eigenvector $\boldsymbol{\psi}_k^{(0)}$ is also determined up to an arbitrary factor. In particular, for $n_0L/l=-3/2+2m$, $\gamma_i^{\min} l=\pi$, $k_d^{\min} l=\pi/2$ we obtain

$$\boldsymbol{\psi}_k^{(0)} = \begin{pmatrix} -i\,\exp\left[-i\,\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right]+\exp\left(\frac{3}{4}\,i\pi\right)\exp\left(\frac{\pi z}{2l}\right)\\ i\,\exp\left[\frac{i}{2}\,\pi\left(\frac{z}{l}-\frac{1}{2}\right)\right]+\exp\left(-\frac{3}{4}\,i\pi\right)\exp\left(\frac{\pi z}{2l}\right)\\ i\,\exp\left[\frac{i}{2}\,\pi\left(\frac{z}{l}-\frac{1}{2}\right)\right]+\exp\left(\frac{1}{4}\,i\pi\right)\exp\left(\frac{\pi z}{2l}\right)\\ -i\,\exp\left[-\frac{i}{2}\,\pi\left(\frac{z}{l}-\frac{1}{2}\right)\right]+\exp\left(-\frac{1}{4}\,i\pi\right)\exp\left(\frac{\pi z}{2l}\right)\\ -i2\cos\left[\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right]\\ i2\cos\left[\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right] \end{pmatrix}. \tag{33}$$

For arbitrary values of L the explicit form of this vector

is very complicated and will not be given here. For each specific value of the parameters $\gamma_i^{\min} l$, $k_d^{\min} l$, L the eigenvector $\boldsymbol{\psi}_k^{(0)}$ was found using the program Mathematica 2.2.

In order to find the three- and four-wave interactions it is also necessary to solve a linear boundary value problem, which is the Hermitian conjugate of the linear portion of the boundary value problem Eqs. (19) and (25) with respect to the scalar product

$$\langle \boldsymbol{\psi}_k^c | \boldsymbol{\psi}_k \rangle = \int_0^l dz (\psi_k^{c\,i})^* \psi_k^i. \tag{34}$$

Here repetition of the label i implies summation from 1 to 6, and $\boldsymbol{\psi}_k^c$ is an eigenvector of this Hermitian-conjugate problem that satisfies the system of equations

$$\mathscr{L}^+ \boldsymbol{\psi}_k^c = 0 \tag{35}$$

at the instability threshold. The operator $\mathscr{L}^+ \equiv (\mathscr{L}^T)^*$ is the Hermitian conjugate of (22), and the following boundary conditions are imposed on the components of the vector $\boldsymbol{\psi}_k^c$:

$$\begin{aligned} &B_k^c(0) = B_{-k}^{c*}(0) = 0, \\ &B_k^c(l) = -F_k^c(l) \exp(-2ik_d n_0 L), \\ &B_{-k}^{c*}(l) = -F_{-k}^{c*}(l) \exp(2ik_d n_0 L), \end{aligned} \tag{36}$$

obtained by integrating the Hermitian operator id/dz by parts.

The solution of the Hermitian-conjugate boundary value problem (35), (36) is analogous to the solution of the original boundary value problem (27), (25), and the threshold condition for the Hermitian-conjugate problem coincides with the threshold condition (28) for the direct problem. The eigenvector for the conjugate problem takes the following form when $n_0 L/l = -3/2 + 2m$, $\gamma_i^{\min} l = \pi$, $k_d^{\min} l = \pi/2$:

$$\boldsymbol{\psi}_k^{c(0)} = \begin{pmatrix} \exp\left[-i\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right] + \exp\left(i\frac{\pi}{4}-\frac{\pi z}{2l}\right) \\ \exp\left[i\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right] + \exp\left(-i\frac{\pi}{4}-\frac{\pi z}{2l}\right) \\ \exp\left[i\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right] + \exp\left(i\frac{3\pi}{4}-\frac{\pi z}{2l}\right) \\ \exp\left[-i\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right] + \exp\left(-i\frac{3\pi}{4}-\frac{\pi z}{2l}\right) \\ i\pi \exp\left[-i\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right] \\ i\pi \exp\left[i\frac{\pi}{2}\left(\frac{z}{l}-\frac{1}{2}\right)\right] \end{pmatrix}. \tag{37}$$

4. THREE-WAVE INTERACTION OF SIDEBAND WAVES

The investigation of three- and four-wave interactions given here will follow several ideas taken from Refs. 18 and 19, in which the generation of hexagonal cells was discussed at the surface of a liquid dielectric in an external electric field (Refs. 18) and under conditions of weakly supercritical convection (Refs. 19).

Let us expand the general solution $\boldsymbol{\Psi}_k$ of the nonlinear boundary value problem (19), (25) within the ring $|\mathbf{k}| \simeq \sqrt{2k_d^{\min} k_0 n_0}$ in eigenvectors $\boldsymbol{\psi}_{k,n}$ of the linear boundary value problem (27), (25):

$$\boldsymbol{\Psi}_k = \sum_n \boldsymbol{\psi}_{k,n} A_{k,n}(t), \quad A_{-k} = A_k^*, \tag{38}$$

where the letter n labels the eigenmode of the linear problem for a given value of the wave vector k.

Substituting this expression into the nonlinear system (19) and taking the scalar product of the latter (as in Eq. (34)) with the eigenvector $\boldsymbol{\psi}_k^c$ of the conjugate linear problem leads to the following equation, which is accurate up to quadratic nonlinearities:

$$\frac{\partial A_{k,n}}{\partial t} = \nu_{k,n} A_{k,n} + \frac{1}{2} \sum_{n_1,n_2} \sum_{\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}} U_{k_1,k_2,k}^{n_1,n_2,n} A_{k_1,n_1} A_{k_2,n_2}, \tag{39}$$

where $U_{k_1,k_2,k}^{n_1,n_2,n}$ is the matrix element of the three-wave interaction.

Since for small values of the supercritical parameter $0 < h = (\gamma_i - \gamma_i^{\min})/\gamma_i^{\min} \ll 1$ sideband waves are excited only in the narrow ring $|\mathbf{k}| \simeq \sqrt{2k_d^{\min} k_0 n_0}$ corresponding to the mode with maximum gain and label $n=0$, while the other modes have negative gain, to find the matrix element we need only calculate it at the instability threshold $k_d = k_d^{\min}$, $\gamma = i\gamma_i^{\min}$, $\boldsymbol{\psi}_{k,n} = \boldsymbol{\psi}_k^{(0)}$, $\boldsymbol{\psi}_{k,n}^c = \boldsymbol{\psi}_k^{c(0)}$ for $n=0$. Therefore, in what follows the label n will be omitted. Furthermore, the condition $\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}$ implies that only vectors that make angles of $\pi/3$ with each other will participate in the three-wave interaction.

As a result, we obtain from Eq. (19)

$$U_{k_1,k_2,k} \equiv U = 2 \frac{\langle \boldsymbol{\psi}_k^c | \boldsymbol{\eta}_0 \rangle}{\langle \boldsymbol{\psi}_k^c | \mathbf{J} \boldsymbol{\psi}_k \rangle}, \tag{40}$$

where $\mathbf{J}$ and $\boldsymbol{\eta}$ are defined in Eqs. (21) and (23). The zero label in $\boldsymbol{\eta}_0$ indicates that its arguments are evaluated at the instability threshold, $\boldsymbol{\eta}_0 = \boldsymbol{\eta}(\boldsymbol{\psi}_k^{(0)}, \boldsymbol{\psi}_k^{(0)})$, and Eq. (39) reduces to Eq. (3).

For each specific set of values of the parameters $\gamma_i^{\min} l$, $k_d^{\min} l$, L the matrix element U was found using the program Mathematica 2.2. In the first step, the eigenvectors $\boldsymbol{\psi}_k^{(0)}$, $\boldsymbol{\psi}_k^{c(0)}$ were calculated for the direct and conjugate linear problems, and then the value of U was obtained by integration in Eq. (40). For the special cases $n_0 L/l = -3/2 + 2m$, $\gamma_i^{\min} l = \pi$, $k_d^{\min} l = \pi/2$ an analytic expression for the matrix element follows from Eqs. (33) and (37):

$$U = -\frac{2\sqrt{2}}{5} \frac{1 + 2e^{\pi}}{\cosh(\pi/2)}. \tag{41}$$

Figure 4 shows how the matrix element U for the three-wave interaction at the instability threshold depends on the distance L between the back face of the crystal and the feedback mirror (for the first minimum $m=1$ of the threshold curve $\gamma_i^{\min}(L)$), calculated from Eqs. (40) and (23), where the functions $\gamma_i^{\min}(L)$ and $k_d^{\min}(L)$ are given in Fig. 3. In this case, it was assumed that the coupling constant γ is pure imaginary,

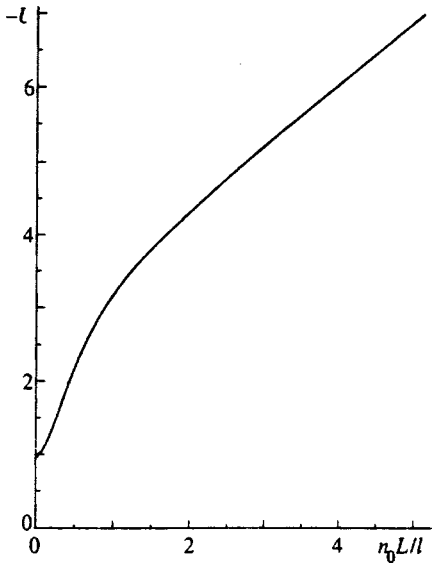

FIG. 4. Dependence of the three-wave interaction matrix element U on n_0L/l.

which is true for $KNbO_3$ and $BaTiO_3$ to high accuracy. The fact that U is a purely real quantity is very important for investigating the explosive instability.

5. FOUR-WAVE INTERACTION OF SIDEBAND WAVES

The explosive three-wave instability can be saturated by nonlinearities of fourth and higher orders. Those light beams whose wave vectors lie in the narrow ring near the instability threshold $|\mathbf{k}|\simeq\sqrt{2k_d^{\min}k_0n_0}\equiv k_{0\perp}$ will be referred to as ''fundamental'' spatial harmonics. The quadratic nonlinearities, represented by the vector $\boldsymbol{\eta}$ in the fundamental system of equations (19) and written out in Eq. (23), give rise to all possible sum and difference harmonics arising from the three-wave interaction $\mathbf{k}=\mathbf{k}_1+\mathbf{k}_2$, $|\mathbf{k}_{1,2}|=k_{0\perp}$.

Equations for the harmonics $\mathbf{k}=\mathbf{k}_1+\mathbf{k}_2$, $|\mathbf{k}_{1,2}|=k_{0\perp}$ follow from Eqs. (19) and (25), where $k_d=k^2/2k_0n_0$. The summation in the quadratic nonlinearities runs over fundamental harmonics at the instability threshold $\boldsymbol{\psi}_k=\boldsymbol{\psi}_k^{(0)}$, and time derivatives and cubic nonlinearities are neglected because the corrections they produce are of higher order (fifth order and higher). Thus, when the amplitudes of the fundamental harmonics are specified the boundary-value problem for combination harmonics reduces to the solution of a linear system of ordinary differential equations with constant coefficients and an inhomogeneous part determined by the amplitudes of the fundamental harmonics. The boundary conditions for this system are, as before, given by Eqs. (25). Solution of this system presents no fundamental difficulties; however, explicit expressions for the combination harmonics are not given here because they are extremely involved. It is important to note that these harmonics are second order in the amplitudes of the sideband waves, and therefore they contribute to the four-wave processes via the interaction with the fundamental harmonics. Furthermore, the four-wave interaction contains a contribution from the intrinsic process $2\rightarrow2$ due to interaction of the fundamental spatial harmonics that arise from the cubic nonlinearities in the fifth and sixth equations of the system (19).

Thus, the general solution $\boldsymbol{\Psi}_k^{\mathrm{tot}}$ of the nonlinear boundary value problem (19), (25) can be cast in the form

$$\boldsymbol{\Psi}_k^{\mathrm{tot}}=\boldsymbol{\Psi}_k+\delta\boldsymbol{\Psi}_k, \tag{42}$$

where $\boldsymbol{\Psi}_k$ is an expansion of (38) in fundamental harmonics and

$$\delta\boldsymbol{\Psi}_k=\sum_{\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}}A_{k_1}A_{k_2}\delta\boldsymbol{\Psi}_k$$

is an expansion in combination harmonics. The vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ lie at the instability threshold $|\mathbf{k}_1|=|\mathbf{k}_2|=k_{0\perp}$.

Substituting Eqs. (42) and (43) into the nonlinear system (19) and taking the scalar product of this system according to (34) with the vector $\boldsymbol{\psi}_k^c$ of the conjugate linear problem (27) leads to the following equation, which is accurate to within cubic nonlinearities:

$$\begin{aligned}\frac{\partial A_k}{\partial t}=\nu_kA_k+\frac{U}{2}\sum_{\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}}A_{k_1}A_{k_2}-\sum_{\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3=\mathbf{k}}&\\ \times\{[-\langle\boldsymbol{\psi}_k^{c(0)}|\boldsymbol{\eta}(\boldsymbol{\psi}_{k_1}^{(0)},\delta\boldsymbol{\Psi}_{k_2+k_3})-\boldsymbol{\eta}(\delta\boldsymbol{\Psi}_{k_2+k_3},\boldsymbol{\psi}_{k_1}^{(0)})\rangle&\\ -\langle\boldsymbol{\psi}_k^{c(0)}|\boldsymbol{\Theta}(\boldsymbol{\psi}_{k_1}^{(0)},\boldsymbol{\psi}_{k_2}^{(0)},\boldsymbol{\psi}_{k_3}^{(0)})\rangle]/\langle\boldsymbol{\psi}_k^{c(0)}|\mathbf{J}\boldsymbol{\psi}_k^{(0)}\rangle\}&\\ \times A_{k_1}A_{k_2}A_{k_3},\end{aligned} \tag{43}$$

where $\mathbf{J}$, $\boldsymbol{\eta}$ are defined in Eqs. (21) and (23), and the summation runs over fundamental harmonics. Thus, we obtain Eq. (3). The matrix element of the four-wave interaction $T_{-kk_1k_2k_3}$ is found by symmetrizing the expression in curly brackets in Eq. (43). The matrix element $T_{-kk_1k_2k_3}$ depends only on the angles between the vectors $\mathbf{k}$, $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$; therefore we will denote this matrix element by T_ϕ, where ϕ is the angle between the vectors $\mathbf{k}_1$ and $\mathbf{k}_2$.

In the special case where only six fundamental harmonics are excited with wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$, $\mathbf{k}_4$, $\mathbf{k}_5$, $\mathbf{k}_6$, forming a hexagon (Fig. 5), we obtain three types of combination harmonics: zero-order $|\mathbf{k}|\simeq0$, second-order $|\mathbf{k}|\simeq2k_{0\perp}$, and ''root-three''-order $|\mathbf{k}|\simeq\sqrt{3}k_{0\perp}$. These combination harmonics are generated by the interaction of pairs of fundamental harmonics at angles π, 0, and $\pi/3$, respectively. In Fig. 5 the second order harmonics are indicated by dotted lines, and the $\sqrt{3}$ harmonics are indicated by dashed lines. The zero-order harmonic renormalizes the pump beams, while the second-order and $\sqrt{3}$ harmonics form the vertices and centers of the faces of secondary hexagons, respectively. This is clear both from Fig. 5 and the experiments (see, e.g., Fig. 2 in Ref. 8). The interaction via second-order harmonics contributes to the matrix element T_0, while interaction via $\sqrt{3}$ harmonics contributes to $T_{\pi/3}$; interaction via zero-order harmonics contributes to both T_0 and $T_{\pi/3}$. The process $2\rightarrow2$ also contributes to both T_0 and $T_{\pi/3}$.

Let us denote the amplitudes of the six fundamental harmonics $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$, $\mathbf{k}_4$, $\mathbf{k}_5$, $\mathbf{k}_6$ by A_1, A_2, A_3, A_4, A_5, A_6. According to Eq. (38), only three amplitudes are indepen-

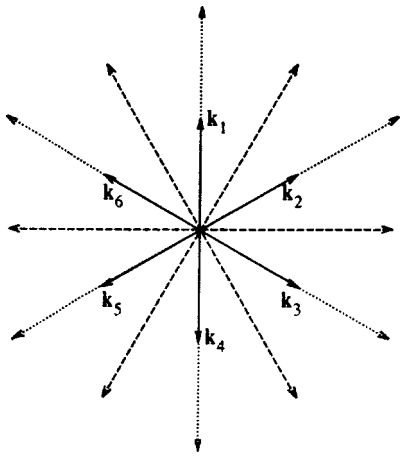

FIG. 5. The six vectors denoted by solid lines represent the fundamental harmonics. The ends of these vectors form the fundamental hexagon. The dotted lines denote the second-order harmonics, while the dashed lines are $\sqrt{3}$ harmonics. The second-order harmonics form the vertices of secondary hexagons, while the $\sqrt{3}$ harmonics are at the centers of their faces.

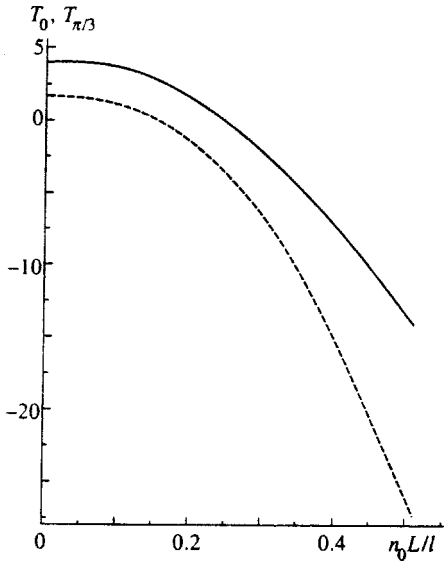

FIG. 6. Dependence of the matrix elements T_0, $T_{\pi/3}$ for four-wave interactions on n_0L/l; the solid curve is T_0, the dashed curve $T_{\pi/3}$.

dent: $A_4=A_1^*$, $A_5=A_2^*$, $A_6=A_3^*$, because $\mathbf{k}_4=-\mathbf{k}_1$, $\mathbf{k}_5=-\mathbf{k}_2$, $\mathbf{k}_6=-\mathbf{k}_3$. Then Eq, (3) can be rewritten in the form

$$\frac{\partial A_1}{\partial t}=\nu_0 A_1+UA_3^*A_5^*-\left[\frac{T_0}{2}|A_1|^2+T_{\pi/3}(|A_3|^2+|A_5|^2)\right]A_1,$$

$$\frac{\partial A_3}{\partial t}=\nu_0 A_3+UA_1^*A_5^*-\left[\frac{T_0}{2}|A_3|^2+T_{\pi/3}(|A_1|^2+|A_5|^2)\right]A_3, \tag{44}$$

$$\frac{\partial A_5}{\partial t}=\nu_0 A_5+UA_1^*A_3^*-\left[\frac{T_0}{2}|A_5|^2+T_{\pi/3}(|A_1|^2+|A_3|^2)\right]A_5.$$

Thus, the original boundary value problem (19), (25) reduces to a system of three amplitude equations.

The Mathematica 2.2 program was used to find the matrix elements T_0, $T_{\pi/3}$ for each specific value of the parameters $\gamma_i^{\min}l$, $k_d^{\min}l$, L, just as in the previous section where the three-wave interaction was discussed. Figure 6 shows the dependence of these matrix elements on the distance L between the back face of the crystal and the feedback mirror (for the first minimum $m=1$ of the threshold curve $\gamma_i^{\min}(L)$), calculated according to Eqs. (23) and (43), where $\gamma_i^{\min}(L)$, $k_d^{\min}(L)$ are given in Fig. 3. Just as for U, all of these matrix elements are found to be purely real quantities.

6. HEXAGON FORMATION DYNAMICS AND STABILITY

In the previous section, the problem of describing the evolution of hexagons was reduced to solution of the system (44) of three ordinary differential equations. When cubic nonlinearities are neglected, this system leads to an explosive instability. For equal and real amplitudes $A_1=A_3=A_5=\mathrm{Re}\,A_1$, we obtain (1) as a special case, whose corresponding solution (2) goes to infinity at finite time. In general, solutions to the system (44) (without cubic nonlinearities) can be expressed in terms of elliptic functions, and for arbitrary initial conditions (except for a set of measure zero) these solutions also exhibit singularities at finite times. It can be shown that in this limit the relative deviations $(|A_1|^2-|A_2|^2)/|A_1|^2$, $(|A_2|^2-|A_3|^2)/|A_2|^2$ go to zero, the total phase $\Phi=\mathrm{Arg}\,A_1+\mathrm{Arg}\,A_3+\mathrm{Arg}\,A_5$ goes to πn, where n is an integer, and each of the phases individually goes to a certain constant. Therefore, at later stages of evolution the system (44) reduces to the following equation for the intensity $I^2=|A_1|^2=|A_2|^2=|A_3|^2$:

$$\frac{1}{2}\frac{\partial I}{\partial t}=\nu_0 I+UI^{3/2}-\left(\frac{T_0}{2}+2T_{\pi/3}\right)I^2, \tag{45}$$

whose solution can be directly compared with experiment for small supercritical parameters when $\nu_0\to 0$.

Thus, the formation of hexagons admits the following physical picture. Due to fluctuations in the medium at time $t=0$, the amplitude A_1 is found to be nonzero for a certain value of wave vector $\mathbf{k}$ lying in the ring $|\mathbf{k}|\simeq\sqrt{2k_d^{\min}k_0n_0}$. The linear instability leads to an increase in $|A_1|$ until the nonlinear terms in Eq. (44) become important, as a result of which the amplitudes A_1, A_3, A_5 all begin to grow explosively at the same time, i.e., sideband waves are found to be excited with wave vectors $\mathbf{k}_1$, $\mathbf{k}_2$, $\mathbf{k}_3$, $\mathbf{k}_4$, $\mathbf{k}_5$, $\mathbf{k}_6$ forming a hexagon (see Fig. 5). The explosive growth of these amplitudes due to the three-wave interaction will continue until the four-wave nonlinearity comes into play. If in this case it turns out that the system parameters $\gamma_i^{\min}l$, $k_d^{\min}l$, n_0L/l are such that the total matrix element $T_0/2+2T_{\pi/3}$ is positive, then the four-wave nonlinearity can stabilize the instability; otherwise, the growth in intensity of the sideband waves con-

tinues and stabilization is achieved only via nonlinearities of higher order. In what follows, we will discuss the stability of stationary hexagonal solutions.

In experiment it has been well established[6,8] that initially two sideband waves are actually excited, with opposite signs of the transverse wave vectors. This exactly corresponds to the initial growth of amplitude A_1 for the ψ_k eigenmode Eq. (38), since this mode consists of sideband waves with transverse wave vectors $\pm\mathbf{k}$ (20). Then a rapid growth of hexagons is observed, and the intensities of all sideband waves are comparable.[8] In Ref. 8 the time dependence of the intensities of the sideband waves was measured. It is clear from Fig. 3 of Ref. 8 that after a short initial stage of exponential growth, the intensity follows a power-law increase that is characteristic of explosive nonlinearity, after which it is stabilized by higher-order nonlinearities. Evidence of this stabilization is the formation of secondary hexagons with low intensity (see Fig. 2 in Ref. 8) generated by second-order and $\sqrt{3}$ harmonics.

The stationary (hexagon) solution to Eq. (44) has the form

$$A_0=\frac{U}{4T_{\pi/3}+T_0}+\operatorname{sign}\ U\sqrt{\frac{2\nu_0}{4T_{\pi/3}+T_0}+\left(\frac{U}{4T_{\pi/3}+T_0}\right)^2}, \tag{46}$$

where $A_0=A_1=A_2=A_3$. This solution is characterized by a ''hard'' excitation regime, with an amplitude discontinuity at threshold (for $\nu_0=0$) given by

$$A_0=\frac{2U}{4T_{\pi/3}+T_0}.$$

The procedure for investigating the internal stability of the stationary solution (46) was analogous to that used in Ref. 19. This solution is stable when

$$-\frac{1}{2}\frac{U^2}{T_0+4T_{\pi/3}}<\nu_0<4\frac{T_0+T_{\pi/3}}{(2T_{\pi/3}-T_0)^2}U^2. \tag{47}$$

This result, when evaluated near the instability threshold where $\nu_0\to 0$, implies that $T_0+T_{\pi/3}>0$. Figure 7 shows the dependence of the hexagon intensity $I=A_0^2$ on n_0L/l in that range of the parameter L where the stability condition (47) holds. Outside this range, explosive growth of the hexagons can be stabilized only by higher-order wave processes.

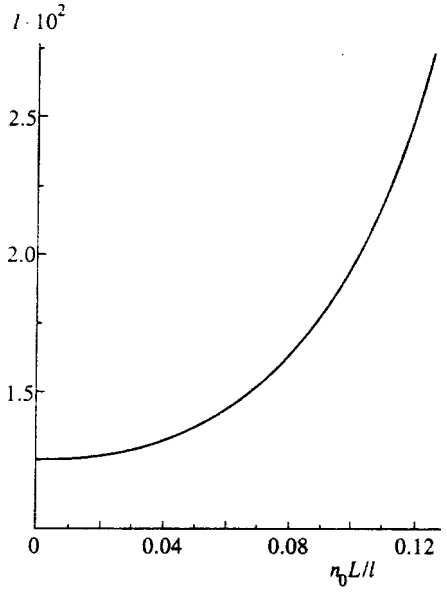

FIG. 7. Intensity I of the fundamental hexagon at the front face of the photorefractive crystal $z=0$ as a function of n_0L/l, which follows from the theory of four-wave interactions in the region where this theory is applicable. The intensity of the hexagon is normalized by the intensity of the pump at the front face of the crystal.

7. NUMERICAL EXPERIMENT

A numerical experiment was performed in order to verify the results of the analytic theory for hexagon generation described above. The goal of this experiment was to investigate the region of large values of $n_0L/l\gtrsim 0.1$, where saturation of the explosive instability is provided by higher-order wave processes (five-wave and higher). In the experiment the boundary value problem (19), (25) was solved numerically, taking into account a larger number of sum and difference harmonics than in the previous section. In order to estimate the number of higher-order harmonics required, and accordingly the order of the wave processes that must be included in order to definitely ensure saturation of the explosive instability, the following integral of the boundary value problem (19), (25) was used:

$$H=\sum_k\ (|F_k|^2-|B_k|^2)+[F_k+F_k^*-B_k-B_k^*]\Big|_{k=0}=0, \tag{48}$$

where the summation runs over all harmonics. The vanishing of this integral for all z at all times t physically expresses conservation of the energy of the optical field, since the dissipation of optical energy, which is small in a photorefractive crystal, has nowhere been taken into account in deriving the system (19), (25). In particular, at $z=0$ (i.e., at the front face of the crystal) the condition $H=0$ implies that the optical power of the pump beam incident on the crystal equals the total optical power of the beams that exit the crystal by its back face (recall that we have neglected reflection from the boundary and have set the reflection coefficient of the feedback mirror equal to unity). According to the boundary condition (25), we have $F_k|_{z=0}=0$ for all k; therefore it follows from Eq. (48) that when sideband waves form the pump is depleted, which is expressed in the growth of the zero-order harmonics $F_k|_{k=0}$, $B_k|_{k=0}$ which renormalize the pump beams. Thus, in the strongly nonlinear theory we must at a minimum take into account all processes in which zero harmonics interact with each other. It is not difficult to see that such processes give contributions up to eighth order in the equation for the fundamental harmonic. Therefore, in the numerical experiment all processes up to eight-wave inclusively were taken into account. In this case, it is necessary to include along with the harmonics 0, 1, 2, $\sqrt{3}$ listed above the harmonics $\sqrt{7}$, 3, $\sqrt{12}$, $\sqrt{13}$, 4 as well. (The ends of the

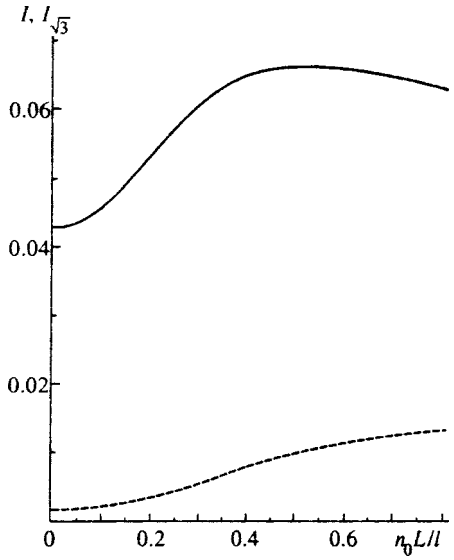

FIG. 8. Dependence of the intensity of the fundamental hexagon I (solid curve) and intensity of the $\sqrt{3}$ harmonic ($I_{\sqrt{3}}$ is dashed) on n_0L/l obtained from the numerical experiment. The normalization of the intensity is the same as in Fig. 7.

wave vectors of harmonics 3, $\sqrt{12}$, 4 form the vertices of hexagons, while the ends of the wave vectors of harmonics $\sqrt{7}$, $\sqrt{13}$ form regular dodecagons).

At each time t the boundary value problem (19), (25) was solved by Newton's method. At each step of this method, values of the amplitude $B_k|_{z=0}$ were specified at the front face of the crystal for all the harmonics. Then the amplitudes $F_k|_{z=l}$, $B_k|_{z=l}$ at the back face of the crystal were found by integrating the first four ordinary differential equations of the system (19) using the fourth-order Runge-Kutta method in the coordinate z. The error used in Newton's method was given by the accuracy with which the boundary condition (25) was satisfied at the back face of the crystal. The time dependence was determined by integrating the space charge field amplitudes $E_{\pm 2k_0,k}$ using a predictor-corrector method (the fifth and sixth equations of the system (19)).

The results of the numerical experiment are shown in Fig. 8 in the form of plots of the intensities of the fundamental harmonic and $\sqrt{3}$ harmonic versus the distance to the feedback mirror for $0 \leqslant n_0L/l \leqslant 0.85$. In the region $0 \leqslant n_0L/l \lesssim 0.1$, the intensity of the fundamental harmonic has a minimum, which is in agreement with the results of the analytic theory according to which fourth-order processes can saturate the explosive growth of the hexagon instability only in this region. However, in this case the analytical $B_k^{\rm an}$ and numerical $B_k^{\rm num}$ values of the light-beam amplitudes differ rather strongly: $|B_k^{\rm num}|/|B_k^{\rm an}||_{z=0} \sim 2$, which indicates a strong renormalization of the amplitudes of the hexagons due to higher-order wave processes. Thus, the predictions of the four-wave theory are valid more qualitatively than quantitatively. For $n_0L/l \gtrsim 0.1$, when the four-wave interactions surely cannot ensure saturation of the explosive instability, the nonlinearity increases with increasing n_0L/l.

Additional numerical investigations showed that, since the amplitudes of the higher-order harmonics rapidly decay as their labels increase, if we assume by definition that all the higher-order combination harmonics equal zero but take into account all possible processes between the 0, 1, 2, $\sqrt{3}$ harmonics, the amplitude of the fundamental harmonic changes by more than 2%. Thus, for the strongly nonlinear theory it actually turns out to be sufficient to take into account wave processes up to eighth order, while including in those processes only interactions with the participation of harmonics 0, 1, 2, $\sqrt{3}$. Note also that only harmonics 0, 1, 2, $\sqrt{3}$ are observed in experiment, while the higher harmonics are too weak to be recorded (see, e.g., Refs. 6 and 8). In these experiments the intensity of hexagons is observed to be from one to three percent of the intensity of the pump, which is somewhat lower than the results obtained from the numerical experiment, which gives a value $|B_k(0)|^2 \simeq 0.04$. This is probably explained by the fact that losses due to reflection at the crystal faces (of order 15% for light incident on the back face of the crystal and just as much again when the light reenters the crystal after reflection from the feedback mirror) have not been included in the system (19), (25), and losses in the reflection of light from the feedback mirror have also been neglected.

We mention in conclusion that for $n_0L/l \gtrsim 0.85$ stationary hexagon solutions turn out to be unstable against repumping of energy back into the $\sqrt{3}$ harmonic. This instability is connected with the fact already noted in Sec. 3 that for $n_0L/l \simeq 1.43$ the depths of the first and second minima of the threshold curve $\gamma_i^{\rm min}(k_dl)$ for the linear instability are comparable (see Fig. 3a). In this case the ratio $k_d^{\rm min}|_{m=2}/k_d^{\rm min}|_{m=1}$ is close to 3 (accurate to 2%). Thus, in the neighborhood of $n_0L/l \sim 1.43$ the amplitudes of the first-order and $\sqrt{3}$ harmonics are the same order, and hence they are coupled by the three-wave interaction. This case requires the inclusion of all wave processes up to eight-wave in the numerical calculations, and thus the treatment of a larger number of harmonics than were included in the numerical experiment described above. Consideration of this problem lies outside the framework of this paper.

8. CONCLUSION

Thus, the following results have been obtained in this paper. A system of equations (19) has been derived with boundary conditions (25) that describes the evolution of the sideband wave amplitudes and space-charge field for an arbitrary level of nonlinearity. Linearization of this system leads to the threshold condition (28) for appearance of a transverse instability, which exactly coincides with the results of Ref. 12. It has been shown that the threshold curve $\gamma_i(k_dl)$ consists of a sequence of minima whose relative depth changes with the parameter n_0L/l, which is proportional to the distance L to the feedback mirror. For $n_0L/l < 1.43$ the deepest minimum (i.e., corresponding to the lowest instability threshold) turns out to be the first, and as the quantity n_0L/l increases the second, third, etc., minima become the deepest in succession. Expression (30) was obtained, which describes with high accuracy the position of all the minima. In this case the magnitudes of the transverse

wave vectors belong to the set 1, $\sqrt{3}$, $\sqrt{5}$,... for the first, second, third, etc. minima. Moreover, a set of explicit solutions (29) has been found for the threshold Eq. (28) which correspond to the deepest values of the minima for all possible values of the parameter n_0L/l. Eigenvectors were calculated for the direct linear boundary value problem (25), (26) and its Hermitian conjugate (36), (35) at the instability threshold as functions of n_0L/l. For the special cases (29) explicit analytic expressions (33), (37) were given for these vectors.

For small values of the supercritical parameter h, when the sideband beams are unstable only within a narrow ring $|\mathbf{k}| \simeq \sqrt{2k_d^{\min}k_0n_0}$, the general solution to the nonlinear boundary value problem (19), (25) was reduced (by expanding (38) in eigenfunctions of the linear problem) to the system of amplitude equations (3). This system consists of a Landau expansion in the amplitude of the growing linear modes. The matrix elements U and $T_{kk_1k_2k_3}$ for three- and four-wave interactions respectively were calculated at the instability threshold as functions of n_0L/l. These matrix elements turn out to be purely real quantities. In the special case (29) the explicit analytic expression (41) was obtained for U. The fact that the matrix element U for the three-wave interaction differs from zero is of fundamental importance, since it leads to the appearance of the explosive three-wave instability and correlation of sideband waves whose wave-vectors make angles with one another that are multiples of $\pi/3$.

The possibility of stabilization of explosive growth of the hexagons due to four-wave interactions was investigated, Eqs. (44) were obtained to describe the temporal dynamics of generation of the steady-state hexagonal solutions (46), and the stability condition (47) for these solutions was found. As a result it was shown that for $n_0L/l \lesssim 0.1$ four-wave interactions can saturate the explosive instability. Outside this region it is definitely necessary to take into account higher-order wave processes. A numerical experiment was performed to investigate the nonlinear boundary value problem (19), (25). It was shown that in general it is necessary to take into account all wave processes up to eight-wave. However, among these processes the only important ones were interactions between the fundamental harmonic and harmonics 0, $\sqrt{3}$, and 2. For $n_0L/l \lesssim 0.1$ the results obtained agreed quantitatively with the results of the four-wave interaction theory; however, it was found that higher-order wave processes lead to a substantial renormalization of the hexagon amplitudes given by the analytic expression (46). It was shown that for $n_0L/l \gtrsim 0.85$ the strongly nonlinear stationary hexagonal solutions (obtained in the numerical experiment) become unstable against repumping of energy into higher harmonics. The reason for this is that in this region the fundamental harmonic and $\sqrt{3}$ harmonic are of the same order due to the closeness of their linear instability thresholds. In this case the transverse wave vectors of these two harmonics correspond to positions of two successive minima of the threshold curve $\gamma_i(k_dl)$, which leads to an intense exchange of energy between them.

In conclusion the author thanks E. A. Kuznetsov for posing the problem and for useful discussions. This work was carried out with the support of the Russian Fund for Fundamental Research (Grant 97-01-00093), the program of government support for leading scientific schools (Grant 96-15-96093), the Landau Scholarship Fund, KFA, Forschungzentrum, Juelich, Germany, and the INTAS Fund (Grant 96-0954).

*)E-mail: lushniko@landau.ac.ru

[1] J. Pender and L. Hesselink, J. Opt. Soc. Am. B **7**, 1361 (1990).
[2] A. Petrossian, M. Pinard, A. Maître, J.-Y. Courtois, and G. Grynberg, Europhys. Lett. **18**, 689 (1992).
[3] R. MacDonald and H. J. Eichler, Opt. Commun. **89**, 289 (1992).
[4] M. Tamburrini, M. Bonavita, S. Wabnitz, and E. Santamato, Opt. Lett. **18**, 855 (1993).
[5] J. V. Moloney and A. C. Newell, *Nonlinear Optics* (Addison-Wesley, Reading MA, 1992).
[6] T. Honda, Opt. Lett. **18**, 598 (1993).
[7] P. P. Banerjee, H.-L. Yu, D. A. Gregory, N. Kukhtarev, and H. J. Caulfield, Opt. Lett. **20**, 10 (1995).
[8] T. Honda and A. Matsumoto, Opt. Lett. **20**, 1755 (1995).
[9] N. V. Kukhtarev, T. Kukhtareva, H. J. Caulfield, P. P. Banerjee, H.-L. Yu, and L. Hesselink, Opt. Eng. (Bellingham) **34**, 2261 (1995).
[10] *Photorefractive Materials and Their Applications*, P. Günter and J.-P. Huignard eds. (Springer–Verlag, Berlin, 1988); M. P. Petrov, S. I. Stepanov, and A. V. Khomenko, *Photorefractive Crystals in Coherent Optical Systems* (Nauka, St. Petersburg, 1992; Springer, New York, 1991).
[11] W. J. Firth and C. Paré, Opt. Lett. **13**, 1096 (1988); G. D'Alessandro and W. J. Firth, Phys. Rev. Lett. **66**, 2597 (1991).
[12] T. Honda and P. P. Banerjee, Opt. Lett. **21**, 779 (1996).
[13] B. I. Sturman and A. I. Chernykh, Zh. Éksp. Teor. Fiz. **111**, 1611 (1997) [JETP **84**, 881 (1997)].
[14] L. D. Landau and E. M. Lifshits, *Fluid Mechanics* (2nd ed. Pergamon Press, Oxford, 1987) [Russ. original Nauka, Moscow, 1988].
[15] B. I. Sturman and V. M. Fridkin, *The Photovoltaic and Photorefractive Effects in Noncentrosymmetric Materials* (Nauka, Moscow, 1992); (Gordon & Breach, New York, 1992); B. I. Sturman, S. G. Odulov, and M. Yu. Goulkov, Phys. Rep. **275**, 197 (1996).
[16] N. V. Kukhtarev, Pis'ma Zh. Tekh. Fiz. **2**, 1114 (1976) [Sov. Tech. Phys. Lett. **2**, 438 (1976)].
[17] M. Saffman, A. A. Zozulya, and D. Z. Anderson, J. Opt. Soc. Am. B **11**, 1409 (1994).
[18] E. A. Kuznetsov and M. D. Spektor, Zh. Éksp. Teor. Fiz. **71**, 262 (1976) [Sov. Phys. JETP **44**, 136 (1976)].
[19] E. A. Kuznetsov and M. D. Spektor, Appl. Math. Theor. Phys. **2**, 76 (1980) [in Russian].

Translated by Frank J. Crowne

ERRATA

Erratum: NMR in $^{55}Mn^{2+}$ nuclei in the quasi-one-dimensional antiferromagnetic $CsMnBr_3$ [JETP 86, 197–205 (January 1998)]

A. S. Borovik-Romanov, S. V. Petrov, and A. M. Tikhonov

Kapitsa Institute of Physics Problems, Russian Academy of Sciences, 117334 Moscow, Russia

B. S. Dumesh

Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

[S1063-7761(98)03103-5]

Equation (9) should read as follows:

$$\mathscr{L}_n = \frac{1}{\gamma_n}\left(\langle \mathbf{m} \rangle, \mathbf{\Omega} + \gamma_n \mathbf{H}_{\text{eff}}\right) \qquad (9)$$

