# Theory of superfluidity of nuclear matter based on the Fermi-liquid approach

A. I. Akhiezer, A. A. Isaev, S. V. Peletminskiĭ, A. P. Rekalo, and A. A. Yatsenko

*Kharkov Physicotechnical Institute, National Scientific Center, 310108 Kharkov, Ukraine*
(Submitted 28 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 3–24 (July 1997)

We use Landau's concept of a Fermi liquid to study the theory of superfluidity of symmetric nuclear matter. For the nucleon–nucleon potential we take the effective Skyrme interaction (the Ska, SkM, SkM*, and RATP potentials). The density-dependence of the transition temperature is studied for different superfluid phases of nuclear matter. We show that the phase in which there is proton–nuclear pairing in the spin-triplet state is realized at densities close to the saturation density. We demonstrate that phase transitions in density from the given phase to a phase with singlet–singlet or triplet–triplet nucleon pairing are possible. The density-dependence at $T=0$ of the energy gap in the quasiparticle spectrum is established for the case of unitary and nonunitary spin states. Finally, we establish that the phase transition to a nonunitary phase is accompanied by the appearance of magnetization, which is found as a function of the nuclear matter density. 
[S1063-7761(97)00107-8]

## 1. INTRODUCTION

It is known that an atomic nucleus can be in a superfluid state. The reality of superfluidity in nuclei is proved by such experimental facts as the presence of a gap in the energy spectra of even–even nuclei, the moments of inertia in deformed nuclei, explained by the presence of a superfluid component in nuclear matter, and the special features of excited states of nuclei and of the probabilities of $\alpha$- and $\beta$-decays.[1–5] A consistent theory of superfluidity in nuclei is much more difficult because of the finite size of nuclei. Hence it would be interesting to study the superfluidity in infinite nuclear matter, especially in connection with astrophysical applications.

In this paper we study the superfluidity of nuclear matter using Landau's semiphenomenological concept of a Fermi liquid.[6] For simplicity we assume that the energy functional is invariant under rotations in the configurational, spin, and isospin spaces (for infinite uniform nuclear matter the spin–orbit coupling is zero). Hence superfluid phases (Cooper pairs) are classified by specifying the values of the quantum numbers related to the total spin of the pair, $S=0,1$, the isospin $T=0,1$, their projections $S_z$ and $T_z$ on the $z$ axis, and the orbital angular momentum $L=0, 1, 2, \ldots$ . The possible values of the orbital angular momentum $L$ for each value of $S$ and $T$ must be found according to the Pauli principle: at $S=0$, $T=0$ or $S=1$, $T=1$ the values of $L$ are $1, 3, 5, \ldots$, and at $S=1$, $T=0$ or $S=0$, $T=1$ we have $L=0, 2, 4, \ldots$ . Each superfluid phase is described by its own set of order parameters: at $S=0$, $T=0$ by a scalar order parameter $\Delta_{00}$, at $S=1$, $T=0$ or $S=0$, $T=1$ by the vector order parameter $\Delta_{k0}$ or $\Delta_{0k}$, respectively ($k=1,2,3$), and at $S=1$, $T=1$ by the tensor order parameter $\Delta_{ik}$ ($i,k=1,2,3$) (we are speaking of scalars, vectors, and tensors in the spin and isospin spaces). These order parameters determine the corresponding gaps in the quasiparticle energy spectrum.

We use the Landau concept of a Fermi liquid to set up the equations for finding the equilibrium normal ($f$) and anomalous ($g$) distribution functions for the initial fermions (self-consistency equations) and, in accordance with the above classification, find a simpler system of equations for the components of the distribution functions (or the order parameters). An analysis of these equations makes it possible to find the transition temperatures for different phases of superfluid nuclear matter and the values of the energy gaps in the Cooper-pair spectrum. Note that in the case of nonunitary states (either $S_z$ or $T_z$ is nonzero) a phase transition to the superfluid state is also a phase transition into a spin-ordered or isospin-ordered state (the latter means that the pair consists of either a proton and another proton or a neutron and another neutron). Here the transition temperatures, in contrast to energy gaps (for symmetric nuclear matter, where the neutron and proton densities are equal), are independent of $S_z$ and $T_z$ but strongly depend on $S$ and $T$.

Note that the method we develop differs from the Green's function method commonly used in phase transition theory and the BCS theory.[7–11] The transition temperatures for the ${}^1S_0$ and ${}^3S_1-{}^3D_1$ pairings (in the case of noncentral forces) for nuclear matter were obtained in Refs. 12–16. Phase transitions in the case of triplet pairing in the spin space within a BCS-theory setting that uses separable potentials were studied in Refs. 17–19. Various potentials have been used to describe the nucleon–nucleon interaction; among these are the Paris, Graz, Mongan, and Reid potentials. Alm *et al.*[20,21] studied ${}^3D_2$ pairing in asymmetric nuclear matter and the possibility of its realization in the core of massive neutron stars.

The general theory developed here requires no specific potential for the nucleon–nucleon interaction. However, to obtain numerical results we employ the effective Skyrme interaction[22] as the nucleon–nucleon potential. This interaction is self-consistent: it depends on the nuclear matter density and effectively takes into account the multiparticle interaction of nucleons.[23] We use the variants of the Skyrme interaction that describe the properties of nuclei in the best possible way.[24]

In this paper we do numerical calculations of the dependence of the transitions temperatures and order parameters for various phases on the density in the case of symmetric nuclear matter. For nonunitary states we study the dependence of magnetization on the density of superfluid nuclear matter.

## 2. BASIC EQUATIONS

A superfluid Fermi liquid is described by two fermion distribution functions: the normal distribution function $f_{\kappa_1\kappa_2}=\langle a^+_{\kappa_2}a_{\kappa_1}\rangle$, with $f^*_{\kappa_1\kappa_2}=f_{\kappa_2\kappa_1}$, where $a^+_\kappa$ and $a_\kappa$ are the creation and annihilation operators for fermions with momentum $\mathbf{p}$, spin (isotopic spin) projection $\alpha$ $(a)$, and $\kappa\equiv(\mathbf{p},\alpha,a)$, and the anomalous distribution function $g_{\kappa_1\kappa_2}=\langle a_{\kappa_2}a_{\kappa_1}\rangle$, with $g_{\kappa_1\kappa_2}=-g_{\kappa_2\kappa_1}$ (here $\langle\cdots\rangle=\mathrm{Tr}\,\rho\cdots$ is the mathematical expectation, or average, of the operators, with $\rho$ the density matrix of the system). It is convenient to combine the distribution functions $f$ and $g$ into a matrix distribution function $\hat f$:

$$\hat f_{\kappa_1\kappa_2}=\begin{pmatrix} f_{\kappa_1\kappa_2} & g_{\kappa_1\kappa_2}\\ g^+_{\kappa_1\kappa_2} & \delta_{\kappa_1\kappa_2}-f_{\kappa_2\kappa_1}\end{pmatrix}. \tag{2.1}$$

We will call $\hat f_{\kappa_1\kappa_2}$ the statistical operator of a nonequilibrium superfluid Fermi liquid.[25]

Landau's concept of a Fermi liquid is based on specifying the energy of the Fermi system as a functional of the normal $(f)$ and anomalous $(g)$ fermion distribution functions, $E=E(f,g)$. The energy functional determines the fermion single-particle energy $\varepsilon$,

$$\varepsilon_{\kappa_1\kappa_2}=\frac{\partial E}{\partial f_{\kappa_2\kappa_1}},\quad \varepsilon^*_{\kappa_1\kappa_2}=\varepsilon_{\kappa_2\kappa_1}, \tag{2.2}$$

and the matrix order parameter of the system,

$$\Delta_{\kappa_1\kappa_2}=2\frac{\partial E}{\partial g^+_{\kappa_2\kappa_1}},\quad \Delta_{\kappa_2\kappa_1}=-\Delta_{\kappa_1\kappa_2}. \tag{2.3}$$

Thermodynamically, the equilibrium state of a superfluid Fermi liquid is determined from the condition for maximizing the entropy at the given values of the system energy $E(f,g)$ and the numbers of particles, $N_a$ and $N_b$, of species $a$ and $b$, which leads to a self-consistency equation for determining the distribution functions $f$ and $g$ (or, what is the same, the energy $\varepsilon_{\kappa_1\kappa_2}$ and the order parameter $\Delta_{\kappa_1\kappa_2}$) as functions of the temperature $T$ and the chemical potentials $\mu_a$ and $\mu_b$ of the system:

$$\hat f=\frac{1}{\exp(Y_0\hat\varepsilon+\hat Y_4)+1}\equiv\frac{1}{\exp(Y_0\hat\xi)+1}. \tag{2.4}$$

Here the matrices $\hat\varepsilon$ and $\hat Y_4$ have the following form:

$$\hat\varepsilon=\begin{pmatrix}\varepsilon & \Delta\\ \Delta^+ & -\tilde\varepsilon\end{pmatrix},\quad \hat Y_4=\begin{pmatrix}Y_4 & 0\\ 0 & -Y_4\end{pmatrix},$$

where $\varepsilon$, $\Delta$, and $Y_4$ are, in turn, matrices in the space of the $\kappa$ variables, with $Y_{4\kappa_1\kappa_2}=Y_{4a_1}\delta_{\kappa_1\kappa_2}$, and the tilde stands for the transposition operation. The parameters $Y_0$, $Y_{4a}$, and $Y_{4b}$ in (2.4) are related to the temperature $T$ and the chemical potentials $\mu_a$ and $\mu_b$ by the following formulas:

$$Y_0=\frac{1}{T},\quad Y_{4a}=-\frac{\mu_a}{T},\quad Y_{4b}=-\frac{\mu_b}{T}. \tag{2.5}$$

General investigations into the properties of the self-consistency equation (2.4) were conducted in Refs. 25–28. These included the study of diagonalization, the construction of the appropriate thermodynamics and hydrodynamics, and the inclusion of external electromagnetic fields (derivation of the London equation[29] and the Ginzburg–Landau equation[30–32]). In the present paper we limit our discussion to the case where the functional $E(f,g)$ is quadratic in $f$ and $g$ and is invariant under rotations in the spin and isospin spaces. We also assume that the state of the system is spatially uniform.

Let us expand the normal distribution function $f$ in the Pauli matrices $\sigma_i$, $\tau_k$, and $\sigma_i\tau_k$, i.e.,

$$\begin{aligned} &f_{p_1\alpha_1a_1,p_2\alpha_2a_2}=f(\mathbf{p}_1)_{\alpha_1a_1,\alpha_2a_2}\delta_{p_1,p_2},\\ &f(\mathbf{p}_1)=f_{00}(\mathbf{p}_1)+f_{k0}(\mathbf{p}_1)\sigma_k+f_{0k}(\mathbf{p}_1)\tau_k+f_{ik}(\mathbf{p}_1)\sigma_i\tau_k,\end{aligned} \tag{2.6}$$

and the anomalous distribution function $g$ in the matrices $\sigma_2$, $\sigma_i\sigma_2$, $\tau_2$, and $\tau_k\tau_2$, i.e.,

$$\begin{aligned} &g_{p_1\alpha_1a_1,p_2\alpha_2a_2}=g(\mathbf{p}_1)_{\alpha_1a_1,\alpha_2a_2}\delta_{p_1,-p_2},\\ &g(\mathbf{p}_1)=g_{00}(\mathbf{p}_1)\sigma_2\tau_2+g_{k0}(\mathbf{p}_1)\sigma_k\sigma_2\tau_2+g_{0k}(\mathbf{p}_1)\sigma_2\tau_k\tau_2\\ &\qquad+g_{kl}(\mathbf{p}_1)\sigma_k\sigma_2\tau_l\tau_2.\end{aligned} \tag{2.7}$$

We write the energy functional $E(f,g)$ in the form

$$E(f,g)=E_0(f)+E_{\mathrm{int}}(f)+E_{\mathrm{int}}(g), \tag{2.8}$$

where

$$E_0(f)=4\sum_{\mathbf{p}}\varepsilon_0(\mathbf{p})f_{00}(\mathbf{p}),\quad \varepsilon_0(\mathbf{p})=\frac{\mathbf{p}^2}{2m}, \tag{2.9}$$

$m$ is the nucleon mass. In accordance with the requirement that the energy functional be invariant under rotations in the spin and isospin spaces, for $E_{\mathrm{int}}(f)$ and $E_{\mathrm{int}}(g)$ we have

$$\begin{aligned} E_{\mathrm{int}}(f)=\frac{2}{\mathscr{V}}\sum_{\mathbf{p},\mathbf{q}}[&f_{00}(\mathbf{p})U_0(\mathbf{p},\mathbf{q})f_{00}(\mathbf{q})\\ &+f_{k0}(\mathbf{p})U_1(\mathbf{p},\mathbf{q})f_{k0}(\mathbf{q})\\ &+f_{0k}(\mathbf{p})U_2(\mathbf{p},\mathbf{q})f_{0k}(\mathbf{q})\\ &+f_{ik}(\mathbf{p})U_3(\mathbf{p},\mathbf{q})f_{ik}(\mathbf{q})],\end{aligned} \tag{2.10}$$

$$\begin{aligned} E_{\mathrm{int}}(g)=\frac{2}{\mathscr{V}}\sum_{\mathbf{p},\mathbf{q}}[&g^*_{00}(\mathbf{p})V_0(\mathbf{p},\mathbf{q})g_{00}(\mathbf{q})\\ &+g^*_{k0}(\mathbf{p})V_1(\mathbf{p},\mathbf{q})g_{k0}(\mathbf{q})\\ &+g^*_{0k}(\mathbf{p})V_2(\mathbf{p},\mathbf{q})g_{0k}(\mathbf{q})\\ &+g^*_{ik}(\mathbf{p})V_3(\mathbf{p},\mathbf{q})g_{ik}(\mathbf{q})],\end{aligned} \tag{2.11}$$

where $\mathscr{V}$ is the volume occupied by the system. The quantities $U$ and $V$ are known as the normal and anomalous amplitudes of the Landau interaction in a superfluid Fermi liquid.

The quasiparticle energy $\varepsilon_{\kappa_1\kappa_2}$ and the order parameter $\Delta_{\kappa_1\kappa_2}$ will also be expanded in the Pauli matrices $\sigma_k$ and $\tau_k$:

$$\varepsilon_{\kappa_1\kappa_2}=\varepsilon_{\alpha_1\alpha_2 a_1 a_2}(\mathbf{p}_1)\,\delta_{p_1,p_2},$$
$$\Delta_{\kappa_1\kappa_2}=\Delta_{\alpha_1\alpha_2 a_1 a_2}(\mathbf{p}_1)\,\delta_{p_1,-p_2}, \tag{2.12}$$

where $\varepsilon_{\alpha_1\alpha_2 a_1 a_2}(\mathbf{p})$ and $\Delta_{\alpha_1\alpha_2 a_1 a_2}(\mathbf{p})$ are the respective matrices in the spin and isospin spaces:

$$\varepsilon(\mathbf{p})=\varepsilon_{00}(\mathbf{p})+\varepsilon_{k0}(\mathbf{p})\sigma_k+\varepsilon_{0k}(\mathbf{p})\tau_k+\varepsilon_{ik}(\mathbf{p})\sigma_i\tau_k,$$
$$\Delta(\mathbf{p})=\Delta_{00}(\mathbf{p})\sigma_2\tau_2+\Delta_{k0}(\mathbf{p})\sigma_k\sigma_2\tau_2+\Delta_{0k}(\mathbf{p})\sigma_2\tau_k\tau_2 \tag{2.13}$$
$$+\Delta_{kl}(\mathbf{p})\sigma_k\sigma_2\tau_l\tau_2.$$

The condition $\Delta_{\kappa_1\kappa_2}=-\Delta_{\kappa_2\kappa_1}$ and the expansion of $\Delta(\mathbf{p})$ in (2.13) imply that $\Delta_{00}(\mathbf{p})$ and $\Delta_{ik}(\mathbf{p})$ are odd functions of $\mathbf{p}$ while $\Delta_{i0}(\mathbf{p})$ and $\Delta_{0i}(\mathbf{p})$ are even functions of $\mathbf{p}$. These symmetry properties lead to a situation in which singlet–singlet and triplet–triplet pairing of nucleons can occur in states with odd values of the orbital angular momentum $L$, while the singlet–triplet and triple–singlet pairing of nucleons can occur in states with even values of $L$.

According to the definitions (2.2) and (2.3), the quantities $\varepsilon$ and $\Delta$ can be related to the functions $f$ and $g$. Precisely, the $\varepsilon$ are related to the $f$ as

$$\varepsilon_{00}(\mathbf{p})=\varepsilon_0(\mathbf{p})+\frac{2}{\mathscr{V}}\sum_{\mathbf{p}'}U_0(\mathbf{p},\mathbf{p}')f_{00}(\mathbf{p}'),$$
$$\varepsilon_{i0}(\mathbf{p})=\frac{2}{\mathscr{V}}\sum_{\mathbf{p}'}U_1(\mathbf{p},\mathbf{p}')f_{i0}(\mathbf{p}'),$$
$$\varepsilon_{0k}(\mathbf{p})=\frac{2}{\mathscr{V}}\sum_{\mathbf{p}'}U_2(\mathbf{p},\mathbf{p}')f_{0k}(\mathbf{p}'), \tag{2.14}$$
$$\varepsilon_{ik}(\mathbf{p})=\frac{2}{\mathscr{V}}\sum_{\mathbf{p}'}U_3(\mathbf{p},\mathbf{p}')f_{ik}(\mathbf{p}'),$$

and the $\Delta$ to the $g$ as

$$\Delta_{00}(\mathbf{p})=\frac{1}{\mathscr{V}}\sum_{\mathbf{p}'}V_0(\mathbf{p},\mathbf{p}')g_{00}(\mathbf{p}'),$$
$$\Delta_{i0}(\mathbf{p})=\frac{1}{\mathscr{V}}\sum_{\mathbf{p}'}V_1(\mathbf{p},\mathbf{p}')g_{i0}(\mathbf{p}'),$$
$$\Delta_{0k}(\mathbf{p})=\frac{1}{\mathscr{V}}\sum_{\mathbf{p}'}V_2(\mathbf{p},\mathbf{p}')g_{0k}(\mathbf{p}'), \tag{2.15}$$
$$\Delta_{ik}(\mathbf{p})=\frac{1}{\mathscr{V}}\sum_{\mathbf{p}'}V_3(\mathbf{p},\mathbf{p}')g_{ik}(\mathbf{p}').$$

The symmetry properties of the interaction potentials $V_0,\ldots,V_3$ must correspond to the symmetry properties of the order parameters $\Delta$:

$$V_0(-\mathbf{p},\mathbf{p}')=V_0(\mathbf{p},-\mathbf{p}')=-V_0(\mathbf{p},\mathbf{p}'),$$
$$V_1(-\mathbf{p},\mathbf{p}')=V_1(\mathbf{p},-\mathbf{p}')=+V_1(\mathbf{p},\mathbf{p}'),$$
$$V_2(-\mathbf{p},\mathbf{p}')=V_2(\mathbf{p},-\mathbf{p}')=+V_2(\mathbf{p},\mathbf{p}'), \tag{2.16}$$
$$V_3(-\mathbf{p},\mathbf{p}')=V_3(\mathbf{p},-\mathbf{p}')=-V_3(\mathbf{p},\mathbf{p}').$$

To obtain the self-consistency equation we must express the functions $f$ and $g$ in terms of $\varepsilon$ and $\Delta$. This problem is studied in Secs. 3 and 4.

## 3. TRANSITION TEMPERATURES FOR DIFFERENT SUPERFLUID PHASES

Near the phase transition points the order parameters $\Delta$ are small. Hence, if the self-consistency equation have been found, the equations for determining the temperatures of transition to various superfluid phases can be obtained by linearizing the self-consistency equations in $\Delta$. We will use another approach, however. The expressions for the anomalous distribution functions $g$ can be found in the $\Delta$-linear approximation by expanding the statistical operator $\hat{f}$ in a series in powers of $\Delta$ and keeping only terms that are linear in $\Delta$. To this end we write the operator $\hat{\xi}$ in the form of a sum:

$$\hat{\xi}=\hat{\xi}_0+\hat{\Delta},\quad \hat{\xi}_0=\begin{pmatrix}\xi & 0\\ 0 & -\widetilde{\xi}\end{pmatrix},\quad \hat{\Delta}=\begin{pmatrix}0 & \Delta\\ \Delta^+ & 0\end{pmatrix}. \tag{3.1}$$

Then in the $\Delta$-linear approximation we have

$$\hat{f}=\hat{f}_0+\hat{f}',$$
$$\hat{f}_0=\begin{pmatrix}f_0 & 0\\ 0 & 1-\widetilde{f}_0\end{pmatrix},\quad \hat{f}'=-\begin{pmatrix}0 & (1-f_0)\int_0^\beta e^{-\lambda\xi}\Delta e^{-\lambda\widetilde{\xi}}(1-\widetilde{f}_0)d\lambda\\ \widetilde{f}_0\int_0^\beta e^{\lambda\widetilde{\xi}}\Delta^+ e^{\lambda\xi}f_0 d\lambda & 0\end{pmatrix}. \tag{3.2}$$

Here

$$f_0=\frac{1}{\exp(\beta\xi)+1}. \tag{3.3}$$

Comparing (2.1) with (3.2), we arrive at the following expression for the anomalous distribution function $g$:

$$g=-(1-f_0)\int_0^\beta e^{-\lambda\xi}\Delta e^{-\lambda\widetilde{\xi}}(1-\widetilde{f}_0)d\lambda. \tag{3.4}$$

Here $g$ is a matrix in the momentum, spin, and isospin variables, with the dependence on momenta determined by formula (2.7). Now we assume that the phase transition to the

superfluid state begins from a normal state, in which there is no spin order ($f_{i0}=f_{ik}=0$; for symmetric nuclear matter the distribution function $f_{0k}$ is zero, too). Assuming that the amplitudes $U_0(\mathbf{p},\mathbf{p}')$ and $U_2(\mathbf{p},\mathbf{p}')$ are quadratic in momenta,

$$U_0(\mathbf{p},\mathbf{p}')=d_0+e_0(\mathbf{p}^2+\mathbf{p}'^2)+h_0\mathbf{p}\cdot\mathbf{p}',$$

$$U_2(\mathbf{p},\mathbf{p}')=d_2+e_2(\mathbf{p}^2+\mathbf{p}'^2)+h_2\mathbf{p}\cdot\mathbf{p}', \quad p,p'\leqslant p_0,$$

with $p_0$ the cutoff momentum, we can write

$$\xi=(\xi_0\tau_0+\xi_3\tau_3)\sigma_0, \tag{3.5}$$

where

$$\xi_0=\frac{\mathbf{p}^2}{2m_0}+\frac{Y_{4a}^*+Y_{4b}^*}{2Y_0}, \quad \xi_3=\frac{\mathbf{p}^2}{2m_3}+\frac{Y_{4a}^*-Y_{4b}^*}{2Y_0},$$

$$\frac{1}{m_0}=\frac{1}{m}+(\rho_a+\rho_b)e_0, \quad \frac{1}{m_3}=(\rho_a-\rho_b)e_2,$$

with the quantities $Y_{4a}^*$ and $Y_{4b}^*$, renormalized by the Fermi-liquid interaction, determined from the normalization conditions

$$\frac{4}{\mathscr{V}}\sum_{\mathbf{p}} f_{00}(\mathbf{p})=\rho_a+\rho_b, \quad \frac{4}{\mathscr{V}}\sum_{\mathbf{p}} f_{03}(\mathbf{p})=\rho_a-\rho_b.$$

In the above formulas $\sigma_0$ and $\tau_0$ are the identity matrices in the spin and isospin spaces, and $\rho_a$ and $\rho_b$ are the densities of fermions of species $a$ and $b$ (for symmetric nuclear matter, $1/m_3=0$ and $\xi_3=0$). Thus, $\xi$ is a diagonal matrix. Integrating in (3.4), we arrive at the following expression for the matrix element $g_{uv}(\mathbf{p})$ $(u\equiv(\alpha,a))$:

$$g_{uv}(\mathbf{p})=-\frac{\Delta_{uv}(\mathbf{p})}{\xi_u(\mathbf{p})+\xi_v(-\mathbf{p})}[1-f_u^0(\mathbf{p})-f_v^0(-\mathbf{p})]. \tag{3.6}$$

Now we use the expansions (2.7) and (2.13) for $g$ and $\Delta$. Since these expansions are done in linearly independent symmetric and skew-symmetric Pauli matrices, in (3.6) we can replace $g_{uv}$ and $\Delta_{uv}$ by the matrix elements belonging to the corresponding terms. For instance, in the case of singlet–singlet pairing of nucleons we have

$$g_{00}(\mathbf{p})(\tau_2)_{ab}=-\frac{\Delta_{00}(\mathbf{p})}{\xi_a(\mathbf{p})+\xi_b(-\mathbf{p})} \times[1-f_a(\mathbf{p})-f_b(-\mathbf{p})](\tau_2)_{ab}, \quad a=1,2, \tag{3.7}$$

where

$$f_a=\frac{1}{\exp(Y_0\xi_a)+1}, \quad \xi_a=(\xi_0+\xi_3,\xi_0-\xi_3).$$

In (3.7) we do not write the dependence on spin indices explicitly, since this dependence is the same for the left- and right-hand sides of the equations. Assuming $a=1$ and $b=2$, we get

$$g_{00}=-\frac{\Delta_{00}}{4\xi_0}\left[\tanh\frac{Y_0(\xi_0+\xi_3)}{2}+\tanh\frac{Y_0(\xi_0-\xi_3)}{2}\right].$$

Thus, with allowance for (2.15), where we have replaced summation by integration, the equation for determining the temperature of a transition to the superfluid phase with singlet–singlet nucleon pairing ($S=0$ and $T=0$) assumes the following form:

$$\Delta_{00}(\mathbf{p}) =-\frac{1}{4}\int\frac{d^3p'}{(2\pi\hbar)^3}V_0(\mathbf{p},\mathbf{p}')\frac{\Delta_{00}(\mathbf{p}')}{\xi_0(\mathbf{p}')} \times\left[\tanh\frac{Y_0(\xi_0(\mathbf{p}')+\xi_3(\mathbf{p}'))}{2} +\tanh\frac{Y_0(\xi_0(\mathbf{p}')-\xi_3(\mathbf{p}'))}{2}\right]. \tag{3.8}$$

Similarly, in the case of singlet–triplet pairing Eq. (3.6) yields

$$g_{0k}(\mathbf{p})(\tau_k\tau_2)_{ab}=-\frac{\Delta_{0k}(\mathbf{p})}{\xi_a(\mathbf{p})+\xi_b(-\mathbf{p})}[1-f_a(\mathbf{p}) -f_b(-\mathbf{p})](\tau_k\tau_2)_{ab}. \tag{3.9}$$

Setting $a=b=1$ in (3.9), we get

$$ig_{01}+g_{02}=-\frac{1}{2}\frac{i\Delta_{01}+\Delta_{02}}{\xi_0+\xi_3}\tanh\frac{Y_0(\xi_0+\xi_3)}{2}.$$

At $a=b=2$ we have

$$-ig_{01}+g_{02}=-\frac{1}{2}\frac{-i\Delta_{01}+\Delta_{02}}{\xi_0-\xi_3}\tanh\frac{Y_0(\xi_0-\xi_3)}{2}.$$

Finally, for $a=1$ and $b=2$, Eq. (3.9) yields

$$g_{03}=-\frac{\Delta_{03}}{4\xi_0}\left[\tanh\frac{Y_0(\xi_0+\xi_3)}{2}+\tanh\frac{Y_0(\xi_0-\xi_3)}{2}\right].$$

Thus, to determine the temperature of a transition to the superfluid phase with singlet–triplet nucleon pairing ($S=0$ and $T=1$) we have the following equations:

$$i\Delta_{01}(\mathbf{p})+\Delta_{02}(\mathbf{p}) =-\frac{1}{2}\int\frac{d^3p'}{(2\pi\hbar)^3}V_2(\mathbf{p},\mathbf{p}') \times\frac{i\Delta_{01}(\mathbf{p}')+\Delta_{02}(\mathbf{p}')}{\xi_0(\mathbf{p}')+\xi_3(\mathbf{p}')}\tanh\frac{Y_0(\xi_0(\mathbf{p}')+\xi_3(\mathbf{p}'))}{2}, \tag{3.10}$$

$$-i\Delta_{01}(\mathbf{p})+\Delta_{02}(\mathbf{p}) =-\frac{1}{2}\int\frac{d^3p'}{(2\pi\hbar)^3}V_2(\mathbf{p},\mathbf{p}') \times\frac{-i\Delta_{01}(\mathbf{p}')+\Delta_{02}(\mathbf{p}')}{\xi_0(\mathbf{p}')-\xi_3(\mathbf{p}')}\tanh\frac{Y_0(\xi_0(\mathbf{p}')-\xi_3(\mathbf{p}'))}{2}, \tag{3.11}$$

$$\Delta_{03}(\mathbf{p}) =-\frac{1}{4}\int\frac{d^3p'}{(2\pi\hbar)^3}V_2(\mathbf{p},\mathbf{p}')\frac{\Delta_{03}(\mathbf{p}')}{\xi_0(\mathbf{p}')} \times\left[\tanh\frac{Y_0(\xi_0(\mathbf{p}')+\xi_3(\mathbf{p}'))}{2}\right.$$

$$+\tanh\frac{Y_0(\xi_0(\mathbf{p}')-\xi_3(\mathbf{p}'))}{2}\Bigg], \tag{3.12}$$

which describe the pairing of, respectively, two neutrons (isospin projection $T_z=-1$), two protons ($T_z=+1$), and a proton and neutron ($T_z=0$).

The equation for determining the transition temperature corresponding to other types of pairing can be found by reasoning along similar lines. For instance, for triplet–singlet nucleon pairing ($S=1$ and $T=1$) we have

$$\begin{aligned}\Delta_{k0}(\mathbf{p})&=-\frac{1}{4}\int\frac{d^3p'}{(2\pi\hbar)^3}V_1(\mathbf{p},\mathbf{p}')\frac{\Delta_{k0}(\mathbf{p}')}{\xi_0(\mathbf{p}')}\\&\times\Bigg[\tanh\frac{Y_0(\xi_0(\mathbf{p}')+\xi_3(\mathbf{p}'))}{2}\\&+\tanh\frac{Y_0(\xi_0(\mathbf{p}')-\xi_3(\mathbf{p}'))}{2}\Bigg].\end{aligned} \tag{3.13}$$

For triplet–triplet nucleon pairing ($S=1$ and $T=1$) the corresponding equations are

$$\begin{aligned}i\Delta_{k1}(\mathbf{p})+\Delta_{k2}(\mathbf{p})&=-\frac{1}{2}\int\frac{d^3p'}{(2\pi\hbar)^3}V_3(\mathbf{p},\mathbf{p}')\\&\times\frac{i\Delta_{k1}(\mathbf{p}')+\Delta_{k2}(\mathbf{p}')}{\xi_0(\mathbf{p}')+\xi_3(\mathbf{p}')}\\&\times\tanh\frac{Y_0(\xi_0(\mathbf{p}')+\xi_3(\mathbf{p}'))}{2},\end{aligned} \tag{3.14}$$

$$\begin{aligned}-i\Delta_{k1}(\mathbf{p})+\Delta_{k2}(\mathbf{p})&=-\frac{1}{2}\int\frac{d^3p'}{(2\pi\hbar)^3}V_3(\mathbf{p},\mathbf{p}')\\&\times\frac{-i\Delta_{k1}(\mathbf{p}')+\Delta_{k2}(\mathbf{p}')}{\xi_0(\mathbf{p}')-\xi_3(\mathbf{p}')}\tanh\frac{Y_0(\xi_0(\mathbf{p}')-\xi_3(\mathbf{p}'))}{2},\end{aligned} \tag{3.15}$$

$$\begin{aligned}\Delta_{k3}(\mathbf{p})&=-\frac{1}{4}\int\frac{d^3p'}{(2\pi\hbar)^3}V_3(\mathbf{p},\mathbf{p}')\frac{\Delta_{k3}(\mathbf{p}')}{\xi_0(\mathbf{p}')}\\&\times\Bigg[\tanh\frac{Y_0(\xi_0(\mathbf{p}')+\xi_3(\mathbf{p}'))}{2}\\&+\tanh\frac{Y_0(\xi_0(\mathbf{p}')-\xi_3(\mathbf{p}'))}{2}\Bigg].\end{aligned} \tag{3.16}$$

Here Eq. (3.14) determines the temperature of a transition to the phase with triplet pairing in the spin space of two neutrons ($T_z=-1$), Eq. (3.15) determines the temperature of a transition to the phase with triplet pairing of two protons ($T_z=1$), and Eq. (3.16) determines that with triplet pairing of a proton and neutron ($T_z=0$).

Note that Eqs. (3.8) and (3.10)–(3.16) are valid for the general case of asymmetric nuclear matter, where the proton and neutron densities are different.

Below we study these equations for the case of symmetric nuclear matter (equal neutron and proton densities, or $Y^*_{4a}=Y^*_{4b}$ and $\xi_3=0$). Then Eqs. (3.8) and (3.10)–(3.16) become simpler:

$$\Delta_{00}(\mathbf{p})=-\frac{1}{2}\int\frac{d^3p'}{(2\pi\hbar)^3}V_0(\mathbf{p},\mathbf{p}')\frac{\Delta_{00}(\mathbf{p}')}{\xi_0(\mathbf{p}')}\tanh\frac{Y_0\xi_0(\mathbf{p}')}{2}, \tag{3.17}$$

$$\Delta_{30}(\mathbf{p})=-\frac{1}{2}\int\frac{d^3p'}{(2\pi\hbar)^3}V_1(\mathbf{p},\mathbf{p}')\frac{\Delta_{30}(\mathbf{p}')}{\xi_0(\mathbf{p}')}\tanh\frac{Y_0\xi_0(\mathbf{p}')}{2}, \tag{3.18}$$

$$\Delta_{03}(\mathbf{p})=-\frac{1}{2}\int\frac{d^3p'}{(2\pi\hbar)^3}V_2(\mathbf{p},\mathbf{p}')\frac{\Delta_{03}(\mathbf{p}')}{\xi_0(\mathbf{p}')}\tanh\frac{Y_0\xi_0(\mathbf{p}')}{2}, \tag{3.19}$$

$$\Delta_{33}(\mathbf{p})=-\frac{1}{2}\int\frac{d^3p'}{(2\pi\hbar)^3}V_3(\mathbf{p},\mathbf{p}')\frac{\Delta_{33}(\mathbf{p}')}{\xi_0(\mathbf{p}')}\tanh\frac{Y_0\xi_0(\mathbf{p}')}{2}. \tag{3.20}$$

Here we have listed only the equations that refer to different transition temperatures. For instance, Eqs. (3.10)–(3.12), which in the case of asymmetric nuclear matter determine the temperature of a transition to states with different isospin projections ($T_z=0,\pm1$) for singlet–triplet nucleon pairing, are replaced by a single equation (Eq. (3.19)) in the case of symmetric nuclear matter. Hence, actually, in the case of symmetric nuclear matter we are dealing with transitions into superfluid phases with singlet–singlet, triplet–singlet, singlet–triplet, and triplet–triplet nucleon pairing, and each has its own transition temperature, independent of the spin and isospin projections. In what follows we denote these transition temperatures by $T_c(00)$ (singlet–singlet pairing), $T_c(30)$ (triplet–singlet pairing), $T_c(03)$ (singlet–triplet pairing), and $T_c(33)$ (triplet–triplet pairing).

Equations (3.17)–(3.20) contain the anomalous Landau interaction amplitudes $V_0(\mathbf{p},\mathbf{p}'),\ldots,V_3(\mathbf{p},\mathbf{p}')$, which in the low−momentum range can be ordered with respect to the arguments $\mathbf{p}$ and $\mathbf{p}'$. Allowing for the symmetry properties (2.16), we get

$$\begin{aligned}&V_0(\mathbf{p},\mathbf{p}')=c_0\mathbf{p}\mathbf{p}', \quad V_1(\mathbf{p},\mathbf{p}')=a_1+b_1(\mathbf{p}^2+\mathbf{p}'^2),\\&V_2(\mathbf{p},\mathbf{p}')=a_2+b_2(\mathbf{p}^2+\mathbf{p}'^2), \quad V_3(\mathbf{p},\mathbf{p}')=c_3\mathbf{p}\mathbf{p}'.\end{aligned} \tag{3.21}$$

Here the expansion coefficients $a$, $b$, and $c$ characterizing the anomalous interaction amplitudes are independent parameters of the theory and, generally speaking, must be determined from comparison with the experimental data.

Equations (3.17)–(3.20) make it possible to determine the temperatures of transitions to various superfluid phases but generally require doing numerical calculations. Analytic expressions for the transition temperature can be obtained if $\varepsilon_F/2T_c\gg1$, where $\varepsilon_F$ is the Fermi energy.

We start with singlet–singlet pairing. This type of pairing is determined by the potential $V_0(\mathbf{p},\mathbf{p}')=c_0\mathbf{p}\mathbf{p}'$ ($p,p'\leqslant p_0$, with $p_0$ the cutoff momentum). This, together with (3.17), implies that the gap $\Delta_{00}(\mathbf{p})$ has the structure

$$\Delta_{00}(\mathbf{p})=\Delta_0\mathbf{p}\mathbf{n}, \quad p\leqslant p_0, \tag{3.22}$$

where **n** is an arbitrary real-valued unit vector. Integrating over the angles in (3.17) and allowing for (3.22), we arrive at the following equation for determining the transition temperature:

$$1+\frac{c_0}{12\pi^2\hbar^3}\int_{p\leqslant p_0} dpp^4\frac{1}{\xi_0}\tanh\frac{Y_0\xi_0}{2}=0. \tag{3.23}$$

If we calculate the integral in (3.23) in the logarithmic approximation ($\varepsilon_F/2T_c\gg 1$), we obtain the following expression for the temperature $T_c(00)$ of a transition to the superfluid phase with singlet–singlet nucleon pairing:

$$T_c(00)=\frac{8\gamma}{\pi}\varepsilon_f\sqrt{\frac{y_0-1}{y_0+1}}\exp\left(\frac{6}{c_0p_f^2\nu}+\psi_4\right), \tag{3.24}$$

where

$$y_0=\frac{p_0}{p_f},\quad \psi_4=\frac{y_0^3}{3}+y_0-\frac{8}{3},\quad \gamma=e^C,$$

$\nu=m^*p_F/\pi^2\hbar^3$ is the density of states at the Fermi surface, and $C$ is Euler's constant.

For triplet–singlet nucleon pairing the potential is $V_1(p,p')=a_1+b_1(\mathbf{p}^2+\mathbf{p}'^2)$ $(p,p'\leqslant p_0)$, with the results that, with allowance for (3.18), the formula for the gap $\Delta_{30}(\mathbf{p})$ becomes

$$\Delta_{30}(\mathbf{p})=\Delta_1+\Delta_2\mathbf{p}^2,\quad p\leqslant p_0, \tag{3.25}$$

where $\Delta_1$ and $\Delta_2$ are momentum-independent coefficients. Plugging (3.25) into (3.18) and nullifying the determinant of the resulting system of linear equations for $\Delta_1$ and $\Delta_2$, we arrive at the following equation for the transition temperature:

$$1+a_1I_2+2b_1I_4+b_1^2(I_4^2-I_2I_6)=0, \tag{3.26}$$

where

$$I_{2l}=\frac{1}{4\pi^2\hbar^3}\int_{p\leqslant p_0} dpp^{2l}\frac{1}{\xi_0}\tanh\frac{Y_0\xi_0}{2},\quad l=1,2,3.$$

This yields the following expression (in the logarithmic approximation) for the temperature $T_c(30)$ for a transition to the phase with triplet–singlet nucleon pairing:

$$T_c(30)=\frac{8\gamma}{\pi}\varepsilon_f\sqrt{\frac{y_0-1}{y_0+1}}\times\exp\frac{1+\frac{a_1\nu}{2}\psi_2+b_1\nu p_f^2\psi_4+\left(\frac{b_1\nu p_f^2}{2}\right)^2(\psi_4^2-\psi_2\psi_6)}{\frac{a_1\nu}{2}+b_1\nu p_f^2+\left(\frac{b_1\nu p_f^2}{2}\right)^2(2\psi_4-\psi_2-\psi_6)}, \tag{3.27}$$

where

$$\psi_2=y_0-2,\quad \psi_6=\frac{y_0^5}{5}+\frac{y_0^3}{3}+y_0-\frac{46}{15}.$$

Similar reasoning can be carried for the cases of singlet–triplet and triplet–triplet nucleon pairings. For singlet–triplet pairing $V_2(\mathbf{p},\mathbf{p}')=a_2+b_2(\mathbf{p}^2+\mathbf{p}'^2)$, and the transition temperature $T_c(03)$ is determined by formula (3.27) in which $a_1$ must be replaced by $a_2$ and $b_1$ by $b_2$. In the case of triplet–triplet pairing, $V_3(\mathbf{p},\mathbf{p}')=c_2\mathbf{p}\cdot\mathbf{p}'$, and the transition temperature is given by (3.24) with $c_0$ replaced by $c_3$.

These formulas are valid in the limit $2T_c/\varepsilon_F\ll 1$. If this inequality is invalid, Eqs. (3.17)–(3.20) must be solved numerically if we wish to find the transition temperature. To do numerical estimates, we express the Landau amplitudes in terms of the nucleon–nucleon interaction amplitude by using the leading approximation in the interaction. To this end we write the nucleon interaction Hamiltonian in the form

$$V_{\rm int}=\frac{1}{2\mathscr{V}}\sum_{\kappa_1\kappa_2\kappa_3\kappa_4} v(\kappa_1,\kappa_2;\kappa_3,\kappa_4)a^+_{\kappa_1}a^+_{\kappa_2}a_{\kappa_3}a_{\kappa_4}, \tag{3.28}$$

where the interaction amplitude $v(\kappa_1,\kappa_2;\kappa_3,\kappa_4)$ invariant under rotations in the spin and isospin spaces has the following form:

$$\begin{aligned}v(\kappa_1,\kappa_2;\kappa_3,\kappa_4)&\equiv v(\mathbf{p}_1,\mathbf{p}_2;\mathbf{p}_3,\mathbf{p}_4)^{a_1a_2a_3a_4}_{\alpha_1\alpha_2\alpha_3\alpha_4}\\&=[v_0(\mathbf{p},\mathbf{q})\delta_{\alpha_1\alpha_3}\delta_{\alpha_2\alpha_4}\delta_{a_1a_3}\delta_{a_2a_4}\\&\quad+v_1(\mathbf{p},\mathbf{q})\sigma_{\alpha_1\alpha_3}\sigma_{\alpha_2\alpha_4}\delta_{a_1a_3}\delta_{a_2a_4}\\&\quad+v_2(\mathbf{p},\mathbf{q})\delta_{\alpha_1\alpha_3}\delta_{\alpha_2\alpha_4}\tau_{a_1a_3}\tau_{a_2a_4}\\&\quad+v_3(\mathbf{p},\mathbf{q})\sigma_{\alpha_1\alpha_3}\sigma_{\alpha_2\alpha_4}\tau_{a_1a_3}\tau_{a_2a_4}]\\&\quad\times\delta_{\mathbf{p}_1+\mathbf{p}_2,\mathbf{p}_3+\mathbf{p}_4},\end{aligned} \tag{3.29}$$

with $\mathbf{p}=(\mathbf{p}_1-\mathbf{p}_2)/2$ and $\mathbf{q}=(\mathbf{p}_3-\mathbf{p}_4)/2$. Here we have assumed, as is common in nuclear physics, that the amplitudes $v_0$, $v_1$, $v_2$, and $v_3$ are independent of the total momentum of the colliding particles.

To find the energy functional corresponding to the interaction Hamiltonian (3.28), we must average the operator (3.28) over the state of an nonideal gas of particles.[33] For this state the Wick rules for averaging operators are valid. The interaction energy functional $E_{\rm int}(f)$ can be obtained from the operator $V_{\rm int}$ by linking the fermion operators into the following scheme:

$$E_{\rm int}(f)=\frac{1}{2\mathscr{V}}\sum_{\kappa_1\kappa_2\kappa_3\kappa_4} v(\kappa_1,\kappa_2;\kappa_3,\kappa_4)\{\underbrace{a^+_{\kappa_1}a^+_{\kappa_2}a_{\kappa_3}a_{\kappa_4}}+\underbrace{a^+_{\kappa_1}a^+_{\kappa_2}a_{\kappa_3}a_{\kappa_4}}\}.$$

Since $\underbrace{a^+_{\kappa_1}a_{\kappa_2}}=f_{\kappa_2\kappa_1}$, we obtain (at $k=(\mathbf{p}-\mathbf{q})/2$)

$$\begin{aligned}U_0(\mathbf{p},\mathbf{q})&=4v_0(\mathbf{k},\mathbf{k})-v_0(-\mathbf{k},\mathbf{k})-3v_1(-\mathbf{k},\mathbf{k})\\&\quad-3v_2(-\mathbf{k},\mathbf{k})-9v_3(-\mathbf{k},\mathbf{k}),\\U_1(\mathbf{p},\mathbf{q})&=4v_1(\mathbf{k},\mathbf{k})-v_0(-\mathbf{k},\mathbf{k})+v_1(-\mathbf{k},\mathbf{k})\\&\quad-3v_2(-\mathbf{k},\mathbf{k})+3v_3(-\mathbf{k},\mathbf{k}),\\U_2(\mathbf{p},\mathbf{q})&=4v_2(\mathbf{k},\mathbf{k})-v_0(-\mathbf{k},\mathbf{k})-3v_1(-\mathbf{k},\mathbf{k})\\&\quad+v_2(-\mathbf{k},\mathbf{k})+3v_3(-\mathbf{k},\mathbf{k}),\\U_3(\mathbf{p},\mathbf{q})&=4v_3(\mathbf{k},\mathbf{k})-v_0(-\mathbf{k},\mathbf{k})+v_1(-\mathbf{k},\mathbf{k})\end{aligned} \tag{3.30}$$

$$+v_2(-\mathbf{k},\mathbf{k})-v_3(-\mathbf{k},\mathbf{k}).$$

The interaction energy functional $E_{\rm int}(g)$ can be obtained from the operator $V$ by linking the fermion operators into the following scheme:

$$E_{\rm int}(g)=\frac{1}{2\mathscr{V}}\sum_{\kappa_1\kappa_2\kappa_3\kappa_4} v(\kappa_1,\kappa_2;\kappa_3,\kappa_4)\underbrace{a^+_{k_1}}\underbrace{a^+_{\kappa_2}}\underbrace{a_{\kappa_3}}a_{\kappa_4}.$$

Allowing for the fact that $\underbrace{a_{\kappa_2}a_{\kappa_1}}=g_{\kappa_1\kappa_2}$, we obtain

$$\begin{aligned}
V_0(\mathbf{p},\mathbf{q})&=v_0(\mathbf{p},\mathbf{q})-3v_1(\mathbf{p},\mathbf{q})-3v_2(\mathbf{p},\mathbf{q})+9v_3(\mathbf{p},\mathbf{q}),\\
V_1(\mathbf{p},\mathbf{q})&=v_0(\mathbf{p},\mathbf{q})+v_1(\mathbf{p},\mathbf{q})-3v_2(\mathbf{p},\mathbf{q})-3v_3(\mathbf{p},\mathbf{q}),\\
V_2(\mathbf{p},\mathbf{q})&=v_0(\mathbf{p},\mathbf{q})-3v_1(\mathbf{p},\mathbf{q})+v_2(\mathbf{p},\mathbf{q})-3v_3(\mathbf{p},\mathbf{q}),\\
V_3(\mathbf{p},\mathbf{q})&=v_0(\mathbf{p},\mathbf{q})+v_1(\mathbf{p},\mathbf{q})+v_2(\mathbf{p},\mathbf{q})+v_3(\mathbf{p},\mathbf{q}).
\end{aligned}\tag{3.31}$$

Below we use the Skyrme interaction as the nucleon–nucleon potential.[24] For the Skyrme interaction the functions $v_0,\ldots,v_3$ in (3.29) are

$$\begin{aligned}
v_0(\mathbf{p},\mathbf{p}')&=\frac{3}{8}t_0+\frac{1}{16}t_3\rho^\alpha+\frac{3}{16\hbar^2}t_1(\mathbf{p}^2+\mathbf{p}'^2)\\
&\quad+\frac{1}{8\hbar^2}t_2(5+4x_2)\mathbf{p}\mathbf{p}',\\
v_1(\mathbf{p},\mathbf{p}')&=-\frac{1}{8}t_0(1-2x_0)-\frac{1}{48}t_3\rho^\alpha(1-2x_3)\\
&\quad-\frac{1}{16\hbar^2}t_1(1-2x_1)(\mathbf{p}^2+\mathbf{p}'^2)+\frac{1}{8\hbar^2}t_2\\
&\quad\times(1+2x_2)\mathbf{p}\cdot\mathbf{p}',\\
v_2(\mathbf{p},\mathbf{p}')&=-\frac{1}{8}t_0(1+2x_0)-\frac{1}{48}t_3\rho^\alpha(1+2x_3)\\
&\quad-\frac{1}{16\hbar^2}t_1(1+2x_1)(\mathbf{p}^2+\mathbf{p}'^2)+\frac{1}{8\hbar^2}t_2\\
&\quad\times(1+2x_2)\mathbf{p}\cdot\mathbf{p}',\\
v_3(\mathbf{p},\mathbf{p}')&=-\frac{1}{8}t_0-\frac{1}{48}t_3\rho^\alpha-\frac{1}{16\hbar^2}t_1(\mathbf{p}^2+\mathbf{p}'^2)\\
&\quad+\frac{1}{8\hbar^2}t_2\mathbf{p}\cdot\mathbf{p}'.
\end{aligned}\tag{3.32}$$

Here $\rho$ is the density of symmetric nuclear matter, and $t_0,\ldots,t_3$, $x_0,\ldots,x_3$, and $\alpha$ are phenomenological parameters. In the above formulas we assume that the momenta $p$ and $p'$ do not exceed a cutoff momentum $p_0$, which in turn is somewhat larger than the Fermi momentum $p_F$, or $p_0\gtrsim p_F$ (for numerical calculations we assume that $p_0=1.1p_F$; see Ref. 14). The local Skyrme interaction describes the effective interaction of a pair of nucleons in the presence of a nucleon medium, which is reflected by the term dependent on density. This term is the result of allowing for three-particle interactions. For symmetric nuclear matter with Skyrme interaction the effective nucleon mass is given by the following expression:[24]

$$\frac{\hbar^2}{2m^*}=\frac{\hbar^2}{2m}+\frac{1}{16}[3t_1+t_2(5+4x_2)]\rho,\tag{3.33}$$

where $m$ is the ''bare'' nucleon mass (cf. Eq. (3.5)).

There are sets of parameters $t$, $x$, and $\alpha$ that are used to describe the physical characteristics of the ground states of nuclei (the mass and radius of a nucleus, the binding energy, the moment of inertia, etc.) and the data on giant resonances and fission barriers for heavy nuclei (see Ref. 23, the review in Ref. 24, and the literature cited therein). We use the most recent versions of the Skyrme interaction taken from the literature (the Ska, SkM, SkM*, and RATP potentials; see Table I), in which, in contrast to the earlier work with exponent $\alpha=1$, in which the term dependent on the density contains a variable power of the density, $1/6\leqslant\alpha\leqslant 1/3$.

TABLE I. Parameters of the different versions of the Skyrme interaction used in the present paper (according to the review article in Ref. 24).

| Potential | Ska | SkM | SkM* | RATP |
|---|---|---|---|---|
| $t_0$, MeV·fm$^3$ | −1602.78 | −2645.00 | −2645.00 | −2160.00 |
| $t_1$, MeV·fm$^5$ | 570.88 | 385.00 | 410.00 | 513.00 |
| $t_2$, MeV·fm$^5$ | −67.70 | −120.00 | −135.00 | 121.00 |
| $t_3$, MeV·fm$^{3+3\alpha}$ | 8000.00 | 15 595.00 | 15 595.00 | 11 600.00 |
| $x_0$ | −0.02 | 0.09 | 0.09 | 0.418 |
| $x_1$ | 0.00 | 0.00 | 0.00 | −0.36 |
| $x_2$ | 0.00 | 0.00 | 0.00 | −2.29 |
| $x_3$ | −0.286 | 0.00 | 0.00 | 0.586 |
| $\alpha$ | 1/3 | 1/6 | 1/6 | 1/5 |

There are sets of parameters $a$, $b$, and $c$ that characterize the anomalous Fermi-liquid interaction amplitudes $V_0,\ldots,V_3$ in the case of a Skyrme interaction:

$$\begin{aligned}
a_{1,2}&=t_0(1\pm x_0)+\frac{1}{6}t_3\rho^\alpha(1\pm x_3),\\
b_{1,2}&=\frac{1}{2\hbar^2}t_1(1\pm x_1),\quad c_{0,3}=\frac{1}{\hbar^2}t_2(1\mp x_2).
\end{aligned}\tag{3.34}$$

As Eqs. (3.21) and (3.34) imply, singlet–singlet and triplet–triplet nucleon pairings in the corresponding superfluid phases occur in states with orbital angular momentum $L=1$, while triplet–singlet and singlet–triplet nucleon pairings occur in states with $L=0$.

Note that if the parameter $x_2$ in the effective Skyrme interaction is zero, $V_0$ and $V_3$ are equal and the temperatures of transition to phases with singlet–singlet and triplet–triplet pairing coincide, $T_c(00)=T_c(33)$ (this is true for the Ska, SkM, and SkM* potentials). Figure 1 depicts the results of the numerical solution of Eqs. (3.17)–(3.20) represented in the form of a density vs transition-temperature diagram. Qualitatively the behavior of the transition curves is the same for all four Skyrme potentials, although some numerical differences do indeed exist. Note that at the saturation value of

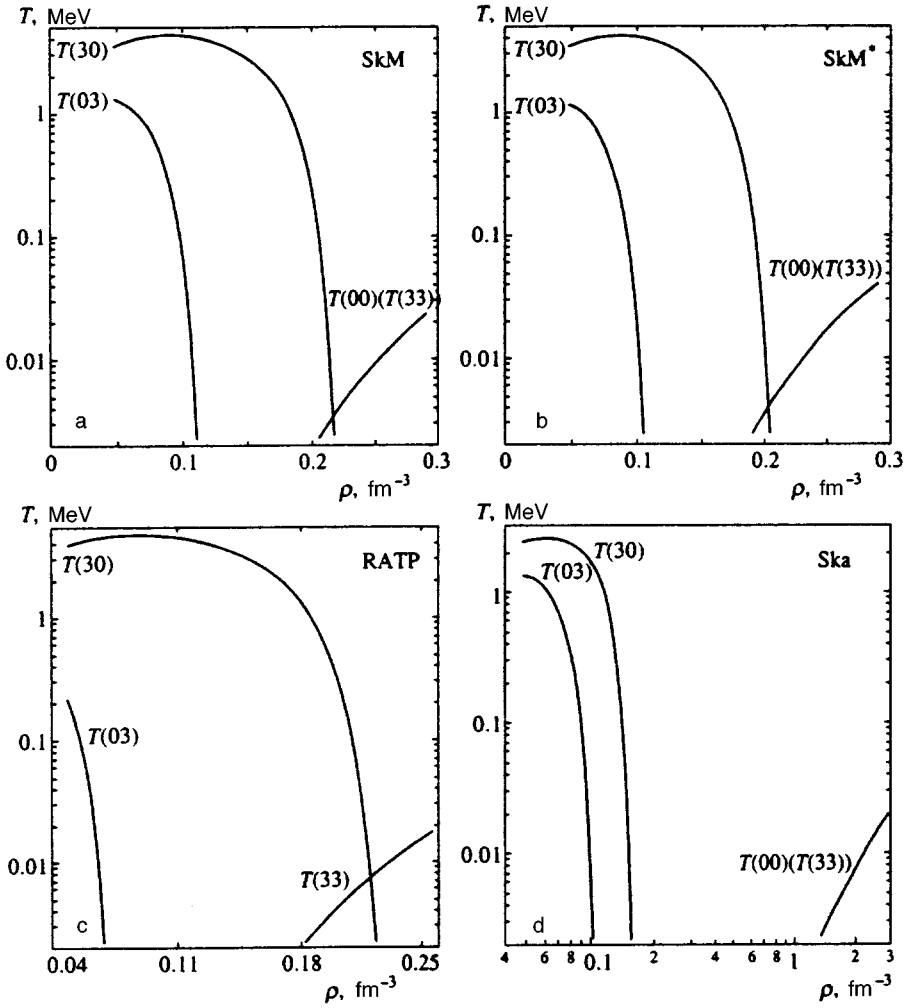



FIG. 1. Transition temperature as a function of density for different pairing types in SkM (a), SkM* (b), RATP (c), and Ska (d) potentials.

nuclear matter density,[1)] or $\rho_\infty$, the temperature of transition to the state with triplet–singlet nucleon pairing (see Table II) is the highest for all types of interaction.

The superfluidity of nuclear matter with the Skyrme interaction was also examined by Su *et al.*[14] They, however, studied nucleon pairing without allowing for the spin and isospin structures of a pair, which means in deriving the energy functional they ignored the spin exchange interaction between nucleons (all the parameter $x_i$ in the Skyrme potentials are set equal to zero). Our treatment makes it possible to determine the structure of the phase diagrams for different superfluid phases and to establish how the variation of the nuclear matter density determines the spin and isospin state in a which there is nucleon pairing.

TABLE II. Values of the transition temperature $T_c(30)$, the energy gap for unitary ($\Delta^u_{p_F}$) and nonunitary ($\Delta^{nu}_{p_F}$) states, and the magnetization density $m$ (nonunitary states) in the case of triplet–singlet nucleon pairing at a density equal to the saturation density $\rho_\infty$ of nuclear matter.

| Potential | Ska | SkM | SkM* | RATP |
|---|---|---|---|---|
| $T_c(30)$, MeV | 0.0069 | 2.36 | 1.75 | 2.57 |
| $\Delta^u_{p_F}$, MeV | 0.012 | 4.08 | 3.01 | 4.52 |
| $\Delta^{nu}_{p_F}$, MeV | 0.0086 | 2.88 | 2.13 | 3.19 |
| $m$, $\mu_{\text{nucl}}$ fm$^{-3}$ | $1.4\times10^{-7}$ | 0.0013 | 0.0010 | 0.0009 |
| $\rho_\infty$, fm$^{-3}$ | 0.1554 | 0.1603 | 0.1603 | 0.1599 |

Note the special features of the resulting phase curves. Figure 1 shows that for the SkM and SkM* potentials there is a certain critical density, the point at which the curves $T_c(30)$ and $T_c(00)$ (or $T_c(33)$) intersect, such that for $\rho<\rho_c$ the highest transition temperature is that of the transition to the phase with triplet–singlet nucleon pairing, while for $\rho>\rho_c$ the highest transition temperature is that of the transition to the singlet–singlet or triplet– triplet transition (for the SkM and SkM* potentials the two coincide). Thus, for the SkM and SkM* potentials there is a phase transition in density from the triplet–singlet phase to the singlet–singlet or triplet–triplet phase.

A similar conclusion holds for the RATP potential, the only difference being that the phase transition in density from the triplet–singlet phase can only end in the triplet–triplet phase, since for the RATP potential the equation for the temperature of transition to a phase with singlet–singlet nucleon pairing has no solutions.

For the Ska potential the curves $T_c(30)$ and $T_c(00)$ (or $T_c(33)$) do not intersect,[2)] so that a phase transition from the triplet–singlet phase to the singlet–singlet or triplet–triplet phase proceeds through the normal phase. Such a transition, however, is purely formal since this section on the phase

diagram corresponds to high nuclear matter densities ($\approx 1.3$ fm$^{-3}$).

To compare our results with those of by Su *et al.*,[14] we note that in their calculations they used the old Skyrme potentials SkI–SkVI (see Ref. 23), known from the beginning of the 1970s, and of the new potentials only the SkM* potential was selected. For the SkM* potential, Su *et al.*[14] concluded that the equation for the transition temperature of the second-order phase transition into the superfluid state at a density equal to the saturation density has no solutions. Our calculations with the SkM* potential show that at a density equal to the saturation density it is not the normal phase that is realized but the superfluid phase with triplet–singlet nucleon pairing and the corresponding transition temperature $T_c(30)$ of 1.75 MeV. The difference can be explained by the fact that, as noted earlier, Su *et al.*[14] did not take into account the terms in the Skyrme potential that describe spin exchange interaction between nucleons. Allowing for these terms, as we have seen, is important and leads to qualitatively different results.

#### 4. ENERGY GAP

The analysis of the phase curves done in Sec. 3 shows that at densities close to the saturation density a superfluid phase appears with triplet–singlet nucleon pairing ($S=1$ and $T=0$). In this section we determine the energy gap and the order parameters for triplet–singlet pairing. To this end we turn to the self-consistency equation (2.4). Using the diagonalization procedure described in Ref. 25, we can reduce this equation to a form that contains operators that act only in the momentum, spin, and isospin spaces. As a result, for the equilibrium distribution functions $f$ and $g$ we obtain

$$f=Kn+X(1-\widetilde{n})X^{+}K, g=\widetilde{KnX}+K(1-n)X, \tag{4.1}$$

where

$$n=\{\exp[Y_0(\xi-X\Delta^{+})]+1\}^{-1}, \quad \xi=\varepsilon+\frac{Y_4}{Y_0},$$

$$K=(1+XX^{+})^{-1}, \tag{4.2}$$

and the matrix $X$ satisfies the equation

$$\xi X+X\widetilde{\xi}+\Delta-X\Delta^{+}X=0, \quad \widetilde{X}=-X. \tag{4.3}$$

For triplet–singlet nucleon pairing the anomalous distribution function $g$ has the form

$$g_{\kappa_1\kappa_2}=g_{k0}(\mathbf{p}_1)\delta_{p_1,-p_2}(\sigma_k\sigma_2)_{\sigma_1,\sigma_2}(\tau_2)_{a_1,a_2}, \tag{4.4}$$

where $g_{k0}(-\mathbf{p})=g_{k0}(\mathbf{p})$. Combining this with (2.3) yields

$$\Delta_{\kappa_1\kappa_2}=\Delta_{k0}(\mathbf{p}_i)\delta_{p_1,-p_2}(\sigma_k\sigma_2)_{\sigma_1,\sigma_2}(\tau_2)_{a_1,a_2},$$

$$\Delta_{k0}(-\mathbf{p})=\Delta_{k0}(\mathbf{p}).$$

To simplify the formulas below we introduce the following notation:

$$\mathbf{g}_p\equiv(g_{10}(\mathbf{p}),g_{20}(\mathbf{p}),g_{30}(\mathbf{p})),$$

$$\mathbf{\Delta}_p\equiv(\Delta_{10}(\mathbf{p}),\Delta_{20}(\mathbf{p}),\Delta_{30}(\mathbf{p})).$$

Solving Eq. (4.3) for the $X$ matrix requires knowing the structure of $\xi$ in the spin and isospin spaces. According to (2.2) and (4.2), this structure is determined by the form of the energy functional $E(f,g)$. For triplet–singlet nucleon pairing we specify the functional $E(f,g)$ by the following formula (compare this with the general expressions (2.8)–(2.11)):

$$E(f,g)=4\sum_{\mathbf{p}}\varepsilon_0(\mathbf{p})f^0_{\mathbf{p}}+\frac{2}{\mathscr{V}}\sum_{\mathbf{pp}'}f^0_{\mathbf{p}}U_0(\mathbf{p},\mathbf{p}')f^0_{\mathbf{p}'}$$

$$+\frac{2}{\mathscr{V}}\sum_{\mathbf{pp}'}\mathbf{g}^{*}_{\mathbf{p}}V_1(\mathbf{p},\mathbf{p}')\mathbf{g}_{\mathbf{p}'}, \quad f^0_{\mathbf{p}}\equiv f_{00}(\mathbf{p}). \tag{4.5}$$

Here, in the view of the weakness of magnetization, we ignored the exchange interaction in the normal part of the Fermi-liquid energy (not to be confused with the microscopic spin exchange interaction between nucleons, taken into account in the Skyrme potentials via the parameters $x_i$). Combining (4.5) with (2.13), (2.14), and (4.2), we see that for symmetric nuclear matter ($Y_{4a}=Y_{4b}=Y_4$), $\xi$ is proportional to the identity matrices in the spin and isospin spaces:

$$\xi_{\kappa_1\kappa_2}=\xi_{p_1}\delta_{p_1,p_2}(\sigma_0)_{\sigma_1,\sigma_2}(\tau_0)_{a_1,a_2},$$

where

$$\xi_{\mathbf{p}}=\varepsilon_0(\mathbf{p})+\frac{Y_4}{Y_0}+\frac{2}{\mathscr{V}}\sum_{\mathbf{p}'}U_0(\mathbf{p},\mathbf{p}')f^0_{\mathbf{p}'}. \tag{4.6}$$

Then, solving Eq. (4.3) and allowing for (4.1), we arrive at the following expression for the anomalous distribution function:[35]

$$\mathbf{g}_p=\frac{1}{2\zeta_pG_p}\left\{\frac{1}{2}(1-n^0_p-n^0_{-p})[(2\zeta^2_p-\mathbf{\Delta}_p\cdot\mathbf{\Delta}^{*}_p)\mathbf{\Delta}_p\right.$$

$$+(\mathbf{\Delta}^2_p)\mathbf{\Delta}^{*}_p]+\frac{n^{\|}_p+n^{\|}_{-p}}{|[\mathbf{\Delta}_p\cdot\mathbf{\Delta}^{*}_p]|}\zeta^2_p\left[\left(2\left(\zeta^2_p-\frac{1}{4}(\xi_p+\xi_{-p})^2\right)\right.\right.$$

$$\left.\left.-\mathbf{\Delta}_p\cdot\mathbf{\Delta}^{*}_p\right)\mathbf{\Delta}_p-(\mathbf{\Delta}_p)^2\mathbf{\Delta}^{*}_p\right]\Bigg\}. \tag{4.7}$$

Here

$$n^0_p=\frac{1}{2}(n^{+}_p+n^{-}_p), \quad n^{\|}_p=\frac{1}{2}(n^{+}_p-n^{-}_p),$$

$$n^{\pm}_p=\left\{\exp\left[Y_0\left(-\zeta_p+\frac{\xi_p-\xi_{-p}}{2}\pm\frac{|[\mathbf{\Delta}_p\cdot\mathbf{\Delta}^{*}_p]|}{2\zeta_p}\right)\right]+1\right\}^{-1},$$

and $\zeta_p$ and $G_p$ are given are defined as follows:

$$\zeta^2_p=\frac{1}{2}\left[\mathbf{\Delta}_p\cdot\mathbf{\Delta}^{*}_p+\frac{1}{4}(\xi_p+\xi_{-p})^2+G_p\right],$$

$$G_p=\sqrt{\left[\mathbf{\Delta}_p\cdot\mathbf{\Delta}^{*}_p+\frac{1}{4}(\xi_p+\xi_{-p})^2\right]^2-|[\mathbf{\Delta}_p\cdot\mathbf{\Delta}^{*}_p]|^2}.$$

The structure of the normal distribution function $f$ in the momentum, spin, and isospin spaces is given by the following expression:

$$f_{\kappa_1\kappa_2}=[f^0_{p_1}\delta_{\sigma_1,\sigma_2}+\mathbf{f}_{p_1}(\boldsymbol{\sigma})_{\sigma_1,\sigma_2}]\delta_{p_1,p_2}\delta_{a_1,a_2},$$

$$\mathbf{f}_p \equiv f_{k0}(\mathbf{p}),$$

where, as (4.1) and the solution of Eq. (4.3) imply, the functions $f_p^0$ and $\mathbf{f}_p$ have the form

$$f_p^0 = \frac{1}{2G_p}\left\{(1+n_p^0-n_{-p}^0)G_p + \frac{\zeta_p(\xi_p+\xi_{-p})}{2}\left[1-n_p^0-n_{-p}^0 + \frac{2[\zeta_p^2-(1/4)(\xi_p+\xi_{-p})^2-\mathbf{\Delta}_p\cdot\mathbf{\Delta}_p^*](n_p^{\parallel}+n_{-p}^{\parallel})}{|[\mathbf{\Delta}_p\cdot\mathbf{\Delta}_p^*]|}\right]\right\}, \tag{4.8}$$

$$\mathbf{f}_p = -\frac{i[\mathbf{\Delta}_p\times\mathbf{\Delta}_p^*]}{2G_p}\left\{\frac{\xi_p+\xi_{-p}}{4\zeta_p} \times\left[1-n_p^0-n_{-p}^0-\frac{2\zeta_p^2(n_p^{\parallel}+n_{-p}^{\parallel})}{|[\mathbf{\Delta}_p\cdot\mathbf{\Delta}_p^*]|}\right]+\frac{n_p^{\parallel}-n_{-p}^{\parallel}}{|[\mathbf{\Delta}_p\cdot\mathbf{\Delta}_p^*]|}\right\}. \tag{4.9}$$

Thus, according to Eqs. (2.15) and (4.6)–(4.9), $\xi_p$ and $\mathbf{\Delta}_p$ can be found by solving the following equations:

$$\xi_p = \varepsilon_0(\mathbf{p}) + \frac{Y_4}{Y_0} + \frac{1}{\mathscr{V}}\cdot \times\sum_{\mathbf{p}'}\frac{U_0(\mathbf{p},\mathbf{p}')}{G_{p'}}\left\{(1+n_{p'}^0-n_{-p'}^0) \times G_{p'}+\frac{\zeta_{p'}(\xi_{p'}+\xi_{-p'})}{2}\left[1-n_{p'}^0-n_{-p'}^0 + \frac{2(\zeta_{p'}^2-(1/4)(\xi_{p'}+\xi_{-p'})^2-\mathbf{\Delta}_{p'}\cdot\mathbf{\Delta}_{p'}^*)(n_{p'}^{\parallel}+n_{-p'}^{\parallel})}{|[\mathbf{\Delta}_{p'}\cdot\mathbf{\Delta}_{p'}^*]|}\right]\right\}, \tag{4.10}$$

$$\mathbf{\Delta}_p = \frac{1}{2\mathscr{V}}\cdot\sum_{\mathbf{p}'}\frac{V_1(\mathbf{p},\mathbf{p}')}{\zeta_{p'}G_{p'}} \times\left\{\frac{1-n_{p'}^0-n_{-p'}^0}{2}[(2\zeta_{p'}^2-\mathbf{\Delta}_{p'}\cdot\mathbf{\Delta}_{p'}^*) \times\mathbf{\Delta}_{p'}+\mathbf{\Delta}_{p'}^2\mathbf{\Delta}_{p'}^*]+\frac{(n_{p'}^{\parallel}+n_{-p'}^{\parallel})\zeta_{p'}^2}{|[\mathbf{\Delta}_{p'}\cdot\mathbf{\Delta}_{p'}^*]|} \times\left[\left(2\left(\zeta_{p'}^2-\frac{(\xi_{p'}+\xi_{-p'})^2}{4}\right)-\mathbf{\Delta}_{p'}\cdot\mathbf{\Delta}_{p'}^*\right)\mathbf{\Delta}_{p'} -\mathbf{\Delta}_{p'}^2\mathbf{\Delta}_{p'}^*\right]\right\}. \tag{4.11}$$

These equations completely solve the general problem of triplet–singlet nucleon pairing in superfluid nuclear matter. Formulas (4.11) do not require a specific form for the interaction amplitude, the requirement being that the symmetry conditions (2.16) are met. To be exact, Eqs. (4.10) and (4.11) for the quantities $\xi_p$ and $\mathbf{\Delta}_p$ should be solved simultaneously. We, however, ignore the effect of a finite gap on $\xi$ in (4.10) and assume, therefore, that $\xi_p=(p^2-p_F^2)/2m^*$.

Let us first examine the states of superfluid nuclear matter for which the product $\Delta\Delta^+$ is proportional to the identity matrix in the spin space (which is equivalent to the condition that $\mathbf{\Delta}_p\times\mathbf{\Delta}_p^*=0$). Such states are called unitary. A characteristic feature of these states is that the projection of the total spin of a nucleon pair on the quantization axis in the spin space is zero. The order parameter for unitary states has the form $\mathbf{\Delta}_p=\mathbf{d}\Delta_p$, where $\mathbf{d}$ is a real-valued unit vector along the quantization axis, and $\Delta_p$, as (4.11) implies, satisfies the equation

$$\Delta_p = -\frac{1}{2\mathscr{V}}\cdot\sum_{\mathbf{p}'}V_1(\mathbf{p},\mathbf{p}')\frac{\Delta_{p'}}{E_{p'}}\tanh\frac{Y_0E_{p'}}{2}, \qquad E_p=\sqrt{\xi_p^2+|\Delta_p|^2}. \tag{4.12}$$

This equation is of the BCS type and has nonzero solutions $\Delta_p\neq 0$ at temperatures $T<T_c(30)$, where $T_c(30)$ is the transition temperature of the triplet–singlet phase transition specified by (3.18). According to (3.21), the potential $V_1(\mathbf{p},\mathbf{p}')$ in Eq. (4.12) has the following form:

$$V_1(\mathbf{p},\mathbf{p}')=a_1+b_1(\mathbf{p}^2+\mathbf{p}'^2), \quad p,p'\leqslant p_0.$$

This leads to the following structure of the energy gap $\Delta_p$:

$$\Delta_p=\Delta_1+\Delta_2p^2, \quad p\leqslant p_0. \tag{4.13}$$

Solving Eq. (4.12) at $T=0$ on the assumption that $\Delta_{p_F}/\varepsilon_F\ll 1$, we obtain the ordinary relationship linking the energy gap $\Delta_{p_F}$ at the Fermi surface with the transition temperature $T_c(30)$:

$$\Delta_{p_F}(T=0)=\frac{\pi}{\gamma}T_c(30),$$

with $T_c(30)$ given by (3.27). In the general case the solution of Eq. (4.12) can be found only numerically. The results of numerical integration for the gap at $T=0$ and $p=p_F$ are depicted in Fig. 2a. Clearly, the SkM, SkM*, and RATP potentials yield values for $\Delta_{p_F}(T=0)$, in contrast to the Ska potential, which yields somewhat lower values. Table II lists the values of the energy gap for the different variants of the Skyrme potential at a density equal to the saturation density $\rho_\infty$.

Now let us examine the states of superfluid nuclear matter for which the product $\Delta\Delta^+$ is not reduced to the identity matrix in the spin space. Such states are called nonunitary. Among these are states whose projection of the total spin of a nucleon pair on the quantization axis is $\pm 1$. For such states the order parameter $\mathbf{\Delta}_p$ has the form

$$\mathbf{\Delta}_p=\mathbf{d}^{(\pm)}\Delta_p, \quad \mathbf{d}^{(\pm)}=\frac{1}{\sqrt{2}}(\mp i\mathbf{d}_1+\mathbf{d}_2), \tag{4.14}$$

where $\mathbf{d}_1$ and $\mathbf{d}_2$ are real-valued orthonormal vectors. The upper sign in (4.14) stands for the spin projection $S_z=+1$ and the lower sign for the spin projection $S_z=-1$.

The dependence of $\Delta_p$ on momentum, as (4.11) implies, is given by the equation

$$\Delta_p = -\frac{1}{4\mathscr{V}}\cdot\sum_{\mathbf{p}'}\frac{V_1(\mathbf{p},\mathbf{p}')}{G_{p'}E_{p'}}\left\{\frac{1}{2}\left[\tanh\frac{Y_0(2E_{p'}^2+|\Delta_{p'}|^2)}{4E_{p'}} +\tanh\frac{Y_0(2E_{p'}^2-|\Delta_{p'}|^2)}{4E_{p'}}\right](2E_{p'}^2-|\Delta_{p'}|^2)\right.$$

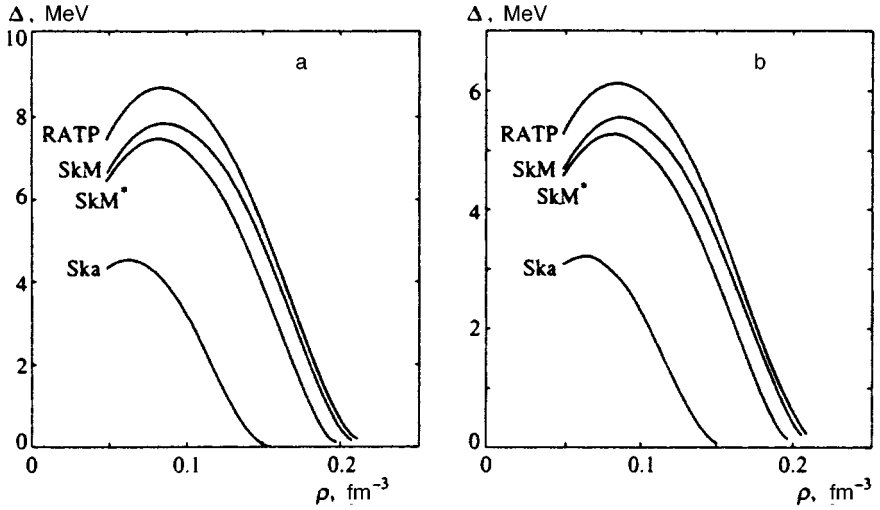



FIG. 2. Energy gap as a function of density for unitary (a) and nonunitary (b) states.

$$+\left[\tanh\frac{Y_0(2E_{p'}^2+|\Delta_{p'}|^2)}{4E_{p'}}-\tanh\frac{Y_0(2E_{p'}^2-|\Delta_{p'}|^2)}{4E_{p'}}\right]$$
$$\times\frac{E_{p'}^2[2(E_{p'}^2-\xi_{p'}^2)-|\Delta_{p'}|^2]}{|\Delta_{p'}|^2}\Bigg\}\Delta_{p'}, \qquad (4.15)$$

where

$$G_p=\xi_p^2\sqrt{1+2\frac{|\Delta_{p'}|^2}{\xi_p^2}}, \quad E_p=\sqrt{\frac{1}{2}(\Delta_p^2+\xi_p^2+G_p)}.$$

As $\Delta_p\to 0$, Eq. (4.15) becomes Eq. (3.18) for determining the temperature of the transition to the superfluid phase with triplet–singlet nucleon pairing. As in the case of unitary states, we should look for the solution of Eq. (4.15) in the form (4.13). Figure 2b depicts the results of numerical calculations of the energy gap for the nonunitary states of nuclear matter at $T=0$ and $p=p_F$. A comparison of Figs. 1a and b shows that for equal densities the gap for nonunitary states is narrower than the gap for unitary states. Table II lists the gap values $\Delta_{p_F}(T=0)$ in the case of nonunitary states for different Skyrme potentials at a density equal to the saturation density. In connection with examining the nonunitary states in the spin space we note that the transition to the superfluid state is accompanied by the appearance of spontaneous magnetization in superfluid nuclear matter, which is caused by the ordering of the spins of the pairing nucleons. The magnetization density **m** is given by the following expression:

$$\mathbf{m}=\mathrm{Tr}_\kappa\boldsymbol{\mu}f, \quad \boldsymbol{\mu}=\frac{1}{2}(1+\tau_3)\mu_p\boldsymbol{\sigma}+\frac{1}{2}(1-\tau_3)\mu_n\boldsymbol{\sigma}, \qquad (4.16)$$

where $\boldsymbol{\mu}$ is the operator of the nucleon's magnetic moment, $\mu_p$ and $\mu_n$ are the magnetic moments of the proton and neutron, respectively, and $f$ is the normal nucleon distribution function. Calculating the trace in (4.16), we arrive at the following expression for the magnetization density $m$ at $T=0$:

$$m=-(\mu_p+\mu_n)\sum_{p\leqslant p_0}\frac{\xi_p|\boldsymbol{\Delta}_p|^2}{E_pG_p}. \qquad (4.17)$$

The results of numerical integration for $m$ are depicted in Fig. 3. Clearly, the magnetization density is fairly low (per nucleon it amounts to roughly $(0.006-0.008)\mu_{\mathrm{nucl}}$ at a density equal to the saturation density for the SkM, SkM*, and RATP potentials, and $9.1\times10^{-7}\mu_{\mathrm{nucl}}$ for the Ska potential, where $\mu_{\mathrm{nucl}}$ is the nuclear magneton). The explanation is that, on the one hand, in this superfluid phase the paired particles are a proton and a neutron whose magnetic moments point in opposite directions, and, on the other, the larger fraction of nuclear matter is in the normal state.

## 5. CONCLUSION

We investigated a semiphenomenological theory of superfluid nuclear matter based on Landau's concept of a Fermi liquid. We classified the various superfluid phases of symmetric nuclear matter and derived the equations for determining the corresponding transition temperatures. In the theory of a superfluid Fermi liquid, the normal and anomalous Landau interaction amplitudes, interpreted phenomeno-

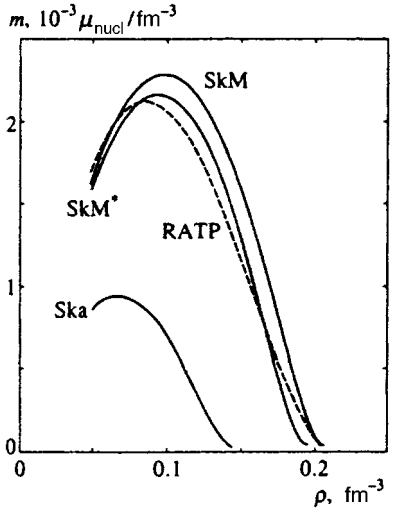



FIG. 3. Magnetization density as a function of density for nonunitary states in the case of triplet–singlet nucleon pairing (the curve corresponding to the RATP potential is depicted by a dashed line).

logically, are considered independent, and the derived equations do not assume any relationship between the normal and anomalous Fermi-liquid amplitudes. However, to be able to estimate the Fermi-liquid amplitudes, we used the leading approximation in the interaction and expressed the normal and anomalous Fermi-liquid amplitudes in terms of the parameters of the effective nucleon–nucleon interaction. For the nucleon–nucleon potential we used the effective Skyrme interaction, with the most recent values of the phenomenological parameters taken from the literature (the SKa, SkM, SkM*, and RATP potentials). Note that we clearly separated the spin and isospin parts of the interaction potential, which makes it possible to easily follow the spin and isospin structures of paired nucleons.

An analysis of the temperatures of transition to various superfluid phases of symmetric nuclear matter shows that there is the possibility of phase transitions in density from the triplet–singlet superfluid phase into the singlet–singlet or triplet–triplet phase (the SkM and SkM* potentials) or from the triplet–singlet phase into the triplet–triplet phase (the RATP potential). For the Ska potential the transition from the triplet-singlet phase into the singlet–singlet or triplet–triplet phase is possible only through the normal phase.

For the Fermi-liquid model of superfluid nuclear matter with the effective Skyrme interaction as the nucleon–nucleon potential we conclude that at densities close to the nuclear-matter saturation density the superfluid phase that occurs is the triplet–singlet one (the particles that form a pair are a neutron and a proton in the triplet spin state). Although this result might seem somewhat unexpected (singlet spin pairing is realized in atomic nuclei), there has been research into the superfluidity of nuclear matter that studied spin-triplet nucleon pairing (with other interaction potentials; see the literature cited in the Introduction).

For triplet–singlet superfluid phases there is the possibility of states of the pairing proton and neutron that are unitary or nonunitary in spin (the projection of the total spin, $S_z$, is either $S_z=0$ or $S_z=\pm 1$, respectively). We have found the density-dependence of the energy gap at $T=0$ in the unitary and nonunitary cases. For nonunitary states the transition to a superfluid phase is accompanied by the appearance of spontaneous magnetization, which incidentally is not very strong (see Table II).

Among the other problems related to superfluid nuclear matter we would like to mention the effect of the asymmetry of nuclear matter on the transition temperature and the energy gap for various superfluid phases (asymmetric nuclear matter), the effect of bound nucleon states (deuterons)[20,36,37] on the superfluid properties,[27] and the problem of setting up the equations of hydrodynamics and of studying kinetic phenomena.[38]

The authors would like to express their deep gratitude to Profs. G. Röpke and P. Schuck for the useful remarks and the discussions concerning the results of the work.

[1)]By the nuclear-matter saturation density we mean the density at which the binding energy per nucleon is at its maximum.

[2)]Note that in our study we limited ourselves to temperatures $\gtrsim 10^{-3}$ MeV, this range being of principal interest owing to possible astrophysical applications (for neutron stars, for instance, the inner temperatures are of the order $10^{-2}$ MeV; see Ref. 34). If we continue the curves $T_c(30)$ and $T_c(00)$ (or $T_c(33)$) still farther in density, they will obviously intersect, but the point of intersection corresponds to a transition temperature $T_c \ll 10^{-3}$ MeV.

[1] A. Bohr, B. Mottelson, D. Pines, Phys. Rev. **110**, 936 (1958).
[2] S. T. Belyaev, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **31**, 11 (1959).
[3] N. N. Bogolyubov, Dokl. Akad. Nauk SSSR **119**, 52 (1958) [Sov. Phys. Doklady **3**, 279 (1958)].
[4] V. G. Solov'ev, Zh. Éksp. Teor. Fiz. **35**, 823 (1958) [Sov. Phys. JETP **8**, 572 (1959)]; **36**, 1869 (1959) [**9**, 1325 (1959)]; Nucl. Phys. **9**, 655 (1958).
[5] V. G. Solov'ev, *Effect of Pair Correlations of the Superconducting Type on the Properties of Atomic Nuclei* [in Russian], Atomizdat, Moscow (1963).
[6] L. D. Landau, Zh. Éksp. Teor. Fiz. **30**, 1058 (1956) [Sov. Phys. JETP **3**, 920 (1956)]; **32**, 59 (1957) [**5**, 101 (1957)].
[7] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **106**, 162 (1957); **108**, 1175 (1957).
[8] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Quantum Field Theoretical Methods in Statistical Physics*, Pergamon Press, New York (1965).
[9] N. N. Bogolyubov, V. V. Tolmachev, and D. V. Shirkov, *New Method in Superconductivity*, Plenum Press, New York (1959).
[10] A. B. Migdal, *Theory of Finite Fermi Systems and Properties of Atomic Nuclei*, Interscience, New York (1967).
[11] V. G. Solov'ev, *Theory of Atomic Nuclei*, Inst. of Physics, Bristol, U.K. (1992).
[12] Th. Alm, G. Röpke, and M. Schmidt, Z. Phys. A **337**, 355 (1990).
[13] L. Amundsen and E. Ostgaard, Nucl. Phys. A **437**, 487 (1985).
[14] R. K. Su, S. D. Yang, and T. T. S. Kuo, Phys. Rev. C **35**, 1539 (1987); M. F. Jiang and T. T. S. Kuo, Nucl. Phys. A **481**, 294 (1988).
[15] M. Baldo, J. Cugnon, A. Lejeune *et al.*, Nucl. Phys. A **451**, 509 (1986).
[16] J. M. Chen, J. W. Clark, E. Krotschek *et al.*, Nucl. Phys. A **451**, 509 (1986).
[17] R. Tamagaki, Prog. Theor. Phys. **44**, 905 (1970).
[18] L. Amundsen and E. Ostgaard, Nucl. Phys. A **442**, 163 (1985).
[19] A. Sedrakian, G. Röpke, and Th. Alm, Nucl. Phys. A **594**, 355 (1995).
[20] Th. Alm, B' L. Friman, G. Röpke *et al.*, Nucl. Phys. A **551**, 45 (1993).
[21] Th. Alm, G. Röpke, A. Sedrakian *et al.*, Preprint MPG-VT-UR 66/95 (1995).
[22] T. H. R. Skyrme, Nucl. Phys. **9**, 615 (1959).
[23] D. Vautherin and D. M. Brink, Phys. Rev. C **5**, 626 (1972); B. I. Barts, Yu. L. Bolotin, E. V. Inopin, and V. Yu. Gonchar, *The Hartree–Fock Method in Nuclear Theory* [in Russian], Naukova Dumka, Kiev (1982).
[24] M. Brack, C. Guet, and H.-B. Hakansson, Phys. Rep. **123**, 275 (1985).
[25] A. I. Akhiezer, V. V. Krasil'nikov, S. V. Peletminskii *et al.*, Phys. Rep. **245**, 1 (1994).
[26] V. V. Krasil'nikov, S. V. Peletminskii, and A. A. Yatsenko, Physica A **162**, 513 (1990).
[27] A. I. Akhiezer, S. V. Peletminskii, and A. A. Yatsenko, Phys. Lett. A **151**, 99 (1990).
[28] A. I. Akhiezer, V. V. Krasil'nikov, S. V. Peletminskii, and A. A. Yatsenko, Usp. Fiz. Nauk **163**, No. 2, 1 (1993) [Phys. Usp. **36**, 35 (1993)].
[29] H. London and F. London, Proc. R. Soc. London **149**, 71 (1935).
[30] V. L. Ginzburg and L. D. Landau, Zh. Éksp. Teor. Fiz. **20**, 1064 (1950).
[31] G. M. Eliashberg, Zh. Éksp. Teor. Fiz. **39**, 1437 (1960) [Sov. Phys. JETP **12**, 1000 (1961)].
[32] L. P.;Gor'kov, Zh. Éksp. Teor. Fiz. **34**, 1518 (1958) [Sov. Phys. JETP **7**, 1048 (1958)].
[33] I. A. Akhiezer and S. V. Peletminskiĭ, *Methods of Statistical Physics*, Pergamon Press, Oxford (1980).
[34] D. M. Sedrakyan and K. M. Shakhabasyan, Usp. Fiz. Nauk **161**, No. 7, 3 (1991) [Sov. Phys. Usp. **34**, 555 (1991)].
[35] A. A. Isaev and S. V. Peletminskiĭ, Ukr. Fiz. Zh. **37**, 952 (1992).
[36] M. Baldo, U. Lombardo, and P. Schuck, Phys. Rev. C **52**, 975 (1995).
[37] G. Röpke, Ann. Phys. **3**, 145 (1994).
[38] K. Morawetz and D. Kremp, Z. Phys. A **351**, 157 (1995).

Translated by Eugene Yankovsky

# Suppression of superconductivity close to the metal–insulator transition in strongly disordered systems

É. Z. Kuchinskiĭ, M. V. Sadovskiĭ, and M. A. Érkabaev

*Electrophysics Institute Ural Branch, Russian Academy of Sciences, 620049 Ekaterinburg, Russia*
(Submitted 24 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 192–199 (July 1997)

By means of the self-consistent theory proposed earlier for a metal–insulator transition in strongly disordered systems, which takes into account interelectron interaction effects, the effects of the suppression of the superconducting-transition temperature $T_c$, caused by the formation of a Coulomb pseudo-gap in the density of states, are studied in a wide interval of disorder values—from a weakly disordered metal to an Anderson insulator. It is shown that the proposed theory gives a satisfactory description of the experimental data for a number of systems that have been studied. © *1997 American Institute of Physics.* [S1063-7761(97)01607-7]

## 1. INTRODUCTION

The problem of the degradation of the superconducting-transition temperature under conditions of strong disordering has attracted the attention of theoreticians for a rather long time.[1] It is closely associated with the question of the breakdown of the superconducting state close to the metal–insulator transition caused by disordering.[2] A number of mechanisms for the suppression of $T_c$ have been proposed, such as an increase of the Coulomb pseudopotential,[3,4] the effect of Coulomb corrections to the density of states,[5] etc. Most of these papers discussed only small corrections to $T_c$ because of these mechanisms.

The theory of the metal–insulator transition proposed in Refs. 6 and 7, which generalizes the self-consistent localization theory[8,9] in the direction of taking into account electron–electron interaction effects, made it possible to study the behavior of a generalized diffusion coefficient over a wide range of variation of the system parameters both in the metallic and in the insulator regions. The substantial influence of electron–electron interaction on the generalized diffusion coefficient was treated. These results were used to study the behavior of the single-particle density of states of the system, taking into account the influence of electron–electron interaction effects.

The results of the corresponding calculations demonstrate the formation and evolution of a Coulomb pseudo-gap in the density of states of a system close to the Fermi level. In the metallic region, the behavior of the density of states close to the Coulomb pseudo-gap corresponds to the ordinary Al'tshuler–Aronov root correction.[10] When one approaches the metal–insulator transition as the disorder parameter increases, the depth of the pseudo-gap increases and the effective width of the region of the root behavior decreases; at the metal–insulator transition point, the density of states at the Fermi level goes to zero, i.e., a Coulomb gap forms. In the insulator region, for the case of a band of finite width in the region of the Coulomb gap, a quadratic dependence of the density of states is obtained. The effective width of the corresponding region increases with increasing disorder parameter. This recalls the well-known behavior of the Efros–Shklovskii Coulomb gap[11] in the insulator region far from the metal–insulator transition point. Such behavior of the density of states gives good qualitative agreement with experiments in a number of disordered systems close to the metal–insulator transition,[1] from amorphous alloys[12–16] to disordered single-crystal metal oxides, including high-temperature superconductors.[17]

In this paper, the results of calculations of the density of states of a system for the case of a band of finite width are used to numerically study how Coulomb pseudo-gap effects in the density of states affect the suppression of superconductivity close to the metal–insulator transition.

Superconductivity in strongly disordered systems will be treated in terms of a simple BCS model. In the weak-binding approximation, the linearized equation for the gap has the following form:[2]

$$\Delta(\xi)=-\int_{-\infty}^{\infty}d\xi' V(\xi,\xi')N(\xi')\,\frac{1}{2\xi'}\tanh\left(\frac{\xi'}{2T_c}\right)\Delta(\xi'). \tag{1}$$

Here $N(\xi)$ is the density of states of the disordered system averaged over the implementations of the disorder, allowing for electron–electron interaction effects, and $V(\xi,\xi')$ is the effective interaction potential. The only difference from the standard approach is that the nontrivial dependence of $N(\xi)$ on electron energy $\xi$ measured from the Fermi level $E_F$ is taken into account here.

It is assumed in BCS theory that an effective electron–electron attraction exists, which is determined by a certain balance between pairing due to electron–phonon interaction and Coulomb repulsion. The following will be regarded as the effective interaction potential:

$$V(\xi,\xi')=V_c(\xi,\xi')+V_{\rm ph}(\xi,\xi'), \tag{2}$$

where $V_c(\xi,\xi')=V_c\theta(E_F-|\xi|)\theta(E_F-|\xi'|)$ and $V_{\rm ph}(\xi,\xi')=-V_{\rm ph}\theta(\omega_D-|\xi|)\theta(\omega_D-|\xi'|)$ are the electron–electron and electron–phonon interaction potentials, respectively, and $\omega_D$ is the Debye frequency. The constants $V_c>0$ and $V_{\rm ph}>0$ correspond to repulsion and attraction, acting in substantially different energy intervals: $E_F\gg\omega_D$.

After substituting this expression into Eq. (1) and transforming, using the parity of the slit function $\Delta(\xi)$, we get

$$\Delta(\xi)=[V_{\rm ph}\theta(\omega_D-\xi)-V_c\theta(E_F-\xi)]\times\int_0^{\omega_D}d\xi' N(\xi')\,\frac{1}{\xi'}\tanh\left(\frac{\xi'}{2T_c}\right)\Delta(\xi')-V_c\theta(E_f-\xi)\int_{\omega_D}^{E_F}d\xi' N(\xi')\,\frac{1}{\xi'}\tanh\left(\frac{\xi'}{2T_c}\right)\Delta(\xi'). \quad (3)$$

As usual, we shall seek the solution of this equation in a two-step form:[18]

$$\Delta(\xi)=\begin{cases}\Delta_{\rm ph}, & |\xi|<\omega_D,\\ \Delta_c, & \omega_D<|\xi|<E_F,\end{cases} \quad (4)$$

where $\Delta_{\rm ph}$ and $\Delta_c$ are certain constants that can be determined, after substituting Eq. (4) into Eq. (3), from a system of homogeneous equations of the following form:

$$\left\{1-(V_{\rm ph}-V_c)N_0(0)K\left(\frac{\omega_D}{2T_c}\right)\right\}\Delta_{\rm ph}+V_cN_0(0)\left[K\left(\frac{E_F}{2T_c}\right)-K\left(\frac{\omega_D}{2T_c}\right)\right]\Delta_c=0,$$
$$V_cN_0(0)K\left(\frac{\omega_D}{2T_c}\right)\Delta_{\rm ph}+\left\{1+V_cN_0(0)\left[K\left(\frac{E_F}{2T_c}\right)-K\left(\frac{\omega_D}{2T_c}\right)\right]\right\}\Delta_c=0. \quad (5)$$

Here $N_0(0)$ is the single-particle density of states of noninteracting electrons at the Fermi level, and we have introduced the notation

$$K(\xi)=\int_0^{\xi}d\xi'\,\frac{1}{\xi}\tanh\,\xi'\left[\frac{N(2T_c\xi')}{N_0(0)}\right]. \quad (6)$$

The condition for this homogeneous system of equations to be solvable is the equation for determining $T_c$:

$$(\lambda-\mu^*)K\left(\frac{\omega_D}{2T_c}\right)=1,$$
$$\mu^*=\mu\left\{1+\mu\left[K\left(\frac{E_F}{2T_c}\right)-K\left(\frac{\omega_D}{2T_c}\right)\right]\right\}^{-1}, \quad (7)$$

where $\mu^*$ is the Coulomb pseudopotential, $\mu=V_cN_0(0)$ is the Coulomb repulsion constant, and $\lambda=V_{\rm ph}N_0(0)$ is the pairing constant due to the electron–phonon interaction. In the pure limit, when the density of states at the Fermi level can be regarded as constant, the usual equation of BCS theory follows from this.

Equation (7) for determining $T_c$ has been studied numerically over a wide region of variation of the system parameters in both the metal and the insulating states. The density of states of the system was computed using the lower-order corrections in the interelectron interaction:[6,7]

$$N(\xi)=-\frac{1}{\pi}\,{\rm Im}\int\frac{d^3\mathbf{p}}{(2\pi)^3}\,G^R(\mathbf{p},\xi), \quad (8)$$

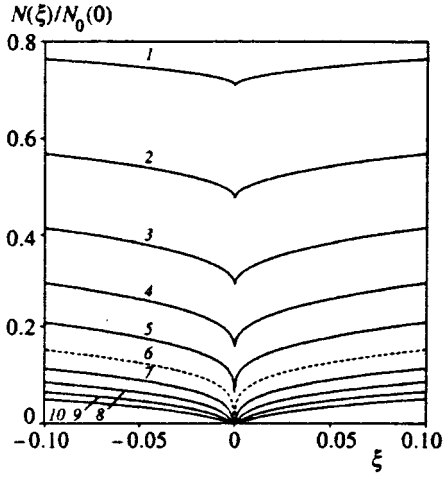



FIG. 1. Density of states of a system in the case of a band of finite width $2E_F$ when $(8/3\pi)\mu=1.0$, for various values of the disorder parameter $(p_Fl)^{-1}$: *1*—0.1...,*5*—0.5 in the metallic region, *7*—0.7...,*10*—1.0 in the insulator region. The dashed curve *6* corresponds to the metal–insulator transition point. Energy $\varepsilon$ is in units of $D_0k_0^2$ in the graph.

where $G^{R(A)}(\mathbf{p},\xi)=[\xi-\xi_p\pm i\gamma-\Sigma_{ee}^{R(A)}(\mathbf{p},\xi)]^{-1}$ is the retarded (advanced) single-particle Green's function, and $\Sigma_{ee}^{R(A)}(\mathbf{p},\xi)$ is the Fock contribution to the eigenenergy part:[6,10]

$$\Sigma_{ee}^{R(A)}(\mathbf{p},\xi)\approx 4i\gamma^2\mu N_0^{-1}(0)G_0^{A(R)}(\mathbf{p},\xi)\times\int_{\xi}^{\infty}\frac{d\omega}{2\pi}\int_{|\mathbf{q}|<k_0}\frac{d^3\mathbf{q}}{(2\pi)^3}\times\frac{1}{[-i\omega+D(\omega)q^2]^2}. \quad (9)$$

Here $D(\omega)$ is a generalized diffusion coefficient, which satisfies the following self-consistent nonlinear integral equation:[6,7]

$$\frac{D(\omega)}{D_0}=1-\frac{1}{\pi N_0(0)}\frac{D(\omega)}{D_0}\int_{|\mathbf{q}|<k_0}\frac{d^3\mathbf{q}}{(2\pi)^3}\times\frac{1}{-i\omega+D(\omega)q^2}+\frac{8i}{3\pi}\frac{\mu D_0}{\pi N_0(0)}\int_{\omega}^{\infty}d\Omega\int_{|\mathbf{q}|<k_0}\frac{d^3\mathbf{q}}{(2\pi)^3}\times\frac{q^2}{(-i(\Omega+\omega)+D(\Omega+\omega)q^2)(-i\Omega+D(\Omega)q^2)^2}, \quad (10)$$

where $D_0=E_F/3m\gamma$ is the classical diffusion coefficient, $\gamma=1/2\tau$ is the Born damping, $\tau$ is the free path time, $k_0=\min\{p_F,l^{-1}\}$ is the cutoff momentum, $p_F$ is the Fermi momentum, and $l$ is the free path length. The values shown below for static conductivity were also obtained by numerically solving Eq. (10).[6,7]

Figure 1 shows the behavior of the density of states of the system close to the Fermi level, demonstrating the evo-

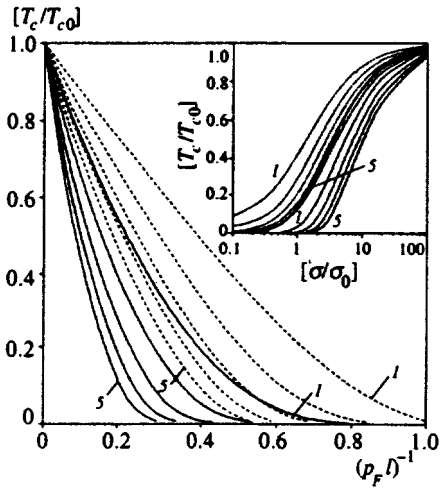



FIG. 2. Degradation of $T_c$ as a function of the disorder parameter $(p_F l)^{-1}$ for a fixed pairing constant $\lambda$ ($\lambda = 0.5$—continuous curves, $\lambda = 1.0$—dashed curves) for various values of the Coulomb repulsion constant $(8/3\pi)\mu$: *1*—0.2,⋯ ,*5*—1.0. The inset shows the dependence of $T_c$ on the static conductivity $\sigma$ of the system for the corresponding values of the pairing constant $\lambda$ and Coulomb repulsion constant $\mu$.

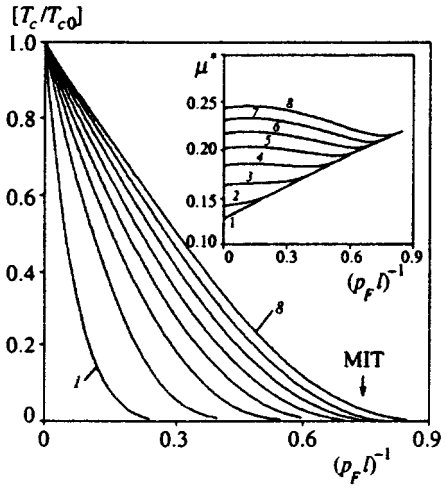



FIG. 3. Degradation of $T_c$ as a function of the disorder parameter $(p_F l)^{-1}$ for fixed Coulomb repulsion constant $(8/3\pi)\mu = 0.4$ for various values of the pairing constant $\lambda$: *1*—0.3, *2*—0.4,..., *8*—1.0. The inset shows the dependence of the Coulomb pseudopotential $\mu^*$ on the disorder parameter $(p_F l)^{-1}$ for the corresponding values of the pairing constant $\lambda$ and the Coulomb repulsion $\mu$. The arrow shows the position of the metal–insulator transition point $\mu$.

lution of the Coulomb pseudo-gap as the disorder parameter increases. It is this behavior that results in suppression of the superconducting transition temperature.

The graphs in Fig. 2 demonstrate the suppression of $T_c$ with increasing disorder parameter $(p_F l)^{-1}$ for various values of the Coulomb repulsion constant $\mu$ with fixed pairing constant $\lambda$. For large $\mu$, as disorder $(p_F l)^{-1}$ increases, $T_c$ rapidly decreases and goes to zero in the metallic region far from the metal–insulator transition. When $\mu$ is reduced, the falloff of $T_c$ with increasing disorder $(p_F l)^{-1}$ slows down, and, for small $\mu$ and large $\lambda$ (dashed curves in figure), superconductivity can occur in the insulating region.[2] The latter is clearly demonstrated by the graphs in the inset of Fig. 2, which shows the dependence of $T_c$ on the static conductivity $\sigma$ of the system for corresponding values of the pairing constant $\lambda$ and the Coulomb repulsion constant $\mu$. For large $\mu$, $T_c$ rapidly decreases as conductivity $\sigma$ decreases, and superconductivity is suppressed in the metallic region rather far from the metal–insulator transition. When $\mu$ is reduced, the falloff of $T_c$ slows down with decreasing conductivity $\sigma$, and, for small $\mu$ and rather large $\lambda$ (dashed curves in inset), $T_c$ remains finite in the limit $\sigma \to 0$.

The graphs in Fig. 3 demonstrate the degradation of $T_c$ as the disorder parameter $(p_F l)^{-1}$ increases for various values of the pairing constant $\lambda$ with fixed Coulomb repulsion constant $\mu$. For small $\lambda$, as the disorder $(p_F l)^{-1}$ increases, $T_c$ rapidly decreases and goes to zero in the metallic region far from the metal–insulator transition. When $\lambda$ is increased, the decrease of $T_c$ with increasing disorder $(p_F l)^{-1}$ slows down, and, for sufficiently large $\lambda$, the superconductivity is suppressed only in the insulating region. The dependence of the Coulomb pseudopotential $\mu^*$ on the disorder parameter $(p_F l)^{-1}$ shown in the inset of Fig. 3 for corresponding values of the pairing constant $\lambda$ and the Coulomb repulsion constant $\mu$ demonstrates an insignificant increase of the Coulomb pseudopotential $\mu^*$ with increasing disorder $(p_F l)^{-1}$ close to the superconductivity-suppression point. This is apparently fairly natural, since the different processes that renormalize the matrix element of the Coulomb interaction in Eq. (2) because of Anderson localization effects and electron–electron interaction and that substantially increase the Coulomb pseudopotential close to the metal–insulator transition[2] are not considered in this case.

Similar behavior of $T_c$ as a function of static conductivity $\sigma$ and of the disorder parameter was experimentally observed in a number of disordered systems that remain superconducting close to the metal–insulator transition caused by disordering.[1,2,12–17,19–21] The results of our numerical calculations agree well with experiments in the amorphous alloys $InO_x$,[14] $Nb_xSi_{1-x}$,[15,16] and $Au_xSi_{1-x}$.[19–21]

Reference 14 presented the results of measurements of the disorder parameter $(p_F l)^{-1}$ for the amorphous alloy $InO_x$, as well as data for $T_c$ and static conductivity close to the metal–insulator transition.

According to Refs. 6 and 7, the static conductivity of the system close to the metal–insulator transition has the following form:

$$\sigma = \sigma_0[(p_F l) W_c(\mu) - 1]. \tag{11}$$

Here $\sigma_0$ is some characteristic conductivity scale close to the metal–insulator transition, and $W_c(\mu)$ is the disorder parameter corresponding to the metal–insulator transition, which depends on the Coulomb repulsion constant.

Approximating the experiment for the static conductivity of the amorphous alloy $InO_x$ by Eq. (11) makes it possible to estimate the characteristic conductivity scale $\sigma_0$ and, from the value of $W_c$, the Coulomb repulsion constant $\mu$. Satisfactory correlations (see inset in Fig. 3) are obtained for the following values of the parameters: $\sigma_0 \simeq 324.95$ $(\Omega \cdot \mathrm{cm})^{-1}$, $W_c \simeq 0.606$, and $\mu \simeq 1.0$.

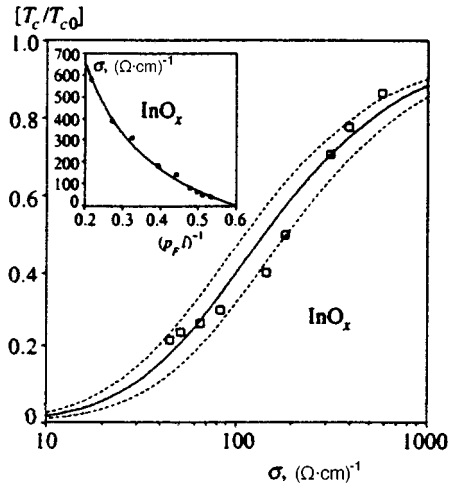



FIG. 4. Behavior of $T_c$ as a function of static conductivity $\sigma$ for the amorphous alloy $InO_x$. The inset shows the results of an approximation of the data for the static conductivity $\sigma$ as a function of the disorder parameter $(p_F l)^{-1}$.

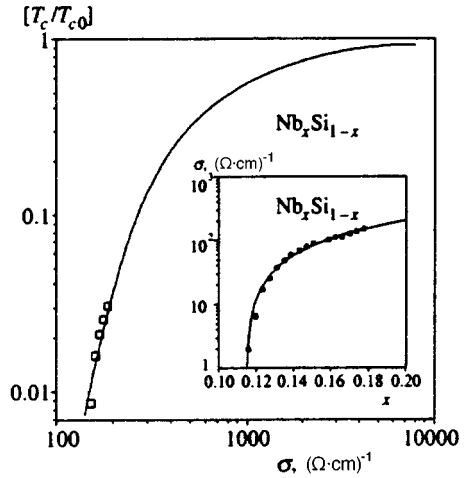



FIG. 5. Behavior of $T_c$ as a function of static conductivity $\sigma$ for the amorphous alloy $Nb_xSi_{1-x}$. The inset shows the results of an approximation of the data for the static conductivity $\sigma$ as a function of the Nb concentration.

Figure 4 shows a comparison of our results with the experimental data for $T_c$ as a function of the static conductivity $\sigma$ of the amorphous alloy $InO_x$, using $T_{c0}=3.41$ K, $\omega_D=112$ K, and $E_F=9.98\times10^4$ K, $[\omega_D/E_F]\simeq1.1\times10^{-3}$ for pure In and the resulting values of $\sigma_0$ and $\mu$, which makes it possible to estimate the pairing constant $\lambda$. Satisfactory correlation is obtained for $\lambda\simeq0.45$. The dashed curves correspond to the values $\lambda\simeq0.4$ and 0.5.

Let us consider the results of studies of the dependence of $T_c$ and the static conductivity on the Si concentration in the amorphous alloys $Nb_xSi_{1-x}$ (Refs. 15, 16) and $Au_xSi_{1-x}$ (Refs. 19–21) close to the metal–insulator transition. Assuming a disorder parameter proportional to the Si concentration for these systems, so that $(p_F l)^{-1}\sim1-x$, we transform Eq. (11) for the static conductivity close to the metal–insulator transition to the form

$$\sigma=\sigma_0\,\frac{x-x_c}{1-x}, \tag{12}$$

where $x_c$ is the critical concentration (corresponding to Nb or Au) at the metal–insulator transition point.

Approximating the experiment for the static conductivity of the amorphous alloys $Nb_xSi_{1-x}$ and $Au_xSi_{1-x}$ by Eq. (12) makes it possible to estimate the characteristic conductivity scale $\sigma_0$ and the critical concentration $x_c$. Satisfactory correlations (see the inset in Figs. 5 and 6) are obtained for the following values of the parameters:

$$Nb_xSi_{1-x}:\quad \sigma_0\simeq1963.9\ (\Omega\cdot cm)^{-1},\quad x_c\simeq0.115;$$

$$Au_xSi_{1-x}:\quad \sigma_0\simeq2782.13\ (\Omega\cdot cm)^{-1},\quad x_c\simeq0.14.$$

The graphs in Figs. 5 and 6 demonstrate the comparison of our results with the experimental data for $T_c$ as a function of static conductivity $\sigma$ in the amorphous alloys $Nb_xSi_{1-x}$ and $Au_xSi_{1-x}$, using the values of $\sigma_0$ shown above and the following parameters: $T_{c0}=9.26$ K, $\omega_D=276$ K, and $E_F=6.18\times10^4$ K, $[\omega_D/E_F]\simeq3.0\times10^{-3}$ for pure Nb; $T_{c0}=T_{c\,\max}\simeq0.86$ K, $\omega_D=170$ K, and $E_F=6.42\times10^4$ K, $[\omega_D/E_F]\simeq0.9\times10^{-3}$ for $Au_xSi_{1-x}$. This comparison makes it possible to estimate the pairing constant $\lambda$. Assuming a Coulomb repulsion constant of $\mu\simeq1$ for these systems, satisfactory correlation can be obtained with $\lambda\simeq0.54$ for $Nb_xSi_{1-x}$ and with $\lambda\simeq0.62$ for $Au_xSi_{1-x}$.

Of course, these computations, which are based on the BCS model, are oversimplified. A consistent approach to the problem of computing the superconducting transition temperature must be based on a solution of the Eliashberg equations and must use realistic models of the electron–electron interaction.[18] This is especially true of the results given above for large values of the pairing constant $\lambda$, which demonstrate that superconductivity can exist in the insulating region. At the same time, we have not questioned the genesis of the initial $T_{c0}$ in a pure system in this paper, but have been occupied only with the question of how $T_c$ depends on the disorder. In this sense, the results can be qualitatively applied in the strong bonding region. We should point out that it is still necessary to more consistently take into account disor-

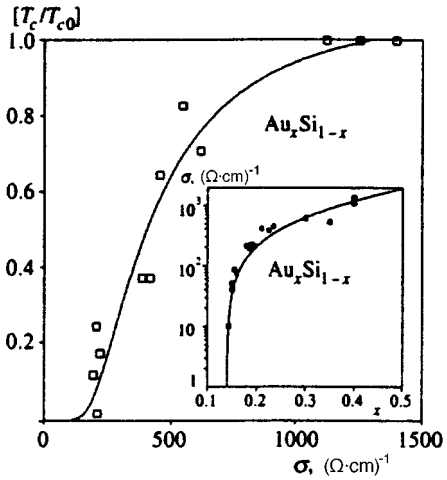



FIG. 6. Behavior of $T_c$ as a function of static conductivity $\sigma$ for the amorphous alloy $Au_xSi_{1-x}$. The inset shows the results of an approximation of the data for the static conductivity $\sigma$ as a function of the Au concentration.

der effects in the Coulomb-interaction matrix element, which is associated with an additional $T_c$-degradation mechanism.[2–4] As was pointed out above, this paper has taken into account only the effects of the formation of a Coulomb pseudo-gap in the density of states. It is possible that the satisfactory agreement with experiment obtained above indicates that the effects of the variation of the density of states play a dominant role in the $T_c$-degradation mechanism, as was noted earlier (at the level of small corrections) in Ref. 5.

This work was carried out with the partial financial support of the Russian Fund for Fundamental Research (Grant 96-02-16065) and also as part of Project IX.1 of the Statistical Physics Program of the State Committee for Science and Technology of Russia.

[1] D. Belitz and T. R. Kirkpatrick, Rev. Mod. Phys. **66**, 261 (1994).
[2] M. V. Sadovskiĭ, Sverkhprovodnost' (KIAE) **8**, 337 (1995) submitted to Phys. Rep. (1996).
[3] P. W. Anderson, K. A. Muttalib, and T. V. Ramakrishnan, Phys. Rev. B **28**, 117 (1983).
[4] L. N. Bulaevskii and M. V. Sadovskii, J. Low Temp. Phys. **59**, 89 (1985).
[5] D. Belitz *et al.*, Phys. Rev. B **40**, 111 (1989).
[6] É. Z. Kuchinskiĭ, M. V. Sadovskiĭ, V. G. Suvarov, and M. A. Érkabaev, Zh. Éksp. Teor. Fiz. **107**, 2027 (1995) [ JETP **80**, 1122 (1995)].
[7] É. Z. Kuchinskiĭ and M. A. Érkabaev, Fiz. Tverd. Tela **39**, 412 (1997) [ Phys. Solid State **39**, 357 (1997)].
[8] D. Volthardt and P. Wolfle, in *Electronic Phase Transitions*, ed. W. Hanke and Yu. Kopaev (North-Holland, Amsterdam, 1992), p. 1.
[9] M. V. Sadovskii, Sci. Rev. Phys. A **7**, 1 (1986).
[10] B. L. Al'tshuler and A. G. Aronov, Zh. Éksp. Teor. Fiz. **77**, 2028 (1979) [ Sov. Phys. JETP **50**, 968 (1979)].
[11] A. L. Efros and B. I. Shklovskii, in *Electron–Electron Interaction in Disordered Systems*, ed. M. Pollak and A. L. Efros (North-Holland, Amsterdam, 1984), p. 409.
[12] W. L. McMillan and J. Mochel, Phys. Rev. Lett. **46**, 556 (1981).
[13] Y. Imry and Z. Ovadyahu, Phys. Rev. Lett. **49**, 841 (1982).
[14] A. T. Fiory and A. F. Hebard, Phys. Rev. Lett. **52**, 2057 (1984).
[15] G. Hertel, D. J. Bishop, E. G. Spencer, J. M. Rowel, and R. C. Dynes, Phys. Rev. Lett. **50**, 743 (1983).
[16] D. J. Bishop, E. G. Spencer, and R. C. Dynes, Solid-State Electron. **28**, 73 (1985).
[17] H. Srikanth, K. P. Rajeev, G. V. Shivashankar, and A. K. Raychaudhuri, Physica C **195**, 87 (1992).
[18] S. V. Vonsovskiĭ, Yu. A. Izyumov, and É. Z. Kurmaev, *Superconductivity of the Transition Metals, Their Alloys and Compounds* (Nauka, Moscow, 1997).
[19] N. Nishida, M. Yamaguchi, T. Furubayashi, M. Morigaki, H. Ishimoto, and K. Ono, Solid State Commun. **44**, 305 (1982).
[20] T. Furubayashi, N. Nishida, M. Yamaguchi, M. Morigaki, and H. Ishimoto, Solid State Commun. **55**, 513 (1985).
[21] N. Nishida, T. Furubayashi, M. Yamaguchi, M. Morigaki, and H. Ishimoto, Solid-State Electron. **28**, 81 (1985).

Translated by W. J. Manthey

# Anomalies of the baric and temperature dependences of the elastic characteristics of ice during solid-phase amorphization and the phase transition in the amorphous state

E. L. Gromnitskaya, O. V. Stal'gorova, and V. V. Brazhkin

*L. F. Vereshchagin Institute of High-Pressure Physics, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia*
(Submitted 29 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 200–208 (July 1997)

The elastic characteristics of ice up to pressures of 1.7 GPa are determined for the first time at a temperature of 77 K, along with features of their variation associated with the phase transformation of hexagonal ice I*h* into high-density amorphous ice *hda*. The elastic instability of the ice lattice before solid-phase amorphization is experimentally confirmed. Elastic instability during a transition from one amorphous state to another amorphous state was also observed for the first time; this took place when *hda* ice was warmed at $p=0.05$ GPa from $T=77$ K. © *1997 American Institute of Physics.* [S1063-7761(97)01707-1]

## 1. INTRODUCTION

It is well known that water is not only a widely distributed substance on earth and in outer space, but also is a very interesting object for scientific studies with practical applications. The phase diagram of water, studied with varying temperatures $T$ and pressures $p$ that largely correspond to the conditions actually existing on the planets of the solar system, demonstrates more than ten crystal modifications of ice that possess different physical properties and is an object of numerous but far from complete investigations.[1–16]

Moreover, the phase transformation detected not long ago of hexagonal ice I*h* (the atmospheric-pressure phase) into high-density amorphous ice (*hda*) and the transition of ice from *hda* into low-density amorphous ice (*lda*) are of interest from the viewpoint of studying the phenomenon of solid-phase amorphization and transitions from one amorphous state to another. It has been established[1,2] that, at $T=77$ K, under the action of high pressure, the transformation I$h-hda$ occurs at $p\approx 1.1$ GPa, with the *hda* phase being maintained when the pressure is removed. Warming *hda* ice at atmospheric pressure, as shown by x-ray diffraction methods, successively causes the transitions $hda-lda$ at $T\approx 120$ K, then $lda-$I$c$ (the cubic modification) at $T\approx 140-145$ K, and finally I$c-$I$h$ at $T\approx 160$ K.[1–4,6,8]

More than a hundred publications are devoted to the phenomenon of solid-phase amorphization under pressure, and there are fairly complete reviews (see, for example, Refs. 5 and 17). At the same time, the mechanism of this phenomenon remains largely unclear. The most popular concept to explain the phenomenon of solid-state amorphization is currently the ''elastic instability'' model of the crystal lattice at definite $p$ and $T$. The softening of the elastic characteristics before amorphization, occurring under the action of irradiation, during supersaturation of solid solutions, or obtained by other methods, has been observed many times.[18,19] However, no experimental studies have been made until recently of the elastic properties of the lattice during solid-phase amorphization under the influence of pressure.

As far as transformations in the amorphous state are concerned, the facts of polymorphism between amorphous modifications under pressure are fairly reliably established by now, and the structure, the optical properties, and the bulk characteristics of $H_2O$, $GeO_2$, and $SiO_2$ during transitions in the amorphous state have also been experimentally studied.[1–5,20–23] However, the development of a microscopic theory of transformations in the amorphous state requires additional experimental information.

The I$h-hda$ and $hda-lda$ transitions in ice occur in an experimentally accessible range of pressures and temperatures, which makes ice an extremely attractive model object for verifying various concepts of solid-phase amorphization and transitions in the amorphous state. A study of the elastic properties of both crystalline and amorphous modifications of ice, as well as an investigation of their features during solid-phase amorphization and during transitions in the amorphous state would contribute greatly to the solution of these problems. However, there are not many corresponding experimental papers.[10–16] The elastic properties of I*h* ice were studied earlier by Brillouin scattering under a pressure of up to 0.28 GPa at $T=237.5$ K (Ref. 10) and in ultrasound measurements up to $p=0.7$ GPa at $T=248$ K.[11] Recent studies by Brillouin scattering made it possible to determine the elastic moduli of single crystals of ice VI and VII at room temperature and ice III at $T=253$ K.[13–16] In Ref. 12, we presented for the first time experimental confirmation of the *elastic instability* of the crystal structure of I*h* ice before the transition to the amorphous state under the action of high pressure at $T=77$ K.

## 2. EXPERIMENTAL TECHNIQUE

This paper is a more complete presentation and refinement of the results of the studies of the elastic properties of ice by a pulsed ultrasound method begun earlier in Ref. 12. The $p-T$ dependences of the velocities of longitudinal and transverse ultrasound waves for ice are obtained for the first time in the pressure range from zero to 1.7 GPa and a temperature of $T=77$ K, as well as in the natural warming regime from $T=77$ K to $T=180$ K at a constant pressure of

$p=0.05$ GPa. The features of the behavior of the elastic properties of ice during the I$h-hda$, $hda-lda$, and $lda-$I$c$ transitions are established and studied.

The ultrasound piezometer described in Ref. 24 was used for the low-temperature studies. The samples of ice were prepared from distilled water by rapid cooling in order to obtain homogeneous polycrystalline samples. After preparation, the samples, in the form of cylinders 8–10 mm high and $\approx 17$ mm in diameter in thin-wall ($\approx 0.04$ mm) lead shells, were placed into a high-pressure cell and were isolated from the punches by thin copper foil (0.02 mm) over the ends. During the experiment, the changes of the times of flight of the longitudinal ultrasound waves, $\Delta t_l(p)$ or $\Delta t_l(T)$, and of the transverse ultrasound waves, $\Delta t_t(p)$ or $\Delta t_t(T)$, were determined by the method of visual superposition of the signals.[25] We used $x$-cut and $y$-cut quartz plates, with carrier frequencies of 5 and 3 MHz, respectively, as piezoelectric sensors. The change in the sample height, $\Delta l(p)$ or $\Delta l(T)$, was determined by means of dial-type micrometer indicators remote from the low-temperature zone.

## 3. MEASURED RESULTS AND DISCUSSION

### 3.1. Measurements at $T=$ const = 77 K

Measurements were carried out at $T=77$ K with the high-pressure unit totally immersed in the liquid-nitrogen reservoir. The temperature in the working volume was measured with four copper–constantan thermocouples mounted in the immediate vicinity of the sample. The pressure in the chamber was determined from the force of the press and a multiplication factor to take into account the deformation of the piezometer channel with pressure and temperature. It was shown that this pressure-measurement method is sufficiently accurate by determining the CsI–CsII phase-transition pressure in cesium: $p=(2.25\pm 0.02)$ GPa.[26]

The baric dependences $\Delta t_l(p)$, $\Delta t_t(p)$, and $\Delta l(p)$ were measured in the pressure interval from zero to 1.7 GPa while increasing and relieving the pressure after the temperature stabilized at $T=77$ K at all points of the ultrasound piezometer. The resulting behavior was used to calculate the velocities of the longitudinal ultrasound wave, $v_l(p)$, and the transverse ultrasound wave, $v_t(p)$, which made it possible to determine the compression $x=V(p)/V_0$ of the ice, the adiabatic bulk modulus $K_S(p)$, the shear modulus $G(p)$, Young's modulus $E(p)$, and the Poisson ratio $\sigma(p)$, which are shown in Figs. 1 and 2. The results of five experiments on the determination of $v_l(p)$ and of four on the determination of $v_t(p)$ were used in the calculations. The calculations took into account corrections for friction in the piezometer channel, as well as for deformation of the chamber with pressure and temperature, which were determined in separate experiments.

The variation of the properties of ice over the range of pressures studied here can be broken up into several sections. As the pressure is raised from zero to 1.1 GPa, the resulting dependences characterize the I$h$ hexagonal phase of ice; then, from 1.1 GPa to 1.4 GPa, the phase transformation of the I$h$ hexagonal ice into the high-density amorphous phase of ice $hda$ occurs; next, as the pressure is raised from 1.4 GPa to 1.7 GPa and then as the pressure is removed, the $hda$ ice maintains its metastable state.

The compression of the I$h$ hexagonal ice increases its density by about 10% before the transition begins, with the bulk modulus $K_S(p)$ increasing nonlinearly with pressure. As the transition into the amorphous state is approached ($p\sim 1.0$ GPa), the increase of the modulus $K_S(p)$ slows down, so that the derivative with respect to pressure, $\partial K_S(p)/\partial p$, decreases by a factor of 1.5 by comparison with the original value. Such behavior of $K_S(p)$ can be explained by an increase of the negative contribution of the elastic modulus $c_{13}(p)$:[27]

$$K_S^2=\frac{[c_{33}(c_{11}+c_{12})-2c_{13}^2][2(c_{11}+c_{12})+c_{33}+4c_{13}]}{9(c_{11}+c_{12}+2c_{33}-4c_{13})},$$

which, according to Ref. 10, increases more rapidly with pressure than the other elastic moduli $c_{ij}$.

The most interesting phenomenon that precedes the amorphization of the I$h$ ice is the decrease with pressure of the transverse ultrasound wave velocity $v_t(p)$ and, accordingly, of the shear modulus $G(p)$. Such a change of the shear characteristics indicates that the crystal lattice has lost stability because of softening of one of the shear modes ($TA_1$), which is discussed in more detail in Ref. 12. The strong increase of the Poisson ratio with pressure (27%) is evidence of a decrease of the degree of covalence of the bonds or, in other words, of decreased rigidity of the bonds in certain crystallographic directions in the hexagonal structure of ice (I$h$).

In a number of cases, the process of compressing the I$h$ ice was accompanied by a sharp dry crackling, which can be caused by ''penetration and slippage'' of the corresponding planes in which the microdefects accumulate that determine the decrease of the resistance to stress. It was pointed out in Ref. 28 that the mechanical resistance to shear for I$h$ ice decreases with pressure, and that extrapolation of its baric dependence to zero leads to a pressure of 1.0–1.5 GPa, which agrees well with the pressure of the I$h-hda$ transition. The behavior of ice under pressure at low temperatures thus gives a clear example of the connection of the softening of the elastic constants determined by interatomic interaction at the microscopic level and the corresponding falloff of the macroscopic mechanical characteristics.

The I$h-hda$ phase transition was recorded from the marked change of all the measured parameters ($\Delta t_l/t_0=19.0\%$, $\Delta t_t/t_0=38.0\%$, $\Delta l/l_0=20.0\pm 0.5\%$), which occurred especially intensely at a pressure of $p\approx 1.1\pm 0.1$ GPa, and this transition can be considered mainly complete by $p=1.4$ GPa. We took the pressure $p\approx 1.1\pm 0.1$ GPa to be the pressure of the I$h-hda$ phase transition, which is in good agreement with the work of other authors.[1–4,6,8]

The strong density increase ($\approx 20.0\pm 0.5$)% accompanying the I$h-hda$ transition measured in this work virtually coincides with the jump determined in Refs. 2, 4, 6, 8. As was already pointed out in Ref. 12, the transverse ultrasound wave velocity $v_t(p)$ increases significantly (19%), whereas the longitudinal ultrasound wave velocity increases by only

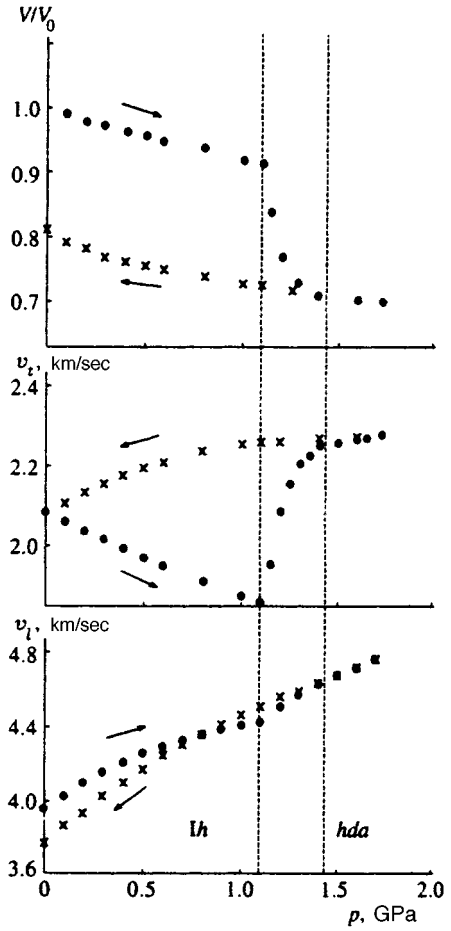



FIG. 1. Solid-phase amorphization of ice, the I$h-hda$ transition. Baric dependences at $T=\text{const}=77$ K of the compression $V/V_0$, the longitudinal velocity $v_l$ and transverse velocity $v_t$ of ultrasound waves as the pressure is increased (●) and relieved (×); $v_l$ ($p=0$, $T=77$ K) and $v_t$ ($p=0$, $T=77$ K) are from the data of Refs. 9–11.

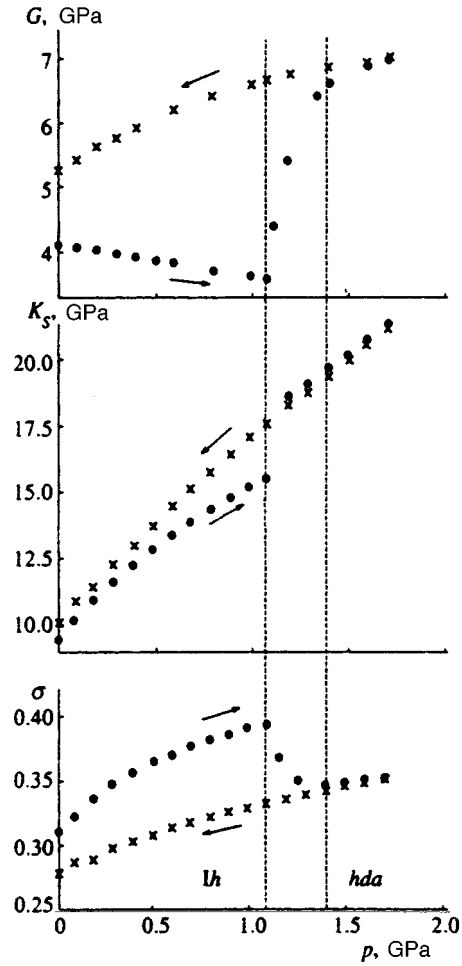



FIG. 2. Solid-phase amorphization of ice, the I$h-hda$ transition. Baric dependence at $T=\text{const}=77$ K of the adiabatic bulk modulus $K_S$, the shear modulus $G$, and the Poisson ratio $\sigma$ as the pressure is increased (●) and relieved (×).

2% during the transition. The increase of the adiabatic bulk modulus (26.0%) and the shear modulus (70.0%) and the decrease of the Poisson ratio (16.0%) are evidence that the rigidity of the bonds increases during the transition of I$h$ ice into the $hda$ phase.

The baric dependences of the elastic characteristics coincide and do not show any anomalies as the pressure is increased from 1.4 GPa to 1.7 GPa and when the load is subsequently completely removed (Figs. 1 and 2), from which it follows that amorphous high-density ice stays in a metastable state all the way to atmospheric pressure. We should point out that amorphous high-pressure ice and I$h$ ice have extremely similar pressure dependences that characterize the ability of the substance to react to volume deformations under pressure $[v_l(p),\ K_S(p)]$. It follows from the behavior under pressure of the shearing elastic characteristics $v_t(p)$ and $G(p)$, as well as the Poisson ratio, that $hda$ ice possesses greater rigidity and a larger degree of covalency than does I$h$ hexagonal ice.

The absolute values of the moduli $K_S(p)$ and $G(p)$ of $hda$ amorphous ice appreciably exceed the corresponding values for crystalline ices I$h$, II, III, and V extrapolated to our $p-T$ levels and are 10–30% less than the moduli of ices VI, VII, and VIII.[11,13–15] As a rule,[29–32] the amorphous phase has $K_S(p)$ values 5–30% less and $G(p)$ values 15–50% less than the corresponding crystalline prototypes. On the basis of experimental data concerning the structure and spectra of vibrations, Refs. 33 and 34 concluded that the close-order structures of $hda$ ice and of crystal modifications VI and VII are similar. The values of the elastic moduli obtained in our work also support this conclusion.

### 3.2. Measurements with $p=\text{const}=0.05$ GPa

The sequence of phase transformations of ice $hda-lda-$I$c$ was studied under conditions of warming at fixed pressure $p=0.05$ GPa. The warming rate was $\approx 0.8$ K/min in the temperature interval 77 K–110 K and 0.5 K/min in the interval 110–200 K. The warming characteristics and the use of the same type of high-pressure chambers show that the experimental conditions of Mishima *et al.*[1–4] virtually coincide with ours.

The temperature dependences of the ultrasound velocities $v_l$ and $v_t$ (Figs. 3 and 4) during warming of ice, first reported in this paper, showed features associated with phase transitions of ice in the temperature interval 77–180 K.

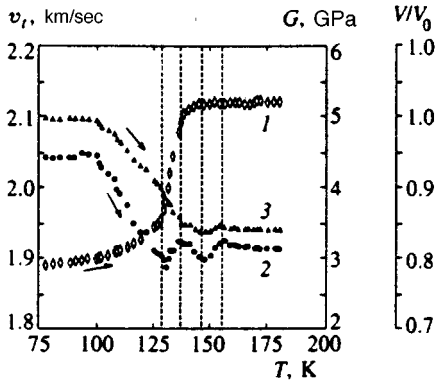



FIG. 3. Phase transition of high-density amorphous ice into low-density amorphous ice ($hda-lda$). Temperature dependences for $p=$const $=0.05$ GPa of the compression $V/V_0$ (*1*), the velocity $v_t$ of transverse ultrasound waves (*2*), and the shear modulus $G$ (*3*).

Three experiments each were carried out on the determination of $v_l(T)$ and $v_t(T)$. Note that the ultrasound method of studying transitions between disordered states and during crystallization of an amorphous phase is highly sensitive. Anomalies associated with these transitions appeared in the form of extrema of the $v_l(T)$ and $v_t(T)$ dependences. As we assume from Refs. 1–4, it is possible to discriminate sections corresponding to different stages of phase transformations from the positions of these anomalies: the section of the existence of amorphous high-density ice $hda$ (77–130 K), the region of its transition to amorphous low-density ice $lda$ (130–140 K), a narrow section of the metastable existence of $lda$ ice (140–148 K), and the crystallization of $lda$ ice into I$c$ ice (the cubic modification) at $T\approx 148-157$ K.

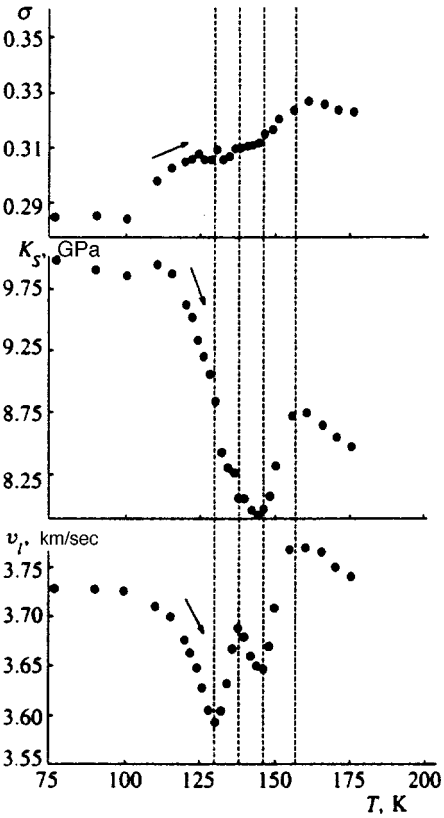



FIG. 4. Phase transition of high-density amorphous ice into low-density amorphous ice ($hda-lda$). Temperature dependences for $p=$const $=0.05$ GPa of the velocity $v_l$ of longitudinal ultrasound waves, the adiabatic bulk modulus $K_S$, and the Poisson ratio $\sigma$.

There is significant interest in the phenomenon of a transition from one amorphous state ($hda$) to another amorphous state ($lda$), which differ from each other in density by approximately 20%. The intense increase in volume $V(T)/V_0$ on the 130–140 K section and the minima of the $v_l(T)$ and $v_t(T)$ temperature dependences at $T=130$ K are logically associated with the $hda-lda$ transition. The unusually strong decrease of the $v_l(T)$ and $v_t(T)$ values preceding this and the character of the $V(T)/V_0$ variation at $T=110-130$ K indicate softening of the amorphous network of the ice before the transition from one amorphous state to the other amorphous state, and this was also observed for the first time. The strong dropoffs in the $K_S(T)$ and $G(T)$ variation at the threshold of the $hda-lda$ transition, which begins after $T=100$ K, and the acceleration of the growth of the derivatives $\partial v_t(T)/\partial T$, $\partial K_S(T)/\partial T$, and $\partial G(T)/\partial T$ as the temperature $T=130$ K is approached are also anomalous.

Note that the strong softening of the moduli in the $hda$ phase during heating from 100 K to 130 K [25% for $G(T)$ and 11% for $K_S(T)$] cannot be explained by the usual temperature dependence because of thermal broadening, since the characteristic variations of $G(T)$ and $K_S(T)$ with temperature for different crystalline modifications of ice equal about 0.1–0.2% per degree. The strong falloffs of the moduli are thus also apparently associated with anomalous softening of the amorphous network before the $hda-lda$ transformation.

It is natural to associate the maxima on the $v_l(T)$ and $v_t(T)$ behavior at $T=140$ K and the completion of the bulk effects by that time with the completion of the $hda-lda$ transition, while the experimental dependences of the elastic characteristics at $T=140-148$ K are assumed to be characteristic of the $lda$ phase of ice (the temperature interval agrees with Mishima's data[2–4]). Note that the moduli $K_S(T)$ and $G(T)$ are 5–10% lower in magnitude for the amorphous phase of ice $lda$ than the corresponding values for ices I$c$ and I$h$ (Figs. 3 and 4). This is additional evidence that the close-order structure of amorphous ice $lda$ is apparently based on the I$c$ or I$h$ structure of ice.

As established in Ref. 2, the density of amorphous ice $lda$, ice of the cubic modification I$c$, and hexagonal ice I$h$ are very close to each other in the $p-T$ region studied here, and therefore the $V(T)/V_0$ dependence has no singularities for $T>145$ K. However, there are extrema on the experimental $v_l(T)$ and $v_t(T)$ curves at $T=148$ K and $T=157$ K, which evidently reflect anomalies observed in the substance during the $lda-$I$c$ phase transition ($T=148-157$ K). As already noted, the transition from the amorphous phase to the crystalline phase is accompanied by appreciable increases of the adiabatic bulk modulus ($\approx 11\%$) and shear modulus ($\approx 3\%$); this is associated with an increase of the ordering and a strengthening of the bonds in the lattice. In this case, the Poisson ratio $\sigma(T)$ changes insignificantly (it increases by 2–3%); consequently, there is no appreciable change in

the degree of covalency of the structure, as should be observed when an amorphous phase crystallizes into the crystalline prototype with similar close order.

## 4. CONCLUSION

The baric dependence of the elastic characteristics of ice has thus been determined for the first time up to pressures of 1.7 GPa at a temperature of 77 K, as well as the behavior associated with the phase transition of hexagonal ice I*h* into amorphous high-density ice *hda*. The elastic instability of the lattice before such a phase transition has been experimentally confirmed. Elastic instability has also been observed for the first time during the transition from one amorphous state into another amorphous state that took place when *hda* ice was warmed from $T=77$ K at $p=0.05$ GPa. The use of an ultrasound technique made it possible to establish the complete sequence of *hda*–*lda*–I*c* phase transformations during this warming, which it is hard to do when bulk measurements are used.

This paper has thus used ice as a model object in order to study the behavior of the elastic characteristics before the crystal–amorphous-phase and amorphous-state–amorphous-state transformations, as well as during these transformations. The elastic-instability model was confirmed as the cause of solid-phase amorphization. For the transformation between the two amorphous phases, softening of the elastic characteristics in the region preceding the transformation was also observed for the first time. This experimental fact should evidently be taken into account when developing theoretical models of the polymorphism of the amorphous state. It was concluded in Refs. 35–37 that spinodals are present for the *lda*–*hda* transition, close to which the shapes of the *hda* and *lda* ices lose thermodynamic and apparently mechanical stability. The experimental data of this paper indirectly confirm the presence of such spinodals.

Moreover, much information has been obtained in this work on the elastic properties of the phases of ice in a wide $p-T$ region. Ice with these $p$ and $T$ parameters is apparently present in the ice satellites of Jupiter and Saturn, such as Ganymede, Callisto, Titan, etc. Information on the elastic properties of the various modifications of ice is of extremely great interest, since the velocities of elastic waves inside the planets of the solar system and their satellites, along with the mass and size of the heavenly body, are often almost the only available information on the physicomechanical properties.

It will be interesting in the future to investigate the elastic characteristics during the *lda*–*hda* transition as $p$ and $T$ are varied and also to study the phase transformations using single crystals of ice.

This work was carried out with the financial support of the Russian Fund for Fundamental Research (Grant 95-02-03677).

[1] O. Mishima, L. D. Calvert, and E. Whalley, Nature (London) **310**, 393 (1984).
[2] O. Mishima, L. D. Calvert, and E. Whalley, Nature (London) **314**, 76 (1985).
[3] O. Mishima, K. Takemura, and K. Aoke, Science **254**, 406 (1991).
[4] O. Mishima, J. Chem. Phys. **100**, 5910 (1994).
[5] E. G. Ponyatovsky and O. I. Barkalov, Mater. Sci. Rep. **8**, 147 (1992).
[6] M. A. Floriano, Y. P. Handa, D. D. Klug, and E. Whalley, J. Chem. Phys. **91**, 7187 (1989).
[7] R. J. Hemley, L. C. Chen, and H. K. Mao, Nature (London) **338**, 638 (1989).
[8] A. Bizid, L. Bosio, A. Defrain, and M. Oumerzzine, J. Chem. **87,** 2225 (1987).
[9] L. F. Baez and P. Clancy, J. Chem. Phys. **103**, 22 (1985).
[10] R. E. Gagnon, H. Kiefte, and M. J. Clouter, J. Chem. Phys. **89**, 4522 (1988).
[11] G. H. Shaw, J. Chem. Phys. **84**, 5862 (1986).
[12] O. V. Stal'gorova, E. L. Gromnitskaya, and V. V. Brazhkin, JETP Lett. **62**, 356 (1995).
[13] H. Shimizu, T. Nabeteni, T. Nishiba, and S. Sasaki, Phys. Rev. B **53**, 6107 (1996).
[14] H. Shimizu, M. Ohnishi, and S. Sasaki, Phys. Rev. Lett. **74**, 2820 (1995).
[15] C. A. Tulk, R. E. Gagnon, H. Kiefte, and M. J. Clouter, J. Chem. Phys. **101**, 2350 (1994).
[16] A. Polian and M. Grimsditch, Phys. Rev. B **27**, 6409 (1983).
[17] V. V. Brazhkin and A. G. Lyapin, High Press. Res. **15**, 9 (1996).
[18] J. Koike, Phys. Rev. B **47**, 7700 (1993).
[19] Ch. Ettl and K. Samyer, Mater. Sci. Eng. A **178**, 245 (1994).
[20] J. Schroeder, T. G. Bilodeau, and X.-S. Zhao, High Press. Res. **4**, 531 (1990).
[21] G. H. Wolf, S. Wang, C. A. Herbst *et al.*, in *High-Pressure Research: Application to Earth and Planetary Sciences*, ed. Y. Syono and M. H. Manghnani (Terra Sci. Pub. Co., Tokyo, 1992), p. 503.
[22] R. J. Hemley, L. C. Chen, and H. K. Mao, Nature (London) **338**, 638 (1989).
[23] M. Grimsditch, Phys. Rev. Lett. **52**, 2379 (1984).
[24] O. V. Stal'gorova, E. L. Gromnitskaya, D. R. Dmitriev, and F. F. Voronov, Prib. Tekh. Éksp. **39(6)**, 101 (1996).
[25] E. P. Papadakia, Rev. Sci. Instrum. **47**, 805 (1976).
[26] F. F. Voronov, O. V. Stal'gorova, and E. L. Gromnitskaya, Zh. Éksp. Teor. Fiz. **105**, 1456 (1994) [ JETP **78**, 785 (1994)].
[27] G. I. Peresada, Phys. Status Solidi A **4**, K23 (1971).
[28] S. H. Kirby, W. B. Durham, M. L. Beeman, H. C. Heard, and M. A. Daley, J. Phys., Coll. C1 **48**, 227 (1987).
[29] N. Soga, R. Ota, and M. Kunugi, in *Mechanical Behavior of Materials*, ed. S. Taira and M. Kunugi (Soc. Mat. Sci. Japan, Kyoto, 1974), vol. 4, p. 366.
[30] L. R. Testardi and J. J. Hauser, Solid State Commun. **21**, 1039 (1977).
[31] N. Soga, M. Kunugi, and R. Ota, J. Phys. Chem. Solids **34**, 2143 (1973).
[32] B. Golding, B. G. Bagley, and F. S. Hsu, Phys. Rev. Lett. **29**, 68 (1972).
[33] M.-C. Bellissent-Funel, J. Teixeira, L. Bosio *et al.*, in *Physics and Chemistry of Ice*, ed. N. Maeno (Hokkaido Univ. Press, Sapporo, 1992), p. 98.
[34] A. I. Kolesnikov, V. V. Sinitzyn, E. G. Ponyatovsky *et al.*, J. Phys.: Condens. Matter **6**, 375 (1994).
[35] E. G. Ponyatovskiĭ, V. V. Sinitsyn, T. A. Pozdnyakova, JETP Lett. **60**, 360 (1994).
[36] P. H. Pool, U. Essmann, F. Sciortino *et al.*, Phys. Rev. E **48**, 4605 (1993).
[37] H. E. Stanley, C. A. Angell, U. Essmann *et al.*, Physica A **205**, 122 (1994).

Translated by W. J. Manthey

# Influence of a nonmagnetic impurity on the properties of the quasi-one-dimensional antiferromagnet $CsNiCl_3$

S. S. Sosin, I. A. Zaliznyak, L. A. Prozorova, Yu. M. Tsipenyuk, and S. V. Petrov

*P. L. Kapitza Institute of Physics Problems, Russian Academy of Sciences, 117334 Moscow, Russia*
(Submitted 5 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 209–220 (July 1997)

Various magnetic properties of the diluted quasi-one-dimensional antiferromagnet $CsNi_{1-x}Mg_xCl_3$ are investigated experimentally for several impurity concentrations. The antiferromagnetic resonance spectrum and the phase diagrams are found to depend significantly on the amount of added Mg. The field and temperature dependences of the static magnetization is measured for crystals with two different contents $x$. A substantial increase in the magnetization is observed at low temperature, where the additional susceptibility is approximately proportional to the concentration. The physical mechanisms underlying the observed strong influence of magnetic defects formed at breaks in the spin chains in a quasi-one-dimensional antiferromagnet on its magnetic properties in the ordered state and for $T>T_N$ are discussed. © *1997 American Institute of Physics.* [S1063-7761(97)01807-6]

## 1. INTRODUCTION

Quasi-one-dimensional systems with antiferromagnetic exchange have been a topic of considerable interest in the last decade. The hypothesis advanced by Haldane regarding the existence of a gap in the spectrum of magnetic excitations of one-dimensional chains with integral spins[1] has stimulated a series of experimental studies to test the hypothesis, along with subsequent theoretical investigations. Among the latter is an exactly solvable model proposed by Affleck *et al.* for such a chain with additional biquadratic exchange $(J/3)\Sigma_i(\mathbf{S}_i\cdot\mathbf{S}_{i+1})^2$ ($J$ is the Heisenberg exchange constant).[2] The ''valence-bond solid'' (VBS) ground state of the system is constructed as follows. The wave function of each spin $S=1$ is represented by the symmetrized product of the wave functions of two spins $S=1/2$. So-called valence bonds are then introduced between neighboring spins $S=1$, i.e., the wave function of the system is antisymmetrized on the pairs of indices associated with the spins $S=1/2$ at adjacent nodes. This model has a singlet ground state with a short correlation length $1/\ln 3$ and a gap spectrum of excitations, i.e., it is an example of a Haldane system. However, despite the simplicity of the formal construction of the VBS state, its physical interpretation involving the segregation of the initial electron spins $S=1/2$ at each node is not a trivial matter, because the energy binding them into $S=1$ within one electron orbital is incomparably higher than the exchange interaction energy between neighboring ions.

All the same, the simplicity of the approach has made it a very attractive object for experimental testing. In particular, it has been noted that the fourfold degeneracy of the ground state of a finite chain of spins $S=1$ in this model can actually be represented by ''broken valence bonds'' at its ends.[1)] EPR studies[4,5] have revealed the presence of split paramagnetic lines in the Haldane antiferromagnet NENP lightly doped ($\sim 0.5\%$) with various impurities (preferably replacing Ni in chains with $S=1$), consistent with the presence of $S=1/2$ degrees of freedom at the ends. However, numerous investigations of the magnetization in diluted NENP and other known Ni-organic Haldane systems (NiNO and TMNIN) (Ref. 6), having confirmed the occurrence of a paramagnetic contribution that increases with the impurity concentration, have shown that the absolute value of this contribution is usually substantially smaller than the value calculated for the model of free spins $S=1/2$ at the ends of each chain. An investigation of the magnetic contribution to the heat capacity[7] of the Zn-doped inorganic Haldane antiferromagnet $Y_2BaNiO_5$ definitely refutes the model of free spins $S=1/2$ in favor of equally populated finite chains of triplet and singlet states.

The occurrence of exchange between chains in real quasi-one-dimensional compounds with $S=1$ can cause them to become ordered at a finite temperature. The breaking of chains by impurities in such an ordered system induces magnetic defects that are not free, but are weakly bound with other spins (which form Néel order on the average). It is natural to expect the magnetic moment associated with finite chains to be a potentially significant factor governing the process and properties of magnetic ordering. If the isolated chains of triplet and singlet states are assumed to be equally populated with breaks, the quantum case $S=1$ becomes analogous to the classical spin case treated in Ref. 8 on the basis of the theory of spin waves. In particular, ''quasi-one-dimensional enhancement'' of the influence of impurity concentration should be observed, producing large variations in the observed quantities far from the classical percolation limit.

Here we report an experimental study of the diamagnetically diluted compound $CsNiCl_3$, which at temperatures above the three-dimensional antiferromagnetic ordering temperature exhibits properties typical of a Haldane gap system with a singlet ground state. This case affords the possibility of investigating magnetic defects formed by finite chains with $S=1$ both in the free state and in a state associated with antiferromagnetic ordering of the system.

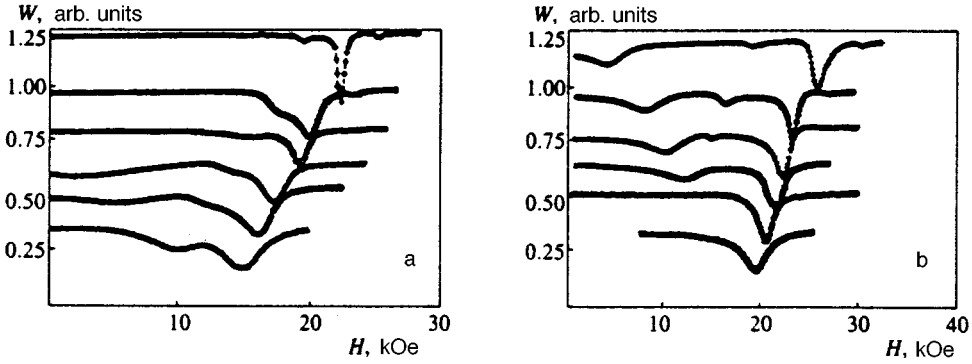



FIG. 1. Resonance absorption lines at frequencies of 41 GHz (a) and 57 GHz (b), $T=1.3$ K. Reading from the top down: $x=0.0$, 0.01, 0.018, 0.03, 0.044, 0.075.

## 2. SAMPLES AND MEASUREMENT PROCEDURE

One of the most important parts of the present study was the preparation of homogeneously diluted $CsNi_{1-x}Mg_xCl_3$ single crystals and the monitoring of their impurity content. The samples were prepared by the following procedure: The initial CsCl, $NiCl_2$, and $CsMgCl_3\cdot 6H_2O$ powders were mixed in the prescribed stoichiometric ratio and placed in a quartz ampoule. To dehydrate the carnallite for subsequent preparation, hydrogen chloride was passed through the ampoule for one hour as the mixture was gradually heated to the melting point (800 °C). The ampoule was then evacuated at a temperature of 150–200 °C, sealed off, and moved through a 760 °C furnace at the rate of 1 mm/h. The resulting crystals were annealed for 7–10 days at a temperature of approximately 400 °C. The samples were transparent, dark red in color, with dimensions of $\sim 1$ cm$^3$, and were easily drilled along the binary planes, facilitating their visual orientation. Spot radiographic examinations did not reveal, within the experimental error limits ($\sim 0.5\%$), any changes in the unit-cell parameters of the prepared crystals from the pure $CsNiCl_3$. It can therefore be assumed that $Mg^{2+}$ impurity ions replace $Ni^{2+}$ ions in the lattice, and the closeness of the lattice parameters of $CsNiCl_3$ and $CsMgCl_3$ ($a=7.17$ Å, $c=5.94$ Å and $a=7.27$ Å, $c=6.19$ Å, respectively[9]), together with the low concentration and uniform distribution of the impurity throughout the crystal precluded the possibility of detecting any variation of the lattice by this technique.

The magnesium concentration of the measured single crystals was tested by $\gamma$-activation analysis. The investigated crystals together with magnesium foil (as the standard) were irradiated on a microtron by 26-MeV electron bremsstrahlung for 10 min. The photonuclear $(\gamma,p)$ reaction at the $^{24}$Mg nuclei produces a radioactive $^{24}$Na isotope, whose emission spectrum contains a readily measurable $\gamma$-line with energy $E_\gamma=2753$ keV (half-life $T_{1/2}=15$ h). This emission was recorded by a high-resolution germanium semiconductor detector 15–20 h after irradiation (to ensure the decay of all other short-lived, background-forming nuclides). The magnesium content was determined according to the intensity of the recorded emission from the standard ($I_{\rm st}$) and the sample ($I_{\rm samp}$) with allowance for their respective masses $m^{\rm st}_{\rm Mg}$ and $m_{\rm Mg}$. The relative scatter of the concentrations determined from the results of several measurements is $\sim 10\%$. The mass content $\rho_{\rm Mg}$ was calculated from the equations

$$m_{\rm Mg}=\frac{I_{\rm samp}}{I_{\rm st}}m^{\rm st}_{\rm Mg},\quad \rho_{\rm Mg}=\frac{m_{\rm Mg}}{M},$$

where $M$ is the mass of the crystal, and the required atomic content $x$ was easily obtained from the results. The impurity concentration, reliably determined from the results of these measurements, was approximately half the initial stoichiometric value in every case, most likely on account of the specific characteristics of the fusibility diagram of the components.

The procedure used for the resonance measurements was simple and has been described many times. We used a set of open-ended, direct-amplification spectrometers with various types of absorbing cells designed for frequencies from 25 GHz to 80 GHz. The magnetic field was generated by a superconducting solenoid with scanning up to 65 kOe. The measurements at temperatures above 4.2 K were performed on instruments with a vacuum jacket with a heater and a semiconductor thermometer mounted in its interior. The lowest temperature obtained by helium vapor pumping was 1.3 K. The experimental errors for the various parameters were: temperature $\pm 0.05$ K; resonance field $\pm 0.1$ kOe; frequency $\pm 0.1$ GHz.

The static magnetization was measured on a standard Quantum Design SQUID magnetometer, which is capable of operating in fields up to 55 kOe in the temperature range 1.8–300 K and has negligible relative error at the amplitudes of our measured signals. Samples with masses of several tens of milligrams were drilled out of the large single crystals and were bonded to a nonmagnetic quartz glass holder by cyanolite. In all the measurements the samples were oriented visually within $\sim 1°$ error limits.

## 3. EXPERIMENTAL RESULTS

### 3.1. Antiferromagnetic Resonance Spectrum and Phase Diagram

The main part of the work consisted in the investigation of the antiferromagnetic resonance spectrum of eight $CsNiCl_3$ crystals with various Mg contents (from 0.5 at.% to 7.5 at.%) at the minimum temperature. All the absorption lines corresponding to different branches of the spectrum are found to be significantly shifted, even at the lowest impurity concentration. The situation thereafter continues to change

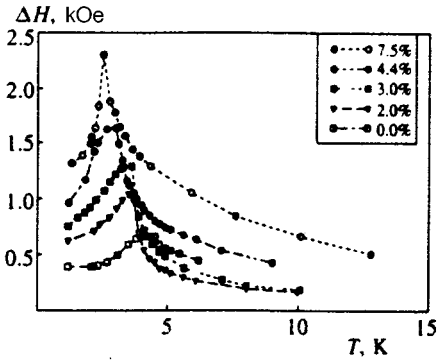



FIG. 2. Temperature dependence of the width of the electron spin resonance lines (for the $T<T_N$ branch of the antiferromagnetic resonance $\omega_1$ at $H>H_c$) for various impurity concentrations.

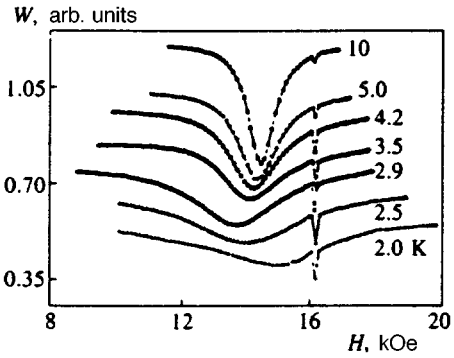



FIG. 4. Variation of the resonance line at the frequency $\omega=45$ GHz in a sample containing 4.4% impurity as the temperature is lowered from 10 K to 2.0 K (the sharp secondary resonances in a 16-kOe field are DPPH tracer marks.)

considerably until the resonance fields at $x=7.5\%$ differ by a factor of two or three from their initial levels.

Figures 1a and 1b show experimental records of the signal transmitted through the resonator containing different samples at two frequencies. The distinct resonance lines are easily traced, and their shifts can be estimated as the concentration $x$ is varied. Also obvious is an appreciable broadening of the lines, which for the most part have a regular Lorentzian profile, owing to the nature of the homogeneous vibrations in the spatially inhomogeneous system. Figure 2 shows temperature curves of the width (full width at half maximum) of the absorption line corresponding to the $\omega_1$ mode for $T<T_N$ (see Fig. 3) in crystals with various impurity contents. The measurement frequencies were chosen so that the derivative $\partial\omega/\partial H$ would exhibit the weakest possible temperature dependence in fields $H=H_{\rm res}$ (i.e., $H_{\rm res}$ is far from $H_c$). At the minimum temperature $T=1.3$ K $\Delta H$ depends linearly on the impurity concentration, the amount of its increase being much smaller than the variation of the corresponding resonance field. Consequently, an increase in $\Delta H$ does not imply the excitation of spatially inhomogeneous modes, but is more likely associated with an increase in the damping of long-wavelength vibrations due to microscopic inhomogeneity of the magnetic system.

We have used the measurement results to plot the antiferromagnetic resonance spectra over the entire accessible frequency range. Figure 3 shows several such spectra, which provide an example for tracing all the characteristic variations attending dilution. The principal effects are, first, a major reduction in the relativistic gap $\omega_1(0)$ (from 55 GHz in the pure crystal to 30 GHz at $x\simeq 0.075$) and, second, the shift of the spin-flop field $H_c$ (from 19.5 kOe to 10 kOe, respectively). Moreover, beginning with a concentration $\sim 3\%$, the antiferromagnetic resonance spectra acquire a field dependence, which cannot be described by the customary equations for a triangular antiferromagnet with easy-axis anisotropy.[10] In particular, the abrupt frequency jump of the $\omega_1(H)$ branch, corresponding to the spin-flop field for $H\|C_6$ (see the solid curve at the top of Fig. 3), spreads out over the field axis by as much as $\sim 2-3$ kOe, suggesting that the first-order phase transition has vanished. All these effects are discussed in detail in the next section.

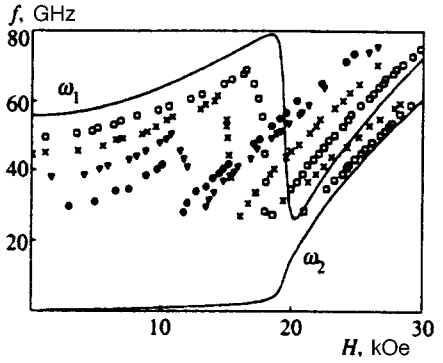



FIG. 3. Antiferromagnet resonance spectrum in $CsNi_{1-x}Mg_xCl_3$. Solid curves) $x=0$, plotted according to equations in Ref. 10; □) $x=0.01$; ×) 0.018; ▽) 0.044; ●) 0.075.

We concluded the antiferromagnetic resonance experiments by studying the temperature dependence of the resonance peak near the points of transition of the magnetic system to the three-dimensionally ordered state. These measurements were performed with a view toward disclosing possible anomalies of the diluted magnetic system in the critical region and investigating the dependence of the Néel temperature on the impurity concentration. A typical series of resonance absorption plots for a crystal with 4.4% impurity as the temperature passes through $T_N$ is shown in Fig. 4.

The shift of the resonance peak to the left as the temperature drops in the disordered state implies an increase in the effective $g$ factor as three-dimensional correlations grow stronger. After passing through $T_N$, the resonance becomes antiferromagnetic and, if the resonance field at the given frequency is greater than $H_c$, begins to shift to the right. The transition point can be determined fairly accurately by varying the temperature in 0.1 K intervals (the corresponding plots of the resonance field are shown in Fig. 5). Unfortunately, this is the only region of the phase diagram $H(T)$ ($H\|C_6$, $H>H_c$) that can be reliably investigated by this technique. In all other cases the $H_{\rm res}(T)$ curves do not exhibit abrupt changes at the points $T_N$, and the transitions have been determined from prominent features of the magnetization curves (see the next section).

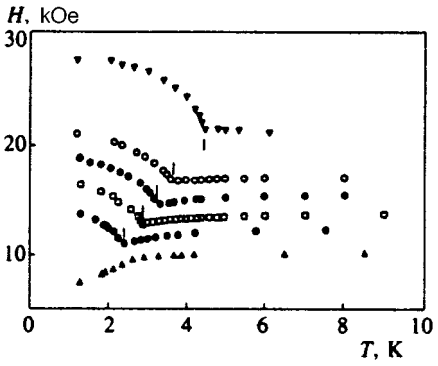



FIG. 5. Graphs of $H_{res}(T)$ for various impurity concentrations. Reading from the top down: $x=0.0$, 0.018, 0.03, 0.044, 0.075, 0.15. The tick marks indicate phase transition points.

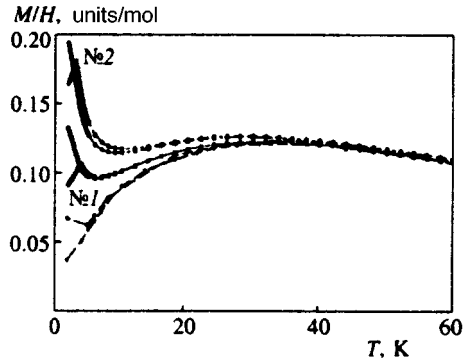



FIG. 7. Temperature dependence of $M/H$ in a field of 4 kOe for the pure crystal (from Ref. 12) and in samples No. *1* and No. *2* for $H\|C_6$ and $H\perp C_6$.

### 3.2. Measurement of the static magnetization

To clarify the nature of this phenomena, we have performed a series of measurements of the static magnetization on the SQUID magnetometer in samples with 3.3% (No. *1*) and 7% (No. *2*) impurity contents. We have plotted field curves of the magnetization $\mathbf{M}(H)$ in two principal orientations of the samples at the lowest possible temperature, 1.8 K, and we have investigated their temperature behavior in low fields.

It is evident from Fig. 6 that the crystals containing impurities acquire a large additional magnetization contribution, which is manifested in two characteristic properties. First, it is strongly nonlinear and reaches saturation in fields of 25–30 kOe; second, it has a certain anisotropy, so that the magnetic moment is several percent higher for easy-axis alignment of the field. It must be borne in mind that nonlinearity and anisotropy of opposite signs are observed in the undiluted system as a result of the zeroth vibrational mode.[11]

The temperature dependence of the magnetization at $H=4$ kOe for the same samples in two field orientations ($H\|C_6$ and $H\perp C_6$) is shown in Fig. 7. The magnetizing field must be chosen, on the one hand, so as to keep the sample in the single-domain state[2)] and, on the other, to avoid entry into a range where the process of paramagnetic saturation begins to assert itself. We can assume that the linear relation $M_\alpha=\chi_{\alpha\beta}H_\beta$ is not seriously violated anywhere down to the lowest temperature for our choice of field $H=4$ kOe, so that the dependence $M(T)$ is consistent with the more customary $\chi(T)$. It is evident from the figure that in both orientations the magnetization increment observed in the ordered phase now occurs at a temperature of 25–30 K, i.e., in the range of the onset of one-dimensional correlations in the chains. As the temperature is lowered further, it increases approximately as $1/T$ down to the Néel point, gradually becoming anisotropic in the presence of the easy axis. Upon entering the three-dimensionally ordered state, the susceptibility in the field $H\|C_6$ makes a sharp bend and begins to decrease, while in the field $H\perp C_6$ it passes through a removable singularity and continues to increase. Consequently, at $T<T_N$ the susceptibility tensor $\chi_{\alpha\beta}$ acquires the typical profile for a noncollinear exchange antiferromagnet, but the quantities $\chi_\perp$, $\chi_\|$, and $\chi_\perp/\chi_\|$ increase far above their initial values (i.e., the value in the pure system).[3)]

We have also performed a series of measurements of the temperature and field dependences of the magnetization in the vicinity of critical points for the purpose of plotting more complete phase diagrams of the investigated samples and comparing them with antiferromagnetic resonance data. It is evident from Fig. 8 that the plotted phase-transition lines determined from the kinks of the magnetization curves are in good agreement with the results of our antiferromagnetic resonance measurements. They also show that the intermediate temperature range $T_{N_2}<T<T_{N_1}$ broadens considerably

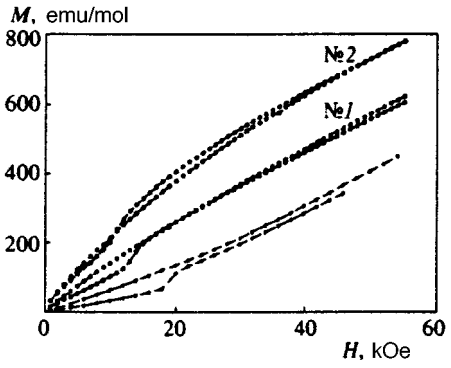



FIG. 6. Graphs of $M(H)$ in the pure crystal (from Ref. 11) and in samples No. *1* and No. *2* for $H\|C_6$ and $H\perp C_6$.

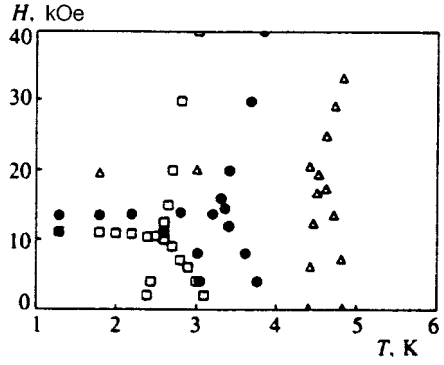



FIG. 8. Phase diagrams of the investigated samples for $H\|C_6$. ⊕) Sample No. *1*; □) No. *2*; △) pure sample (from Ref. 13); the dark symbols represent data of our antiferromagnetic resonance measurements.

as the concentration is raised (see Ref. 13). It has a width of 0.4 K in the pure sample and increases to 0.6 K for a 3% impurity and to 0.8 K for 7%, indicating an increase in the influence of anisotropy in the evolution of three-dimensional ordering. The results are qualitatively consistent with previously published[14] measurements of the phase diagram of a $CsNi_{0.98}Mg_{0.02}Cl_3$ crystal, but the effects observed in the latter was far less pronounced. This discrepancy can be attributed to the error in determining the added-impurity concentration in the crystal. In our experiments the impurity content of the investigated crystals has been accurately measured by $\gamma$-activation analysis.

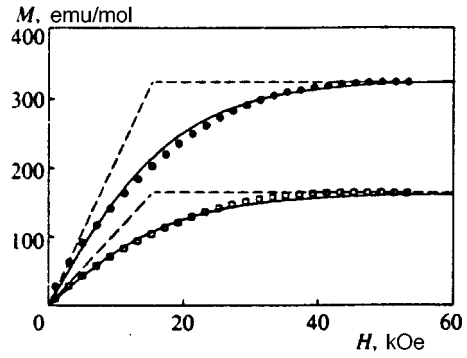



FIG. 9. Difference in the magnetizations of samples No. *1* (□), No. *2* (⊕), and the pure sample (for $H \perp C_6$ (see Fig. 5). Solid curves) Brillouin function for $S=1$, $T=1.8$ K; dashed lines) mean-field approximation at $T=0$.

#### 4. DISCUSSION AND MAIN CONCLUSIONS

Several conclusions regarding the influence of dilution on quasi-one-dimensional antiferromagnets can be drawn from the experimental data. The principal consideration here is the magnetic susceptibility of weakly interacting spin chains fragmented by nonmagnetic inclusions.

As discussed in the Introduction, without three-dimensional ordering (but with one-dimensional correlations, i.e., for $T_N \ll T \ll J$) the behavior of the ''breaks'' of the chains can be described either as additional $S=1/2$ degrees of freedom at the ends[4–6] or as paramagnetism of half the segments existing in the triplet state (like the case of classical spins, where it corresponds to paramagnetism of clusters of odd numbers of antiferromagnetically ordered spins[16]). The temperature dependence of the susceptibility of the chains must have the form $\chi = 2xC_{1/2}/T$ in the first case and $\chi = xC_1/2T$ in the second case, where $C_{1/2}$ and $C_1$ are the Curie constants for spins $S=1/2$ and $S=1$, respectively. Unfortunately, the required temperature interval is too narrow in $CsNiCl_3$ for the reliable experimental confirmation of either approach. All we know is that the second value of the constant coefficient of $1/T$ differs less than the first value from the experimental, consistent with the results of Ref. 7.

When three-dimensional antiferromagnetic ordering sets in, the additional degrees of freedom acquired by the magnetic system as a result of dilution are no longer independent, and all the effects observed in $CsNi_{1-x}Mg_xCl_3$ must begin to exhibit collective behavior. The magnetic properties of such a system have been analyzed previously[8] within the framework of the classical approximation of spin-wave theory at $T=0$. In Ref. 8 the calculated static susceptibility of a diluted antiferromagnet is given by the following expression in the first and second perturbation approximations:

$$\chi(x) \simeq \chi(0)\left(1 + \alpha x \sqrt{\frac{J}{J'}} + \beta x^2 \frac{J}{J'}\right), \tag{1}$$

where $\chi(0)$ is the susceptibility of the undiluted system, $\alpha$ and $\beta$ are numerical coefficients that depend on the configuration of the spins and the number of nearest neighbors, and $J'$ is the interchain interaction constant. Clearly, the role of the small perturbation parameter is not taken by the simple impurity concentration $x$, but by the quantity $x\sqrt{J/J'}$, which is qualitatively consistent with our results ($J/J' \sim 50$ for $CsNiCl_3$).

There is a fairly simple technique for qualitatively explaining the behavior of the magnetic defects at the breaks in the chains, based on the mean-field model. It follows from our experiment that the additional magnetization at saturation is exactly equal to $M_{\rm sat} = g\mu_B N_A x/2$, and the additional susceptibility in weak fields is approximately proportional to $x$. We can assume that in the ordered phase the parts of the chains between two impurities reside in a certain effective field $H_e'$, which is induced by interchain exchange interaction and does not depend on the concentration (see Fig. 9). In this case, when the field is applied along and perpendicular to the easy axis, the total susceptibility of the system at $T \approx 0$ is given by an equation of the form (taking anisotropy into account)

$$\chi_{\rm tot} = g\mu_B\left(\frac{1}{H_e \pm \lambda D} + \frac{x}{2}\frac{1}{H_e' \pm \lambda' D}\right), \tag{2}$$

where $H_e$ is the exchange field inside the chain, $D$ is the anisotropy constant (of the order of 0.3 kOe), and $\lambda$ and $\lambda'$ are numerical factors with values $\sim 1$. The factor $x/2$ in the second term implies that half the segments have positive spin $S=0$ and contribute nothing to the additional magnetization. This equation can also be used to explain the significant anisotropy of the additional susceptibility. For example, according to our experimental data, the susceptibility $\chi_{\rm tot}$ at an impurity concentration $x=0.03$ increases approximately 2.5-fold, so that $H_e' \sim 0.02 H_e \sim 15$ kOe. Consequently, the anisotropy constant, which can be disregarded in comparison with $H_e$, is found to be quite substantial in the second term of Eq. (2). For classical spins, comparing Eqs. (1) and (2), we can obtain the estimate $H_e' \sim \sqrt{JJ'}$, which is of the same order of magnitude as the mean field affecting one spin in the ordered state. It follows, therefore, that the additional susceptibility of diluted $CsNiCl_3$ can be identified with distortion of the triangular structure in a magnetic field.

The paramagnetic character of the $M(H)$ curves in Fig. 9 can be attributed to the high experimental temperature ($T_{\rm min} = 1.8$ K), which is comparable with the exchange field $H_e'$. As a result, the additional magnetization can undergo slow thermal fluctuations, which lower its equilibrium value in static measurements and do not affect it in measurements

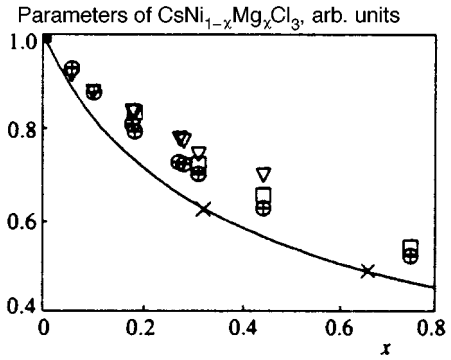



FIG. 10. Parameters of diluted $CsNiCl_3$ in arbitrary units versus impurity concentration. ⊕) $H_c$; ▽) $\omega_1(0)$; □) $T_N$; ×) $1/\sqrt{\chi_\parallel}$; the experimental error limits are contained within the size of the point symbols; the solid curve is plotted according to Eq. (1).

at high resonance frequencies. This phenomenon could be analogous to the superparamagnetism of small particles (see, e.g., Ref. 15).[4)]

The substantial lowering of the Néel temperature upon dilution of the initial system can also be attributed to its quasi-one-dimensionality. Indeed, when infinite one-dimensional chains are fragmented at lengths of $1/x$ in the quasiclassical approximation, their excited levels become discrete and separate from the main gap $\Delta \sim 2xJ$ (Ref. 16). Since this gap will necessarily be suppressed by three-dimensional correlations when exchange ordering sets in, the Néel temperature $T_N \sim 2\sqrt{JJ'}$ decreases by $\sim\Delta$, becoming equal to

$$T_N \sim T_{N_0} - \Delta \sim T_{N_0}\left(1 - 2x\frac{J}{J'}\right). \tag{3}$$

This approach yields fair agreement with our own experimental results up to impurity concentrations $\sim 3\%$, at which the observed shift of $T_N$ is approximately 1.2 K, and we have the estimate $\Delta \simeq 2 \cdot 0.03 \cdot 17 \simeq 1$ K.

The extraordinary resonance characteristics of diluted $CsNiCl_3$ can also be satisfactorily explained on the basis of the notion of the collective properties of the magnetic degrees of freedom in the ordered phase. All the principal parameters of the antiferromagnetic resonance spectrum and the phase diagram of crystals with various impurity concentrations are summarized in Fig. 10. Also shown as a visual aid in this figure are the values of the parallel susceptibility in coordinates $1/\sqrt{\chi_\parallel}$ and the corresponding theoretical curve from Ref. 8.

Assuming that the entire low-temperature magnetization of the antiferromagnet is associated with the ordered phase and is linear in the field, we readily obtain expressions for the critical spin-flop field and the upper relativistic gap of the antiferromagnetic resonance spectrum:[17]

$$H_c^2 = \frac{D}{\chi_\parallel - \chi_\perp}, \quad \left(\frac{\omega_1(0)}{\gamma}\right)^2 = \frac{D}{\chi_\perp}, \tag{4}$$

in which it is required to substitute the actual values of the susceptibility. Consequently, after the severalfold increase of $\chi_\parallel$ and $\chi_\perp$ relative to the initial system, $H_c$ and $\omega_1(0)$ should decrease significantly, as is indeed observed in experiment. The existence of nonlinear magnetization at $T \neq 0$ causes the spin-flop interval to broaden by as much as 2–3 kOe, beginning with a concentration $\sim 3\%$ (see Fig. 3). This effect is attributable to the fact that the potential energy of the exchange system acquires a complex field dependence, which does not pass through a minimum upon flopover of the spin plane at the point $H = H_c$. As a result, it is more likely that spin-flop does not take place at a point, but within a certain field interval, i.e., the first-order phase transition is replaced by two second-order transitions.

However, even when the variation of the total susceptibility of the system is taken into account, the observed field dependence of the antiferromagnetic resonance spectrum can be described only for the lowest impurity concentrations. This fact is mainly attributable to the increase in the effective measurement temperature $T_{\min}/T_N$, so the experimental conditions no longer satisfy the necessary condition $T \ll T_N$. A second cause is more fundamental and lies in the fact that the smallness of relativistic interactions relative to exchange interaction is a necessary condition in describing the long-wavelength spin dynamics of magnetic structures within the approach of Ref. 17. In pure $CsNiCl_3$ the degree of distortion of the exchange structure depends mainly on the ratio $D/J'$, and only in the vicinity of $H \sim H_e \sim 800$ kOe does the influence of the field begin to be felt. Upon dilution, as mentioned more than once, the magnetic structure acquires a new spin-flop degree of freedom, associated with the weak exchange field, where fields $H \sim H_e' \sim 15$ kOe are no longer small for this degree of freedom. Consequently, the range of applicability of the given approach is narrowed considerably, being restricted to low fields. A rigorous microscopic calculation based on the spin Hamiltonian with easy-axis anisotropy also appears to be nearly impossible, so that the question of describing the field dependences of the antiferromagnetic resonance spectrum is still open.

In summary, the following conception of the properties of a quasi-one-dimensional antiferromagnetic system with impurities is formed on the basis of the reported experimental data. When infinite chains are fragmented by nonmagnetic inclusions, new degrees of freedom are formed in them, half the segments acquiring an additional susceptibility to the magnetic field. In the temperature interval $T_N \ll T \ll J$ this susceptibility is paramagnetic and depends on the temperature according to the Curie law. When infinite three-dimensional correlations set in ($T < T_N$), all the chain segments are coupled together, acquiring the properties of a single integral magnetic system. As a result of quasi-one-dimensionality, however, the spin-flop degrees of freedom of the chain segments in a magnetic field continue to exert a major influence on the collective static and resonance properties. As before, the primary consequence of this influence is a significant increase in the static susceptibility, but now it is all associated with the ordered exchange system (i.e., is antiferromagnetic). This effect readily accounts for the pronounced decrease in the critical spin-flop field and the antiferromagnetic resonance gap. In the quasiclassical approach the quasi-one-dimensionality effect is manifested quantitatively by the emergence of the quantity $x\sqrt{J/J'}$ rather than

the simple impurity concentration $x$ as the small perturbation parameter of the initial system [see Eqs. (1) and (4)]. The same is true of the variation of the Néel temperature (3). We note that the foregoing analysis does not require any modification of the constants of the microscopic spin Hamiltonian; such a modification does not have adequate physical justification in the case of weak dilution.

In closing, the authors are grateful to A. F. Andreev, A. S. Borovik-Romanov, and A. I. Smirnov for valuable discussions. This work has received partial support from the Russian Fund for Fundamental Research (Project No. 95-02-04555-a) and the International Association for the Promotion of Cooperation with Scientists from the Independent States of the Former Soviet Union (INTAS Grant No. 94-968). S. Sosin would also like to thank the Forschungszentrum Jülich GmbH for financial support as part of the Landau Scholarship Program.

[1)]The fourfold degeneracy is attributable to the fact that the ground state of a chain comprising an even number of spins $S=1$ is a singlet, and for an odd number it is a triplet (see the theorem in Ref. 3). In the thermodynamic limit these states coincide, and the exponentially decaying (with distance) difference in their energies $E_1(N)-E_0(N)\sim(-1)^L\exp(-L/\xi)$ is less than 0.07 J for the chains of 14 or more spins discussed in this article.

[2)]In pure $CsNiCl_3$ the process of transition to a single-domain state associated with the rotation of all spins forming at $H=0$ different magnetic domains in the plane perpendicular to **H** is accompanied by nonlinear growth of the magnetization, but terminates in a field of 6 kOe.

[3)]In a planar, noncollinear antiferromagnet the direction of the field is more conveniently associated, not with the antiferromagnetism vectors situated in the plane of the spins, but with the vector perpendicular to this plane. This operation causes the symbols $\chi_\perp$ and $\chi_\parallel$ to change places (see Ref. 10).

[4)]The authors are grateful to A. N. Bazhan for calling their attention to this analogy.

[1] F. D. M. Haldane, Phys. Rev. Lett. **50**, 1153 (1983).
[2] I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Phys. Rev. Lett. **49**, 799 (1987).
[3] E. Lieb and D. Mattis, J. Math. Phys. (N.Y.) **3**, 749 (1962).
[4] S. H. Glarum, S. Geschwind, K. M. Lee *et al.*, Phys. Rev. Lett. **67**, 1614 (1991).
[5] M. Hagiwara, K. Katsumata, I. Affleck *et al.*, Phys. Rev. Lett. **65**, 3181 (1990).
[6] N. Fujiwara, J. R. Jeitler, C. Navas *et al.*, J. Magn. Magn. Mater. **140–144**, 1663 (1995); H. Deguchi, S. Takagi, M. Ito, and K. Takeda, J. Phys. Soc. Jpn. **64**, 22 (1995); H. Kikuchi, Y. Ajiro, N. Mori *et al.*, Physica B **201**, 186 (1994).
[7] A. P. Ramirez, S.-W. Cheong, and M. L. Kaplan, Phys. Rev. Lett. **72**, 3108 (1994).
[8] I. Ya. Korenblit and E. F. Shender, Phys. Rev. B **43**, 9478 (1993).
[9] J. Ackerman and E. M. Holt, J. Solid State Chem. **9**, 308 (1974).
[10] I. A. Zaliznyak, V. I. Marchenko, S. V. Petrov *et al.*, JETP Lett. **47**, 211 (1988).
[11] I. A. Zaliznyak, Solid State Commun. **34**, 573 (1992).
[12] N. Achiwa, J. Phys. Soc. Jpn. **27**, 561 (1969).
[13] M. Poirier, A. Caille, and M. L. Plumer, Phys. Rev. B **41**, 4869 (1990).
[14] Y. Trudeau, M. L. Plumer, M. Poirier, and J. Takeuchi, Phys. Rev. B **52**, 378 (1995).
[15] G. T. Rado and H. Suhl (eds.), *Magnetism*, Vol. 3, Academic Press, New York (1963).
[16] L. N. Bulaevskii, Solid State Phys. **11**, 1143 (1969).
[17] A. F. Andreev and V. I. Marchenko, Usp. Fiz. Nauk **130**, 39 (1980) [Sov. Phys. Usp. **23**, 21 (1980)].

Translated by James S. Wood

# Long-lived excited impurity states in diamond-like semiconductors

Ya. E. Pokrovskiĭ, O. I. Smirnova, and N. A. Khvalkovskiĭ

*Institute of Radioengineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia*
(Submitted 20 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 221–236 (July 1997)

The lifetime of charge carriers in the lowest excited states of some impurities of groups III and V in diamond, silicon, and germanium can be several (four to six) orders of magnitude longer that the lifetime of free carriers. Accumulation of carriers in these long-lived states may give rise to several new effects, such as hopping photoconductivity via long-lived excited states of impurities in dc and microwave electric fields, slow relaxation of induced absorption, and infrared absorption at energies lower than the impurity ionization energy. © *1997 American Institute of Physics.* [S1063-7761(97)01907-0]

## 1. INTRODUCTION

The fastest process which controls the rate of recombination of nonequilibrium charge carriers due to extrinsic excitation in doped and compensated diamond-like semiconductors is a capture at attractive impurity centers. Extensive investigation of capture at the centers in semiconductors has been carried out for more than thirty years. The physical nature of giant capture cross sections, which are several orders of magnitude larger than geometrical cross sections of ground states of such impurities, was explained by Lax.[1] The theory of cascade capture at attractive impurity centers developed by now[2] is based on Lax's concept and is in agreement with experimental data. The cascade model is based on the assumption that a carrier is captured not in the ground state of an impurity center, but in a highly excited state and then goes down the ''ladder'' of close energy levels of the impurity center through emission and absorption of acoustic phonons.

The model of cascade capture, however, does not take into account the discrete nature of the impurity spectrum and assumes existence of close energy levels with differences between them smaller than the energy of a characteristic phonon participating in these transitions. The bottleneck in this process is the capture of a free carrier in a highly excited state, and the relaxation rate of excited states is assumed to be relatively high. However, as a carrier goes down the ladder of excited levels, the energy difference between the levels increases. Therefore nonradiative transitions between the states of the impurity center can be slow, and the lifetime $\tau^*$ of carriers in such excited states may be considerably longer than the lifetime $\tau$ of free carriers. Such long-lived excited states of impurities should have considerable influence on relaxation of nonequilibrium charge carriers.

The long lifetime of excited impurity states in diamond-like semiconductors is due to their complex band structure. The lowest excited states of group V donors are $1S(\Gamma_5,\Gamma_3)$, whose configurations are determined by the valley–orbit splitting. In the case of acceptors in diamond and silicon, the lowest excited states are $1S(\Gamma_7^+)$ due to the spin–orbit coupling.[3] Owing to the equal parity, radiative dipole transitions from these excited states to the ground state are forbidden. The calculation of rates of the transitions with participation of phonons is a complicated problem, therefore experimental studies of these states are of fundamental importance.

This paper reports about a study of phenomena related to accumulation of nonequilibrium charge carriers in long-lived excited states of some impurities of groups III and V in diamond, silicon, and germanium.

## 2. DIAMOND

Investigation of photoelectric properties of synthetic diamonds doped with boron attracts attention because the solubility of boron in diamond is high, the boron impurity creates a relatively deep level with the ionization energy $E_i \simeq 370$ meV, and, given the strong interaction between holes and optical phonons, the response time of the impurity photoconductivity to an optical pulse can be shorter than a nanosecond.[4]

We have studied cubic single crystals of semiconducting diamond with volumes of about $10^{-5}$ cm$^{-3}$ synthesized under high pressure and temperature. The samples were doped with boron in the process of synthesis.[5] In the crystals selected for experiments the boron concentration $N_0$ was $10^{16}$–$10^{19}$ cm$^{-3}$, and the degree of compensation was 10–90%. Current contacts were applied to sides of the crystals. The electric conductivity of the samples in the temperature range of 90–370 K was controlled by two mechanisms: at higher temperatures by free holes in the valence band and at lower temperatures by hops between ground states of impurities. Smirnova and Gontar' [5] demonstrated that the hopping conductivity is an exponential function of the average distance between acceptors, and the Bohr radius of the ground state is $\sim 10$ Å.

We have measured the frequency, spectral, and temperature characteristics of impurity photoconductivity signal in dc electric field $U(\mathrm{DC})$, which is proportional to the change in the conductivity. At high temperatures, $T > 200$ K, the boron impurities are only partially ionized, and the conductivity is an exponential function of temperature with an activation energy $E_i = 370$ meV. In this temperature range and at low modulation frequencies of exciting light, $f < 10^4$ Hz, $U(\mathrm{DC})$ was controlled by the bolometric effect, whose time constant was $10^{-2}$–$10^{-3}$ s. By increasing the modulation frequency,

 1063-7761/97/070121-09$10.00  

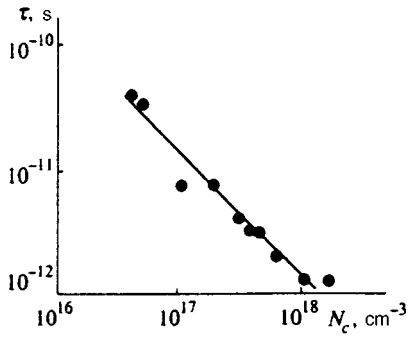



FIG. 1. Lifetime $\tau$ of holes versus concentration $N_c$ of compensating impurity in diamond doped with boron at 300 K.

we could get rid of the bolometric effect and study only the photoconductivity.[6] By comparing the quasistationary photoconductivity with the bolometric signal, we could determine $\mu\tau$, where $\mu$ is mobility. This allowed us to estimate the lifetime $\tau$ of free holes in diamond, which is small in the crystals studied ($\tau<10^{-10}$ s), and its determination by other methods is a complicated task. The lifetime $\tau$ is inversely proportional to the concentration of compensating impurities $N_c$ (Fig. 1), and the cross section of a capture at ionized boron atoms is $\sim 10^{-13}$ cm$^2$ at 300 K, in accordance with Ref. 2.

In lightly doped diamond samples ($N_0<10^{18}$ cm$^{-3}$) the impurity photoconductivity is due to free holes[7] and is a flat function of temperature (curves *1–3* in Fig. 2). But it turned out that in heavily doped ($N>10^{18}$ cm$^{-3}$) and compensated crystals the curves of $U$(DC) versus temperature and photon energy have anomalous shapes (curves *4–7* in Fig. 2), radically different from those of lightly doped samples. A quasistationary photoconductivity due to free holes was observed only at temperatures near the room temperature (section III of curves *4–7* in Fig. 2). The photoresponse $U$(DC) increased exponentially as the temperature decreased (section II in Fig. 2), and in heavily doped compensated semiconductors the signal increased by three orders of magnitude. The activation energy $E_i$ determined from the slope of section II was a periodic function of the exciting photon energy $\hbar\omega$, oscillating between 130 and 170 meV. The signal $U$(DC) also exponentially increased with decreasing temperature when the sample was excited by light with a photon energy corresponding to excitations within an impurity center ($\hbar\omega=347$ meV). The photoresponse magnitude in heavily doped samples was an oscillating function of the photon energy with a period corresponding to the optical phonon energy $\hbar\omega_0\sim 165$ meV.[8,9]

The time constants $\tau^*$ of the photoresponse were much longer than the lifetimes of free holes. At $T=90$ K and photon energies corresponding to the activation energy $E_i\simeq 130$ meV, the time constant $\tau^*$ was 40 ns, and at $E_i\simeq 170$ meV it was 200 ns (see Fig. 12 below), whereas the lifetime of free holes, according to our estimates, was shorter than $10^{-3}$ ns. The comparison between the curves of the photoresponse and dark conductivity versus temperature demonstrated that the photoconductivity increased exponentially with decreasing temperature in the range where the dark conductivity was controlled by hopping.

The increase in the time constant of photoconductivity and drastic changes in the spectral and temperature characteristics of the photoresponse with the increase in the doping level and compensation result from accumulation of carriers

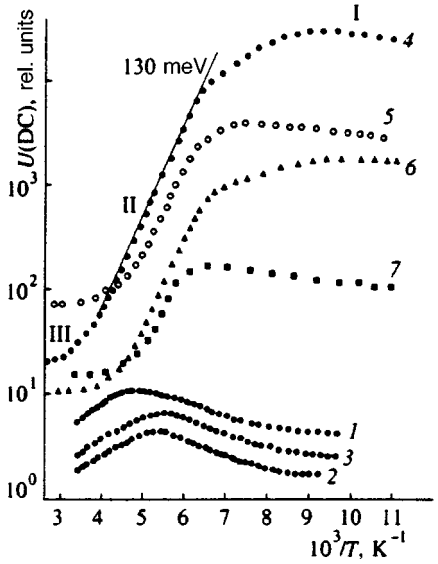



FIG. 2. Photoresponse signal $U$(DC) in boron-doped diamond versus temperature. Impurity concentrations in the samples are the following (cm$^{-3}$):

| Sample | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $N_0$ | $2.6\times10^{17}$ | $3.2\times10^{17}$ | $9.6\times10^{17}$ | $2.6\times10^{18}$ | $2.7\times10^{18}$ | $3\times10^{18}$ | $2.8\times10^{18}$ |
| $N_c$ | $2\times10^{17}$ | $9\times10^{16}$ | $4\times10^{16}$ | $2\times10^{18}$ | $1\times10^{18}$ | $1.2\times10^{18}$ | $4.3\times10^{17}$ |

in long-lived excited states of impurities. Holes created by light in the valence band or in higher bound states of impurities relax to lower excited states and increase their population. The change in population leads to an increase in the hopping conductivity via these deep states. This photoconductivity is determined by the concentration of holes in excited states and the hopping probability. This probability should be an exponential function of the Bohr radius,[10] and for excited states it should be much higher than for the ground state.

The photoconductivity via excited states can be higher than the photoconductivity due to free carriers only if the population of excited levels is very high, and greatly exceeds the concentration of holes in the valence band. This is possible if the lifetime $\tau^*$ in excited states is much longer than the lifetime $\tau$ of free carriers. The model of hopping photoconductivity, based on the assumption that carriers are accumulated in long-lived excited impurity states, accounts for anomalous curves of photoconductivity versus temperature in heavily doped samples.[8]

Indeed, at low temperatures, when thermal ejection of carriers from deep excited states is small and their population is determined by equilibrium between the capture of holes from the valence band or from higher excited states and relaxation to the ground state, the hopping conductivity via these deep levels is essentially independent of the temperature (sections I of curves *4–7* in Fig. 2). The population of long-lived excited states drops with temperature owing to the thermal ionization, and the photoresponse exponentially drops with temperature and it corresponds to the ionization energy of long-lived levels populated by photoexcitation (sections II of curves *4–7* in Fig. 2). A simple calculation yields the temperature $T^*$ of the onset of the exponential drop in the hopping photoconductivity as a function of the ionization energy $E_i$ of an excited level participating in the hopping conductivity and the ratio between the lifetimes in the valence band and in the excited state:

$$kT^* > \frac{E_i}{\ln(N_v(\tau^*/N_c)\tau)}, \tag{1}$$

where $N_v$ is the effective density of states in the valence band, and $N_c$ is the concentration of compensating impurities. At relatively high temperatures, the hopping conductivity is no longer the dominant process, and the photoresponse is determined by the conductivity due to free holes (sections III in Fig. 2).

The two activation energies of photoconductivity measured in experiments, which vary periodically with photon energy $\hbar\omega$, indicate that the hopping current is due to one of the two excited boron states with ionization energies of $\sim$130 meV and $\sim$170 meV, and lifetimes of 40 and 200 ns, respectively. The period of variation in $E_i$ is approximately $\hbar\omega_0$. This leads us to a conclusion that the capture of free holes in long-lived states occurs with emission of optical phonons. The spectral dependence of the photoresponse in heavily doped and compensated samples can be interpreted in terms of population of two excited long-lived states with different $\tau^*$. The temperature $T^*$ calculated by Eq. (1) for these two levels is in good agreement with experimental data.[8]

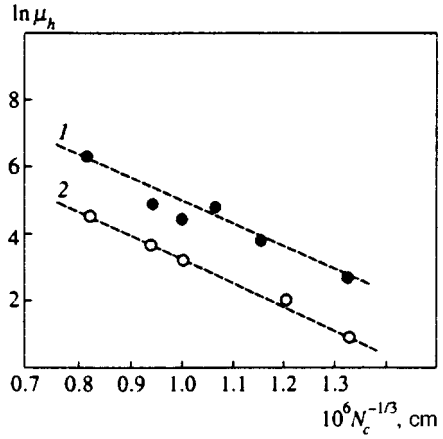



FIG. 3. Hopping mobility $\mu_h \sim U(\mathrm{DC})_\mathrm{I}/U(\mathrm{DC})_\mathrm{III}N_c$ via excited impurity levels in a diamond with (*1*) $E_i \simeq 130$ and (*2*) 170 meV versus concentration $N_c$ of compensating impurity at 90 K.

Carriers in excited states can hop to ionized majority impurities. The hopping probability can be sufficiently high only if the concentration of ionized centers, which is equal to the concentration $N_c$ of compensating impurities, is high. The hopping mobility $\mu_h$ should be an exponential function of the distance between ionized majority impurities. This function can be derived from experimental data by comparing the quasistationary responce $U(\mathrm{DC})_\mathrm{I}$ in the region of hopping photoconductivity (sections I in Fig. 2) and the response $U(\mathrm{DC})_\mathrm{III}$ (sections III in Fig. 2), which is due to free holes, in samples with various $N_c$.

In fact, $U(\mathrm{DC})_\mathrm{I} \sim \tau^*\mu_h$, and $U(\mathrm{DC})_\mathrm{III} \sim \mu/N_e$ (Fig. 1), hence $\mu_h \sim U(\mathrm{DC})_\mathrm{I}/U(\mathrm{DC})_\mathrm{III}N_c$. Figure 3 shows the curves corresponding to population of long-lived states with ionization energies of 130 and 170 meV. It is clear that $\mu_h$ is an exponential function of $N_c^{-1/3}$. Characteristic radii of excited states derived from the curves in Fig. 3 are close to 20 Å.

The states with ionization energies of 130 and 170 meV are much lower than the states of the $P_{3/2}$ series,[3] which are well known from studies of absorption and photoconductivity spectra of diamond doped with boron. A weak optical absorption in the range around 200 meV was detected previously only in heavily doped diamond,[11] which indicates that optical transitions between the ground state and these excited states are forbidden. This is consistent with the long lifetime of holes in these states. Both these facts indicate that the parities in the ground and excited states are the same, and the excited states are split from the ground state by spin–orbit coupling.

### 3. SILICON

We have studied silicon samples doped with impurities of groups III and V, such as B, Ga, In, As, P, Sb, and Bi in the concentrations $N_0 = 10^{15}$–$10^{18}$ cm$^{-3}$ in the process of floating zone melting. The ionization energies of these impurities range from 43 meV in Sb to 157 meV in In.[3] The compensating phosphorus impurity was introduced in *p*-type

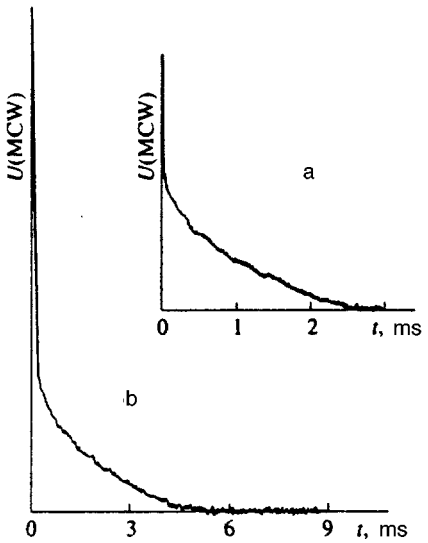



FIG. 4. Microwave photoresponse $U$(MCW) under pulsed excitation as a function of time at $T=5$ K in silicon doped with (a) boron ($N_0=2.8\times10^{16}$ cm$^{-3}$, $N_c=10^{14}$ cm$^{-3}$) and (b) indium ($N_0=5\times10^{16}$ cm$^{-3}$, $N_c=3\times10^{13}$ cm$^{-3}$).

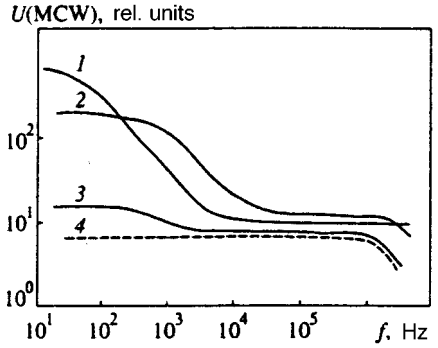



FIG. 5. Photoresponses (*1–3*) $U$(MCW) and (*4*) $U$(DC) versus modulation frequency $f$ of exciting radiation at 5 K in silicon doped with various impurities in concentrations $N_0$ and $N_c$ (cm$^{-3}$): (*1*) Sb ($2.7\times10^{16}$, $1.4\times10^{15}$); (*2*) In ($9\times10^{16}$, $10^{13}$); (*3, 4*) B ($2.8\times10^{16}$, $5\times10^{12}$).

samples in concentrations $N_c=10^{12}-10^{14}$ cm$^{-3}$ using neutron transmutation doping. In other samples, compensating impurities were introduced during zone crystallization. Contacts to samples were fabricated by ion implantation.

As was noted above, hopping photoconductivity was predominant in heavily doped diamond in the ranges of concentration and temperature where the conductivity was also due to hops. It should have been expected that similar effects manifest themselves in silicon at $N_0$, $N_c>10^{17}$ cm$^{-3}$. In fact, the relaxation of $U$(DC) excited by a $CO_2$ laser at 4.2 K in samples with such impurity concentrations was slow ($\sim10^{-2}$ s). But it is known that silicon with such high impurity concentrations is used as a material for cooled bolometers whose relaxation times are approximately the same. As will be shown below, the relaxation times of excited impurities in silicon are of the same order of magnitude, therefore the difference between relaxation times of photoconductivity and bolometric signal cannot be used in silicon, as in a doped diamond. At lower $N_0$ the relaxation time of $U$(DC) was $10^{-7}-10^{-10}$ s and was determined by the lifetime $\tau$ of free charge carriers.

The situation was completely different when the photoresponse was studied in doped silicon in a microwave (36–40 GHz) electric field $U$(MCW).[12]

Figure 4 shows curves of $U$(MCW) versus time for two silicon samples doped with In and B under pulsed excitation by CO and $CO_2$ lasers with a pulse duration $\Delta t\sim50$ $\mu$s. It is clear from Fig. 4 that the signal consists of two components, one of them is fast with a time constant shorter than the exciting pulse width 50 $\mu$s, and the other is slow with a time constant $\tau^*$ varying from sample to sample. The existence of the two components is also clearly seen in Fig. 5 from the magnitude–frequency characteristics for silicon samples doped with various impurities. We have established that the time constant $\tau^*$ of the slow component depends on the chemical nature of both majority and compensating impurities, but is almost constant with concentration for $N_0>10^{16}$ cm$^{-3}$ and temperature $T<15$ K.[13] The values of $\tau^*$ derived from pulsed and/or frequency measurements are presented in Table I. One can see that the slow relaxation occurs for all studied majority impurities, except Ga and Bi. In silicon doped with these impurities, $U$(MCW) had only the fast component due to relaxation of free charge carriers.

The temperature dependencies of the fast $U$(DC) and slow $U$(MCW) components of the photoconductivity for silicon samples doped with Sb under a quasistationary excitation are shown in Fig. 6. The figure demonstrates that, as in the case of diamond, the curves consist of three characteristic sections, namely, section III, where $U$(DC) and $U$(MCW) coincide, section II, where $U$(MCW) grows exponentially with the exponent determined by the activation energy $E_i\sim30$ meV, which is approximately equal to the ionization energy of excited levels $1S(\Gamma_3,\Gamma_5)$ of Sb, and section I, on which $U$(MCW) is two or three orders of magnitude higher

TABLE I. Time constant $\tau^*$ (ms) of slow relaxation of $U$(MCW) in silicon doped with various majority ($N_0$) and compensating ($N_c$) impurities

| $N_0/N_c$ | B | Ga | In | P | As | Sb | Bi |
|---|---|---|---|---|---|---|---|
| B | - | - | - | 0.5 | - | - | 1.6 |
| Ga | - | - | - | $<10^{-5}$ | - | - | - |
| In | - | - | - | 2 | - | - | - |
| P | 0.7 | - | - | - | - | - | - |
| As | 0.6 | - | - | - | - | - | - |
| Sb | 0.8 | 4 | 8 | - | - | - | - |
| Bi | $<10^{-5}$ | - | - | - | - | - | - |

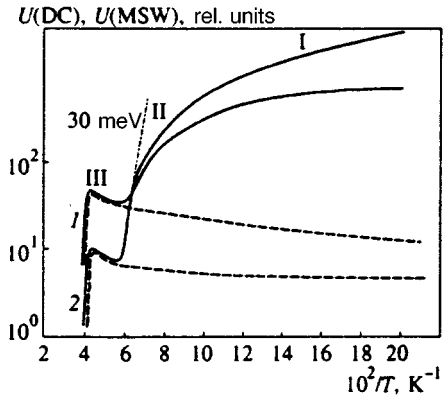



FIG. 6. $U$(DC) (dashed lines) and $U$(MCW) (solid lines) as functions of temperature under quasistationary excitation of silicon doped with Sb and compensated with B in concentrations $N_0$ and $N_c$ (cm$^{-3}$): (*1*) $2.7\times10^{16}$, $1.4\times10^{15}$; (*2*) $10^{17}$, $2\times10^{16}$.

than $U$(DC). Similar curves were obtained for other impurities of groups III and V in silicon.[13,14] The temperature interval I, where the slow component predominates over the conductivity due to free carriers, is determined by both the impurity nature and its concentration, and also by the compensating impurity concentration. The difference between the magnitudes of the fast and slow components was the larger, the higher the concentrations of both majority and compensating impurity. At low temperatures, this difference could be up to three orders of magnitude. The magnitude of the slow component is saturated with increasing excitation rate, and at high excitation intensity the fast component was dominant.[14]

Figure 7 shows the dependencies of $U$(MCW) on magnetic induction $B$ for a silicon sample doped with boron at various excitation rates $G$ and modulation frequencies $f$ of exciting light. One can see that at 5 K, when the fast photoconductivity component predominates (high $G$ or $f$), the curves correspond to the familiar transverse magnetoresistance effect with a free hole mobility $\sim10^4$ cm$^2$/V·s. But at lower $G$ and $f$ the photoresponse $U$(MCW) slowly increases with $B$, which cannot be explained in the terms of free-hole photoconductivity.

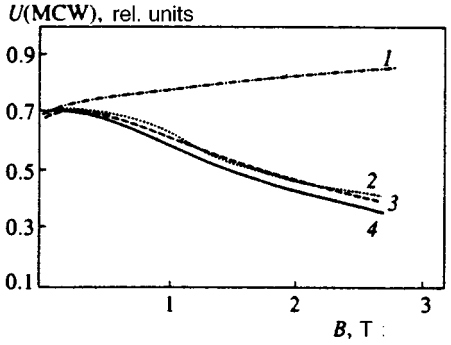



FIG. 7. $U$(MCW) at 5 K versus magnetic field induction $B$ in silicon doped with boron ($N_0=3.5\cdot10^{16}$, $N_c=10^{13}$ cm$^{-3}$) at the excitation modulation frequencies $f$ (*1, 2*) $1.2\times10^3$ Hz; (*3, 4*) $1.8\times10^6$ Hz and excitation rates $G$ (*1, 3*) $G_1$; (*2, 4*) $G_2$ ($G_2/G_1=20$).

The set of presented experimental results indicate that the slow component in $U$(MCW) is due to accumulation of photoinduced carriers in long-lived excited impurity states. The main tendencies in diamond and silicon are similar, but in silicon they manifest themselves in a microwave electric field at lower concentrations and degrees of compensation. Direct participation of long-lived impurity states is confirmed by the absence of slow relaxation of $U$(MCW) in the case of Ga and Bi impurities. Fast intracenter relaxation of nonequilibrium carriers at these impurities can be due to emission of optical phonons.[3] The energies of excited states of B, In, P, Sb, and As are beyond the band of optical phonons, and the carriers can accumulate in long-lived excited states during the process of relaxation.

The quantitative interpretation of experimental data is based on the model of polarization photoconductivity in microwave electric field[15] based on the model of equilibrium hopping conductivity in low-frequency electric field.[16] The polarization conductivity in doped compensated semiconductors is due to the changes in the dipole moment of pairs of ionized majority and compensating impurity atoms induced by a varying electric field. It can be several orders of magnitude higher than the dc conductivity. Under impurity excitation, the hops of carriers between excited and ionized majority impurities contribute to the polarization conductivity. Since the Bohr radius $a^*$ of the excited state is greater than the radius $a$ of the impurity ground state, the polarization conductivity should be greater, and the problem reduces to calculation of the probability of such hops. In our experiments, the electric field frequency was more than six orders higher than in Ref. 16. At this frequency, only the nearest pairs can contribute to the polarization conductivity. Let us consider the situation in $p$-type semiconductors.

Carriers are accumulated in long-lived excited states as a result of capture at both isolated majority impurities and dipoles (pairs of ions of majority and compensating impurities). Processes of photoionization and capture are random, therefore the probability of creating a close pair of an excited impurity state and an isolated ionized majority impurity is low. Excited atoms captured by a dipole are localized near compensating impurity ions, thus a positively charged pair is formed. In this case, the distance between the excited and ionized majority impurities is also fairly large immediately after the capture. However, electrostatic attraction between isolated negatively charged majority impurities and positively charged pairs cause a ''drift'' of the negative charge towards the positively charged complex. The drift is due to hops of carriers among majority impurities. If the drift time is not considerably longer than the lifetime of excited impurities, complexes containing an excited and an ionized majority impurity and a compensating impurity ion are formed. A calculation[15] yields the concentration $N_{*+-}$ of such triplets under stationary excitation at temperatures when thermal ionization of excited impurities can be neglected:

$$N_{*+-}=CG\tau^*/(1+\tau_M/\tau^*), \tag{2}$$

and under pulsed excitation at a pulse width $\Delta t<\tau_M$

$$N_{*+-}=CG\Delta t\ \exp(-t/\tau^*)[1-\exp(-t/\tau_M)]. \tag{3}$$

This concentration is determined by

(a) the fraction of the recombination current captured at dipoles:

$$C=\frac{\alpha(+-)N(+-)}{\alpha(+-)N(+-)+\alpha(-)N(-)}, \tag{4}$$

where $N(-)$ and $N(+-)$ are concentrations of isolated and dipole traps, respectively, and $\alpha(-)$ and $\alpha(+-)$ are the corresponding capture coefficients;

(b) hopping conductivity $\sigma_h$ in a dc electric field via ground states of impurities:

$$\tau_M=\varepsilon/4\pi\sigma_h\sim\exp(\beta r_s/a), \tag{5}$$

where $r_s=(4\pi N_0/3)^{-1/3}$, $\beta$ is the ''percolation'' parameter,[10] and $\varepsilon$ is the dielectric permittivity;

(c) lifetime $\tau^*$ of carriers in the excited state and the photoexcitation rate $G$.

The polarization hopping photoconductivity is determined by the equation[15]

$$\sigma(\mathrm{MCW})=\frac{N_{*+-}}{2r_s^2}\,\frac{\varepsilon}{5\cdot10^{-13}}\left(\frac{a^*}{2}\right)^5\frac{\Gamma(13/2)}{2^{13/2}}. \tag{6}$$

Equation (6) indicates that the concentration of triplets $N_{*+-}$ given by Eqs. (2) and (3) determines $\sigma$(MCW) as a function of the excitation rate, concentrations of the majority and compensating impurities, and temperature.

The model has been tested by taking a large set of silicon samples doped with boron and compensated with phosphorus under quasistationary photoexcitation.[14] We have determined that the slow component of $U$(MCW) grows linearly with the excitation rate $G$ and saturates when the concentration of excited atoms is higher than that of compensating impurities ($G\tau^*>N_c$). This means that the concentration of triplets, which controls the hopping photoconductivity, is maximal at such excitation level, since an ionized and an excited majority impurity are localized near any compensating impurity atom, and a further increase in the excitation level should not lead to an increase in $\sigma$(MCW).

The dependence of $N_{*+-}$ on the compensating impurity concentration is determined by the fraction $C$ of nonequilibrium charge carriers captured by dipoles given by Eq. (4). Moreover, $C(N_c)$ can be derived from the temperature dependencies of $U$(DC), which is due to free holes, measured in the samples with various $N_c$.[15] The measured values of $U$(MCW) as a function of $C$ and $N_c$ are plotted in Fig. 8.

The rate of creation of active triple impurity complexes is determined by $\sigma_h$ and is an exponential function of the majority impurity concentration $N_0$. In a set of samples with $N_c=10^{14}$ cm$^{-3}$ and variable $N_0$, the magnitude of the slow component of $U$(MCW) is a flat function of the majority impurity concentration at low temperature (5 K) if $N_0>2.7\cdot10^{16}$ cm$^{-3}$, and at lower $N_0$ it exponentially drops with the average distance between majority impurities (Fig. 9). This dependence results from Eqs. (2) and (5) and is determined by the ratio between $\tau^*$ and $\tau_M$.

As the temperature increases, $U$(MCW) gradually drops owing to thermal destruction of triplets,[15] then it decreases exponentially in the range $T>15$ K because of thermal ionization of excited states (Fig. 6). The activation energies $E_i$ obtained from the shapes of the exponential sections are 23 meV[14] and 30 meV (Fig. 6) for B and Sb, which corresponds to ionization energies of the lowest excited states of the impurities.[3]

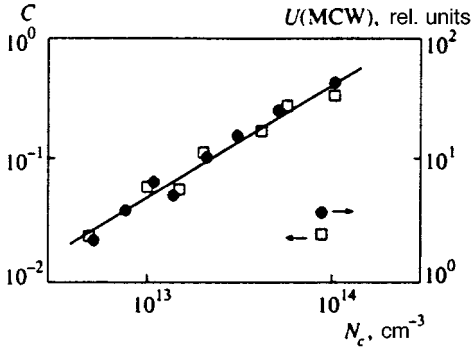



FIG. 8. $U$(MCW) and parameter $C$ defined by Eq. (4) versus phosphorus concentration $N_c$ in silicon doped with boron ($N_0=3.5\times10^{16}$ cm$^{-3}$) at 5 K.

As was noted above, the lifetime $\tau^*$ of excited states depends on the chemical nature not only of the majority impurity, but also of the compensating impurity (Table I). This may be due to participation of active triplets in the process of relaxation. Indeed, the presence of a heavy compensating-impurity ion near an excited majority-impurity atom can perturb the local phonon spectrum and thus change the relaxation time $\tau^*$.

When the long-lived excited states are populated, additional absorption due to ionization of the excited state should appear in absorption spectra of the samples. In spectral measurements, however, background radiation corresponding to the room temperature and emitted by warm cryostat components is always present. The spectrum of background radiation has a maximum close to the maximum of photoconductivity of silicon doped with most of group III and V impurities. Thus, the long-lived excited states of some majority impurities should be populated by the background radiation. In fact, in studies of absorption in the presence of the background radiation, we detected absorption in the IR band at energies lower than the ionization energy of the impurity ground state.[13] At the same time, the background is a pow-

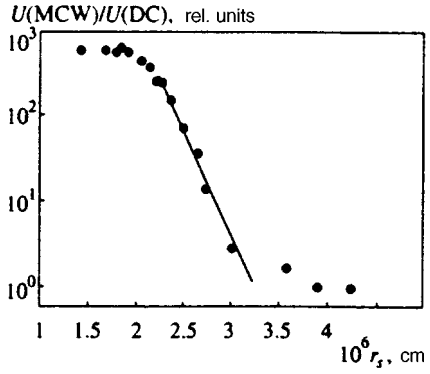



FIG. 9. $U$(MCW)/$U$(DC) versus $r_s=(4\pi N_0/3)^{-1/3}$ in silicon doped with boron in various concentrations $N_0$ and with phosphorus ($N_c=10^{14}$ cm$^{-3}$) at 5 K.

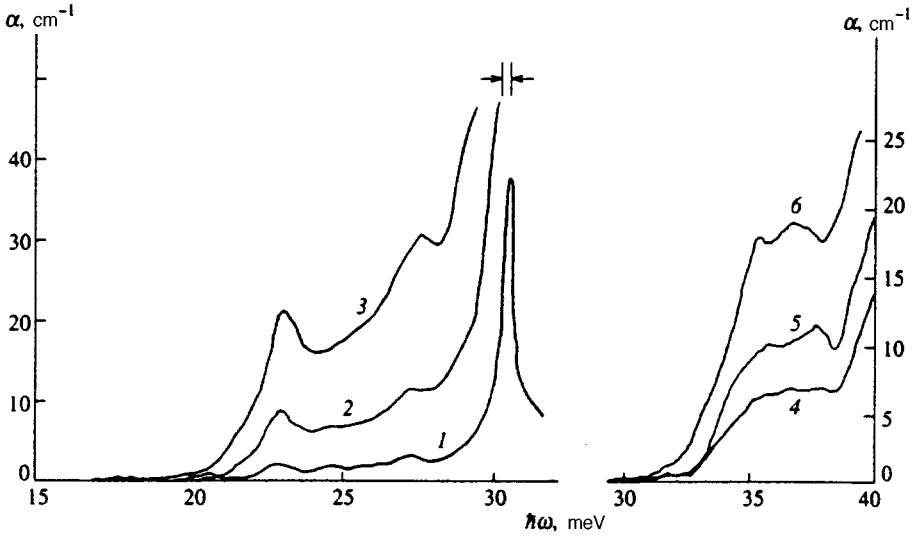



FIG. 10. Absorption spectra of silicon doped in concentrations $N_0$ ($cm^{-3}$) with boron: (1) $3.5\times10^{16}$; (2) $7.6\times10^{16}$; (3) $1.5\times10^{17}$, and arsenic: (4) $12\times10^{17}$; (5) $18\times10^{17}$; (6) $20\times10^{17}$ at 4.2 K.

erful source of IR radiation, whose absorption is a function of the population of long-lived excited impurity states. Modulated impurity excitation of silicon samples leads to modulation of the population of the long-lived excited impurity states, thereby modulating the absorption of background radiation by excited impurities. This induced absorption was detected when the silicon samples were exposed to modulated laser beam.[17] In order to detect a change in the background absorption, a photoresistor screened by an optical filter absorbing laser light was placed downstream of the sample. In order to separate a certain spectral band, we could place optical filters both downstream and upstream of the sample. In fact, a photoresistor signal corresponding to an increase in the silicon absorption was detected when a doped silicon sample was exposed to impurity excitation.[13] Moreover, typical relaxation times $\tau^*$ of induced absorption were close to the decay times of $U$(MCW) in the same samples. Using various optical filters and photoresistors, we have found that the induced absorption of the background radiation in samples is localized in the range of 20–40 meV.

Absorption spectra of silicon samples doped with boron and arsenic under background radiation are given in Fig. 10. The spectra have step-like shapes. The long-wave absorption edge in the sample doped with arsenic is at 32 meV and corresponds to the ionization energy of the lowest excited states $1S(\Gamma_5,\Gamma_3)$ due to the valley–orbit splitting.[3] In the case of the boron impurity, the long-wave absorption edge at 23 meV corresponds to either ionization of the excited state $1S(\Gamma_7^+)$ or excitation of a hole from the ground state to this excited state.[18] Taking into account that the $1S(\Gamma_7^+)$ state is singlet, whereas the ground state is degenerate and should be split by deformation,[3] we recorded absorption spectra of uniaxially compressed samples. Figure 11 shows such spectra under a small strain directed along the [110] axis. One can see that the long-wave absorption edge shifts to higher energies but does not split, and the absorption intensity is almost constant. This indicates that the absorption by boron impurities is due to ionization of the lowest excited state populated by the background radiation.

In silicon doped with phosphorus and antimony, only strong absorption lines due to transitions from the impurity ground state to excited states of the $P_{3/2}$ series were detected in the 10–40-meV band. The ionization energies of the lowest excited states of antimony and phosphorus ($E_i\sim30$ meV) are also in the range of these transitions, and absorption due to transitions from the excited states to the conduction band cannot be selected because of strong $P_{3/2}$ absorption. Nonetheless, the lack of additional absorption in the range of energies between 10 and 20 meV confirms our conclusion that absorption due to transitions between the ground state and lowest excited states of these donors is negligible.

Thus, we have established that some impurities of groups III and V in silicon have long-lived ($\tau^*\sim0.5-8$ ms) excited states. They manifest themselves in the hopping photoconductivity in microwave electric field, spectra and kinetics of induced IR absorption.

#### 4. GERMANIUM

Up to the present time, few studies of long-lived excited impurity states in germanium have been published. Calculations for acceptor states[19] indicate that the rates of transitions between highly excited states are one or two orders of magnitude higher than the rate of the phonon-assisted transition from the $1\Gamma_8^-$ excited state to the $1\Gamma_8^+$ ground state. The

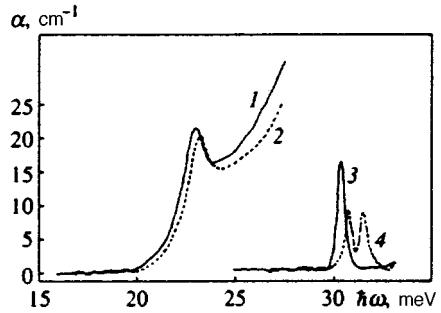



FIG. 11. Absorption spectra of silicon at 4.2 K doped with boron in concentrations $N_0$: (*1, 2*) $1.5\times10^{17}$ $cm^{-3}$; (*3, 4*) $7\times10^{15}$ $cm^{-3}$ and compressed along the [110] axis by pressure $P$: (*1, 3*) 0; (*2, 4*) 44 MPa.

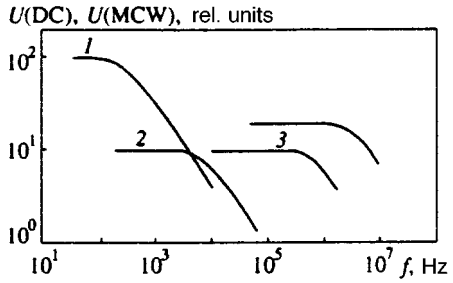



FIG. 12. Photoresponse of diamond-like semiconductors versus modulation frequency $f$ (Hz) of exciting radiation: (*1*) $U$(MCW) in silicon with boron at $T=5$ K; (*2*) $U$(MCW) in germanium with arsenic at 2.2 K; (*3*) $U$(DC) in diamond with boron at 90 K.

lifetime of the $1\Gamma_8^-$ excited state ($\sim 10^{-7}$ s) is in good agreement with measurements of magnetic impurity oscillations[20] and far-infrared photoconductivity[21] in $p$-Ge. Lifetimes of the $2p_0$ excited states of donors in germanium are essentially shorter ($\sim 10^{-9}$ s),[22] and in experimental conditions[21] were no longer than lifetimes of free electrons.

Nonetheless, conditions can be created when a deep long-lived excited state manifests itself in germanium.[23] Unlike the case of diamond and silicon, the spin–orbit splitting in germanium is large (290 meV); therefore the $1S(\Gamma_7^+)$ states of the acceptors excited due to the spin–orbit splitting should be in the valence band[3] and play no role in relaxation of charge carriers. But in the case of donors, the situations in silicon and germanium are similar, so one can expect that the $1S(\Gamma_5)$ donor state in germanium should be long-lived.

The arsenic impurity in germanium is most suitable for investigating the hopping photoconductivity since its ionization energy of 14.18 meV and valley–orbit splitting of 4.24 meV are the highest.[3] The magnitude of the hopping photoconductivity in microwave electric field strongly depends on the ratio between the Bohr radii of the long-lived and ground states of the impurity. Since the ionization energies of the $\Gamma_5$ excited states of group V donors in germanium are the same ($\sim 10$ meV),[3] we prefer the donor with the minimal Bohr radius of the ground state, i.e., with the maximum ionization energy.

We have measured the photoresponse in germanium samples doped with As in concentrations of $1.6\times 10^{14}$ cm$^{-3}$ and $10^{16}$ cm$^{-3}$ and containing compensating acceptors in concentrations of $\sim 10^{13}$ cm$^{-3}$. The photoresponses under dc and microwave electric fields was studied as a function of the modulation frequency $f$, excitation intensity, and temperature in the range between 2.2 and 15 K. At the higher doping level, a slow ($\tau^* \sim 30$ $\mu$s) component of $U$(MCW) appeared (Fig. 12). At 2.2 K, the magnitude of the slow component was almost two orders higher than that of $U$(DC), which is due to free electrons (Fig. 13), and the decay time $\tau^*$ was at least four orders longer than the lifetime of free electrons. As in silicon, $U$(MCW) saturated at some excitation intensity and dropped with temperature. These results are in qualitative agreement with the model of hopping polarization photoconductivity.[15]

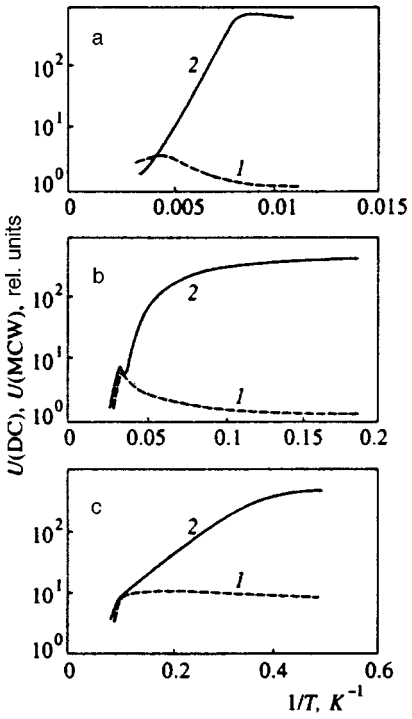



FIG. 13. Temperature dependencies of quasistationary photoresponses in diamond-like semiconductors doped in concentrations $N_0$: (a) $U$(DC) in diamond with boron: (*1*) $N_0=5\times 10^{17}$ cm$^{-3}$; (*2*) $N_0=3\times 10^{18}$ cm$^{-3}$; (b) silicon with boron, $N_0=3.5\times 10^{16}$ cm$^{-3}$: (*1*) $U$(DC); (*2*) $U$(MCW); (c) germanium with arsenic, $N_0=10^{16}$ cm$^{-3}$: (*1*) $U$(DC); (*2*) $U$(MCW).

## 5. CONCLUSION

In conclusion, we note that deep long-lived excited impurity states in diamond-like semiconductors manifest themselves in diamond, silicon, and germanium in different ranges of impurity concentration and temperature (Figs. 12 and 13). This is due, above all, to the various scales of energies of impurity levels, features of their band structure, and energy spectra of their impurities. The existing long-lived excited states are due to the valley–orbit or spin–orbit splitting of impurity ground states. The existence of the deep long-lived excited states, which affect nonequilibrium electronic processes, can be expected in other semiconductors with similar band structures. In fact, the long-lived ($\tau^* \sim 7$ ms) $1S(\Gamma_5)$ excited state of the tellurium donor split from the ground state by valley–orbit interaction has been observed recently in gallium phosphide.[24]

The work was supported by the Russian Fund for Fundamental Research (Projects Nos. 93-2-2070 and 96-2-16243).

[1] M. Lax, Phys. Rev. **119**, 1502 (1960).
[2] V. N. Abakumov, V. I. Perel', and I. N. Yassievich, Fiz. Tekh. Poluprovodn. **12**, 3 (1978) [Sov. Phys. Semicond. **12**, 1 (1978)].
[3] A. K. Ramdas and S. Rodrigues, Rep. Prog. Phys. **44**, 1287 (1981).
[4] *The Properties of Natural and Synthetic Diamond*, ed. by E. Field, Academic Press, London (1992).
[5] O. I. Smirnova and A. G. Gontar', Sverkhtverdye Materialy, No. 1, 12 (1993).
[6] O. I. Smirnova, É. É. Godik, and A. G. Gontar', Fiz. Tekh. Poluprovodn. 19, 646 (1985) [Sov. Phys. Semicond. **19**, 398 (1985)].
[7] O. I. Smirnova, É. É. Godik, and A. G. Gontar', Fiz. Tekh. Poluprovodn.

20, 869 (1986) [Sov. Phys. Semicond. **20**, 547 (1986)].
[8] O. I. Smirnova, É. É. Godik, and A. G. Gontar', Fiz. Tekh. Poluprovodn. 21, 1278 (1987) [Sov. Phys. Semicond. **21**, 774 (1987)].
[9] O. I. Smirnova and A. G. Gontar, Solid State Commun. **93**, 458 (1995).
[10] B. I. Shklovskii and A. A. Éfros, *Electronic Properties of Doped Semiconductors* [in Russian], Nauka, Moscow (1979).
[11] E. C. Lightowlers, P. J. Dean, and W. F. Sherman, Phys. Rev. **140**, 1272 (1965).
[12] Ya. E. Pokrovskii and O. I. Smirnova, JETP Lett. **51**, 429 (1990).
[13] Ya. E. Pokrovskii and O. I. Smirnova, Solid State Commun. **93**, 405 (1995).
[14] Ya. E. Pokrovskii and O. I. Smirnova, Zh. Éksp. Teor. Fiz. **102**, 660 (1992) [Sov. Phys. JETP **75,** 353 (1992)].
[15] Ya. E. Pokrovskii and O. I. Smirnova, Zh. Éksp. Teor. Fiz. 103, 1411 (1993) [JETP **76**, 690 (1993)].
[16] M. Pollak and T. M. Geballe, Phys. Rev. **122**, 1742 (1961).
[17] Ya. E. Pokrovskii and O. I. Smirnova, JETP Lett. **54**, 92 (1991).
[18] H. R. Chandrasekhar, A. K. Ramdas, and S. Rodrigues, Phys. Rev. B **12**, 5780 (1975).
[19] C. V. Meshkov and É. I. Rashba, Zh. Éksp. Teor. Fiz. 76, 2206 (1979) [Sov. Phys. JETP **49**, 1115 (1979)].
[20] V. F. Gantmakher, V. N. Zverev, S. V. Meshkov *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **42**, 1160 (1978).
[21] E. M. Gershenzon, G. N. Gol'tsman, and N. G. Ptitsyna, Zh. Éksp. Teor. Fiz. 76, 711 (1979) [Sov. Phys. JETP **49**, 355 (1979)].
[22] S. V. Meshkov, Fiz. Tverd. Tela **21**, 1114 (1979) [Sov. Phys. Solid State **21**, 647 (1979)].
[23] Ya. E. Pokrovskii, O. I. Smirnova, and N. A. Khval'kovskii, JETP Lett. **61**, 624 (1995).
[24] S. D. Ganichev, E. Zepezauer, I. N. Yassievich *et al.*, Phys. Rev. B **55**, 4 (1997).

Translation was provided by the Russian Editorial office.

# Synchrotron and annihilation channels for axion production in an external electromagnetic field

V. V. Skobelev

*Moscow State Industrial University, 109280 Moscow, Russia*
(Submitted 26 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 25–31 (July 1997)

We consider axion formation processes in the synchrotron ($e^-\to e^-a$) and annihilation ($e^-e^+\to a$) channels in a constant crossed field $F_{\mu\nu}F^{\mu\nu}=F^*_{\mu\nu}F^{\mu\nu}=0$, which approximates constant fields of other configurations in the ultrarelativistic asymptotic limit. The probability and intensity of axion emission are obtained, and we analyze the energy and field asymptotics.

A comparison with the characteristic neutrino channel $e^-\to e^-\nu\overline{\nu}$ yields the constraints on the axion mass and the energy scale for Peccei–Quinn symmetry breaking. Possible astrophysical applications are discussed. © *1997 American Institute of Physics.* [S1063-7761(97)00207-2]

## 1. INTRODUCTION

The line of inquiry associated with the possible existence of the axion as a pseudoscalar Goldstone boson in the mechanism of Peccei–Quinn global symmetry breaking[1] is presently fairly popular (see, for example, the review in Ref. 2). This is due both to the prospect of obtaining a reasonable explanation for the presence (quantum chromodynamics) or absence (the electroweak sector) of CP invariance, and to astrophysical insight into, e.g., the nature of dark matter in the Universe. The nonzero mass $m_a$ of the axion (if it exists) is an inherent attribute of this particle, and can be uniquely expressed in terms of the energy scale $f$ of global symmetry breaking[1)]

$$m_a=\left(\frac{6\times10^6\ \text{GeV}}{f}\right)\text{eV}, \tag{1}$$

the coupling constants of the axion with ordinary particles being proportional to $1/f$. Therefore, estimation of the possible value of $f$ by independent methods is of fundamental importance, especially since the theoretical possibilities extend over an enormous range, from the characteristic electroweak scale $(\sqrt{2}G)^{-1/2}\sim 250$ GeV to the Planck mass $10^{19}$ GeV (an invisible axion).

One of the procedures used involves finding a lower bound on $f$ by requiring that the axion luminosity observed upon the collapse of stellar objects not exceed the neutrino luminosity, which is believed to be one of the principal cooling mechanisms. In particular, there is interest in the synchrotron

$$e^-\to e^-\nu\overline{\nu}, \tag{2a}$$

$$e^-\to e^-a \tag{2b}$$

and annihilation

$$e^-e^+\to\nu\overline{\nu}, \tag{3a}$$

$$e^-e^+\to a \tag{3b}$$

channels for neutrino and axion production, which are open in the strong magnetic fields that are formed, for example, upon collapse in a neutron star (channel (3a) is also open in the absence of a field).

A calculation of the axion luminosity, which is one of the purposes of the present work, requires the specific form of the Lagrangian of the $(aee)$ interaction. The literature describes the use of two forms of the Lagrangian ($m$ is the mass of the electron),

$$\mathscr{L}=-ic\left(\frac{m}{f}\right)a(\overline{\Psi}\gamma^5\Psi), \tag{4a}$$

$$\mathscr{L}=c\frac{1}{2f}\frac{\partial a}{\partial x_\alpha}(\overline{\Psi}\gamma_\alpha\gamma^5\Psi), \tag{4b}$$

with pseudoscalar and pseudovector couplings, respectively, which can lead to results that may or may not be the same[3] ($c$ is the unitary charge of an electron under the global Peccei–Quinn transformation $U_{PQ}$ (1), which is of order unity). The two Lagrangians differ by a total derivative, and in the absence of excitation of the electron–positron vacuum it suffices to consider only the pseudoscalar alternative (4a).

We next use the invariant technique to calculate interaction processes in a constant crossed field

$$F_{\mu\nu}F^{\mu\nu}=F^*_{\mu\nu}F^{\mu\nu}=0, \tag{5}$$

where $F^*_{\mu\nu}=(1/2)e_{\mu\nu\alpha\beta}F^{\alpha\beta}$ is a dual tensor, which was described in a paper by Ritus.[4] This method also yields accurate results in constant fields of arbitrary configuration if

$$\frac{|F_{\mu\nu}F^{\mu\nu}|}{F_0^2},\quad \frac{|F^*_{\mu\nu}F^{\mu\nu}|}{F_0^2}\ll 1,\chi^2, \tag{6}$$

where $F_0=m^2/e=4.41\times10^{13}$ G, and the invariant parameter

$$\chi=\frac{\sqrt{e^2(pF^2p)}}{m^3} \tag{7}$$

is expressed in terms of the ultrarelativistic (according to (6)) electron momentum $p$.

The wave function of an electron in a constant crossed field with potential $a_\alpha\varphi$, where $\varphi=kx$, has the form (we henceforth retain the notation used in Ref. 4, as much as possible)

$$\Psi=\left(1+\frac{e\hat{k}\hat{a}\varphi}{2kp}\right)u(p)\exp\left\{-i\left[\frac{eap}{2kp}\varphi^2-\frac{e^2a^2\varphi^3}{6kp}+px\right]\right\}, \tag{8}$$

$\overline{u}\,u=2m$.

Using (4a) and (8), we can easily obtain an expression for the square of the matrix element of the process $e^-\to e^-a$ after summing and averaging over the polarization of the electron:

$$|M|^2=\frac{4g^2(2\pi)^5}{L_\varphi}\int_{-\infty}^{\infty}ds\;\delta(sk+p-p'-k')\times[(pp'-m^2)|A|^2+4\beta(kk')|A'|^2+\alpha(kk')\mathrm{Im}(A^*A')], \tag{9}$$

$$A(s,\alpha,\beta)=\frac{1}{2\pi}\int_{-\infty}^{\infty}d\varphi\;\exp\left[-i\left(\frac{\alpha\varphi^2}{2}-\frac{4}{3}\beta\varphi^3-s\varphi\right)\right], \tag{9a}$$

$$\alpha=e\left(\frac{ap}{kp}-\frac{ap'}{kp'}\right),\quad \beta=\frac{e^2a^2}{8}\left(\frac{1}{kp}-\frac{1}{kp'}\right), \tag{9b}$$

where $L_\varphi$ is a ''large'' interval of phase $\varphi$, $p'$ and $k'$ are the momenta of the final electron and the axion, $g=cm/f$ is the effective interaction constant, and the derivative $A'(s,\alpha,\beta)$ is taken with respect to the first argument.

Further simplifications are achieved by introducing the new variables

$$\chi=\frac{kp}{m^2}x,\quad \chi'=\frac{kp'}{m^2}x,\quad \kappa=\frac{kk'}{m^2}x,\quad x=\frac{\sqrt{e^2(-a^2)}}{m},$$
$$\rho=\frac{\alpha}{8\beta},\quad \tau=\frac{eF^*_{\mu\nu}p^\mu p'^\nu}{m^4\kappa},\quad u=\frac{\kappa}{\chi'} \tag{10}$$

with an expression for $A$ in terms of Airy functions:

$$A=\frac{(4\beta)^{-1/3}}{\pi}\exp\left[-i\frac{s\alpha}{8\beta}+i\frac{8\beta}{3}\left(\frac{\alpha}{8\beta}\right)^3\right]\Phi(y), \tag{11}$$

$$y=(4\beta)^{2/3}\left[\frac{s}{4\beta}-\left(\frac{\alpha}{8\beta}\right)^2\right]. \tag{11a}$$

Now the differential probability of axion emission per unit time can be written in the form

$$dW_a=\frac{g^2\chi(2\chi/u)^{1/3}}{2\pi^3p_0xL_\varphi\chi'}\frac{u}{1+u}\times\left[\left(\frac{u}{2\chi}\right)^{2/3}(1+\tau^2)\Phi^2+\Phi'^2\right]\frac{d^3k'}{2k'_0}, \tag{12}$$

$$y=\left(\frac{u}{2\chi}\right)^{2/3}\left(1+\tau^2+\widetilde{m}^{\,2}\frac{1+u}{u^2}\right),\quad \widetilde{m}=\frac{m_a}{m}, \tag{12a}$$

where $\chi=\chi'+\kappa$. One characteristic feature of this expression is its lack of dependence on $\rho$, which, after transformation of the phase volume

$$\frac{d^3k'}{2k'_0}\to\frac{xm^2\chi' u}{2\chi(1+u)^2}d\rho\;d\tau\;du, \tag{13}$$

makes it possible to integrate over $\rho$ (see Ref. 4)

$$\int_{-\infty}^{\infty}d\rho=L_\varphi.$$

Using the additional relations

$$y\Phi^2+\Phi'^2=\frac{1}{2}\frac{d^2}{dy^2}\Phi^2,$$

$$\int_{-\infty}^{\infty}d\tau\;\Phi^2(a+b\tau^2)=\frac{\pi}{2\sqrt{b}}\int_{2^{2/3}a}^{\infty}\Phi(z)dz,$$

after integrating over the ''angular'' variable $\tau$, we obtain the distribution with respect to the ''energy'' variable $u$:

$$W_a=-\frac{g^2m^2}{(2\pi)^2p_0}\int_0^\infty du\left[\left(\frac{\chi}{u}\right)^{2/3}\frac{u^2}{(1+u)^3}\Phi'(z)+\frac{\widetilde{m}^{\,2}}{2(1+u)^2}\int_z^\infty\Phi(y)dy\right], \tag{14}$$

$$z=\left(\frac{u}{\chi}\right)^{2/3}\left(1+\widetilde{m}^{\,2}\frac{1+u}{u^2}\right). \tag{14a}$$

Integration by parts reduces the expression for $W_a$ to a single integral,

$$W_a=-\frac{g^2m^2}{(2\pi)^2p_0}\chi\int_0^\infty\frac{du}{1+u}\left[\left(\frac{u}{\chi}\right)^{1/3}\frac{u}{(1+u)^2}\Phi'(z)-\frac{1}{6}\left(\frac{\widetilde{m}}{\chi}\right)^2\left(\frac{\chi}{u}\right)^{1/3}\left(2-\widetilde{m}^{\,2}\frac{4+u}{u^2}\right)\Phi(z)\right], \tag{15}$$

which, of course, no longer describes the distribution with respect to $u$.

The integral can be evaluated for special values of $\widetilde{m}$ and $\chi$.

a) $\widetilde{m}\ll\chi$, $\chi\ll1$. In this case we can generally set $\widetilde{m}=0$. After some simple calculations we obtain

$$W_a\simeq\frac{15g^2m^2}{32\pi\sqrt{3}p_0}\chi^3. \tag{16}$$

b) $\widetilde{m}\ll\chi$, $\chi\gg1$. Similarly, setting $\widetilde{m}=0$ and making the replacement

$$\Phi'(z)\to\Phi'(0)=-\frac{3^{1/6}}{2}\Gamma\left(\frac{2}{3}\right),$$

we find

$$W_a\simeq\frac{g^2m^2\Gamma(2/3)}{2\pi3^{7/3}p_0}\chi^{2/3}. \tag{17}$$

c) $\tilde{m} \gg \chi$. As is easily seen, the argument of the Airy functions in this case is large everywhere in the domain of integration, and its minimum value is

$$z_{\min} = \frac{3}{2^{2/3}} \left( \frac{\tilde{m}}{\chi} \right)^{2/3} \gg 1, \quad \tilde{m} \ll 1. \tag{18}$$

We can use the asymptotic representation

$$\Phi \simeq \frac{\sqrt{\pi}}{2} z^{-1/4} \exp\left( -\frac{2}{3} z^{3/2} \right). \tag{19}$$

We were unable to calculate the corresponding integral, but there is no need to do so, since it is clear from (18) and (19) that the probability is suppressed in any case by the exponential factor

$$\exp\left( -\sqrt{3}\, \frac{\tilde{m}}{\chi} \right) \ll 1. \tag{20}$$

We note that the analogous result from an analysis of the process $e^- \to e^- \nu \overline{\nu}$ with massive neutrinos provided some basis to assert that the small mass of the neutrino can be measured electron storage-ring experiments.[5] It would be very tempting to carry out this program in the case of synchrotron axion emission, since a measurement of the mass would simultaneously also determine the value of the scale factor $f$ (see Eq. (1)). However, this will hardly be possible in the foreseeable future, since it is not clear whether this hypothetical particle exists at all, not to mention the lack of axion detectors.

The mean intensity of axion emission by an electron can be obtained by adding the factor

$$p_0' = p_0 \frac{u}{1+u}$$

to the integrand in Eq. (14), and again integrating by parts

$$I_a = -\frac{g^2 m^2}{(2\pi)^2} \chi \int_0^\infty \frac{du}{(1+u)^2} \left[ \left( \frac{u}{\chi} \right)^{1/3} \frac{u^2}{(1+u)^2} \Phi'(z) - \frac{1}{12} \left( \frac{\tilde{m}}{\chi} \right)^2 \left( \frac{\chi}{u} \right)^{1/3} (1+2u) \left( 2 - \tilde{m}\, \frac{4+u}{u^2} \right) \Phi(z) \right]. \tag{21}$$

As before, we present some asymptotic representations of this expression:

a) $\tilde{m} \ll \chi$, $\chi \ll 1$:

$$I_a = \frac{g^2 m^2}{\pi} \chi^4, \tag{22}$$

b) $\tilde{m} \ll \chi$, $\chi \gg 1$:

$$I_a = \frac{7\Gamma(2/3) g^2 m^2}{2\pi 3^{13/3}} \chi^{2/3}, \tag{23}$$

with the same remarks as for $\tilde{m} \gg \chi$. The luminosity per unit volume, neglecting temperature effects, can be obtained from (21)–(23) by multiplying by the concentration of monoenergetic electrons $n$.

When the annihilation channel $e^+ e^- \to a$ is considered, the replacements $k \to -k$ and $p' \to -p'$ must be made in (9), and summation over the spin states of the final electron must be replaced by averaging. Then, for the annihilation probability per unit time, we obtain

$$W = \frac{g^2}{p_0 p_0' L \varphi V x} \left( \frac{2\chi\chi'}{\kappa} \right)^{1/3} \times \left[ \left( \frac{\kappa}{2\chi\chi'} \right)^{2/3} (1+\tau^2) \Phi^2(y) + \Phi'^2(y) \right], \tag{24}$$

$$y = \left( \frac{\kappa}{2\chi\chi'} \right)^{2/3} \left( 1 + \tau^2 - \tilde{m}^2 \frac{\chi\chi'}{\kappa^2} \right), \tag{24a}$$

where $V$ is the normalization volume, and the result is represented in a symmetrized form with respect to the electron variable $\chi$ and the positron variable $\chi'$, with $\kappa = \chi + \chi'$. The transition to a physically reasonable result[4] is accomplished by multiplying (24) by $VnT$ ($n$ is the electron concentration, and $T$ is the total time) and then making the replacement $T/L_\varphi \to 1/k_0$. This yields the probability of axion annihilation of a positron ''for all time,''

$$W_T = \frac{g^2 n m}{p_0 p_0' F e} \left( \frac{2\chi\chi'}{\kappa} \right)^{1/3} \times \left[ \left( \frac{\kappa}{2\chi\chi'} \right)^{2/3} (1+\tau^2) \Phi^2(y) + \Phi'^2(y) \right], \tag{25}$$

where $F = k_0 \sqrt{-a^2}$ is the field amplitude.

Let us evaluate the scale factor $f$ by comparing the luminosities (or intensities) due to the synchrotron axion $e^- \to e^- a$ and neutrino $e^- \to e^- \nu \overline{\nu}$ emission channels. Typical values of the temperature and the magnetic field induction in neutron star circumstellar shells cover the ranges[6] $T \sim 10^8 - 10^{10}$ K and $F \sim 10^{12} - 10^{14}$ G; such temperatures correspond to an electron energy $p_0/m \sim 10^{-2} - 10^0$ when distribution effects are neglected. On the ''brink'' of fulfillment of (6), we must take the upper energy value and the lower field value together with $\chi \sim 10^{-2}$, which should yield a result that is correct in order of magnitude for a ''pure'' magnetic field. We then clearly have $\tilde{m} \ll \chi$, the effective energy of the axion or neutrino pairs is of order $p_0 \chi$, and in this approximation, instead of comparing the intensities, we can restrict attention to a comparion of probabilities.

Using expression (16) and the results in Refs. 4, 5, and 7, from the requirement $W_a \lesssim W_\nu$ ($c \sim 1$) we obtain

$$f \gtrsim 10^7 \text{eV}. \tag{26}$$

We note that with consideration of (1) this is consistent with the condition $\chi \gg \tilde{m}$. The lower bound for $f$ specified by this relation is approximately three orders of magnitude lower than the value obtained from an analysis of astrophysical data.[2,8] If the latter bound is considered reliable, one of the interpretations of our result will be that the contribution of

the synchrotron mechanism of neutrino and axion emission to the total luminosity for the values of the parameters considered is, in fact, small against the background of the other contributions. If this assumption does not hold, then from (26) and (1) we obtain a possible range for the mass of the axion,

$$m_a \lesssim 1\,\text{eV}. \tag{27}$$

which leaves some hope for measuring it in analogy to the program for measuring the mass of the neutrino described in Ref. 5. In this case the role of the axion as a candidate for a hidden mass ''carrier'' clearly increases.

[1)]See Table 2.1 on p. 17 of Ref. 2 and the notes to it.


[1] R. D. Peccei and H. R. Quinn, Phys. Rev. Lett. **38**, 1440 (1977).
[2] G. G. Raffelt, Phys. Rep. **198**, 1 (1990).
[3] G. Raffelt and D. Seckel, Phys. Rev. Lett. **60**, 1793 (1988).
[4] V. I. Ritus, Tr. Fiz. Inst. im. P. N. Lebedev, Akad. Nauk SSSR **111**, 5 (1979).
[5] V. V. Skobelev, Zh. Éksp. Teor. Fiz. **107**, 322 (1995) [JETP **80**, 170 (1995)].
[6] A. D. Kaminker, K. P. Levenfish, and D. G. Yakovlev, Pis'ma Astron. Zh. **17**, 1090 (1991) [Sov. Astron. Lett. **17**, 450 (1991)].
[7] N. R. Merenkov, Yad. Fiz. **42**, 1484 (1985) [Sov. J. Nucl. Phys. **42**, 938 (1985)].
[8] T. Altherr, E. Petigirard, and T. del Rio Caztelurrutia, Preprint No. CERN-TH 7044/93, Geneva (1993).


Translated by P. Shelnitz

# High quantum efficiency of intersubband transitions in coherent tunneling of electrons through asymmetric double-barrier structures

E. I. Golant and A. B. Pashkovskiĭ

*State Research & Production Corporation ''Istok,'' 141120 Fryazino, Moscow Region, Russia*
(Submitted 25 July 1996)
Zh. Éksp. Teor. Fiz. **112**, 237–245 (July 1997)

A converging perturbation series that can be summed analytically has been obtained for intersubband transitions of electrons coherently tunneling through the middle of a dimensionally quantized level in an asymmetric double-barrier structure in a high-frequency terahertz electric field. The possibility of a substantial increase in tunneling current accompanied by either absorption or emission of a photon has been demonstrated. The quantum efficiency of radiative transitions between dimensionally quantized levels can be up to 66%. © *1997 American Institute of Physics.* [S1063-7761(97)02007-6]

In our earlier publication,[1] we demonstrated that the reflection coefficient of electrons whose energy coincides with the middle of a dimensionally quantized level, and which are driven by a high-frequency (hf) electric field, can be reduced to zero at a field amplitude (in energy units) much lower than the separation between levels. In this process, electrons transfer energy to the hf field, and the current through the structure increases. We have also determined the conditions in which the largest number of electrons (up to 100%) fed to the structure efficiently interact with an hf field at frequency $\omega$, and pass into a neighboring resonance level by emitting or absorbing a photon of energy $\hbar\omega$. There is, however, a question of whether these effects can be detected (much less utilized) in a realistic situation, when energy levels have finite widths and the energy of incident electrons is distributed continuously over some range. In order to solve this problem, one must calculate the change in the electron wave function when its energy and hf field frequency deviate from their exact resonant values.

Following the approach of Ref. 1, we consider an asymmetric double-barrier structure with width $a$ and thin, $\delta$-shaped barriers in a uniform electric field that varies in time as

$$\mathscr{E}\cos\omega t = E(e^{i\omega t}+e^{-i\omega t}), \quad \mathscr{E}=2E.$$

We assume for definiteness that a current of electrons with fixed energy flows from left to right. Then the time-dependent Schrödinger equation has the form

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m^*}\frac{\partial^2\psi}{\partial x^2} + H(x)\psi + H(x,t)\psi,$$

$$H(x) = -U(\theta(x)-\theta(x-a)) - U_1\theta(x-a) + \alpha\delta(x)\psi + \gamma\alpha\delta(x-a), \tag{1}$$

$$H(x,t) = -qE[x(\theta(x)-\theta(x-a)) + a\theta(x-a)] \times (e^{i\omega t}+e^{-i\omega t}).$$

Here $q$ and $m^*$ are the electron charge and mass, $\alpha=\varphi_b b$, $\varphi_b$ and $b$ are the height and width of the first barrier, $\theta(x)$ is the Heaviside step function, $\gamma$ is a numerical factor, $U$ and $U_1$ are the jumps in the conductance band bottom at the barriers, and the interaction potential is written in the form commonly used in the dipole approximation.[2]

Let the initial electron energy coincide with that of level $N$ (the main level) and the hf field frequency correspond to transitions to level $L$. Using the approach developed in Ref. 1, we seek a solution for a sufficiently low field amplitude in the form of a perturbation series without assuming, however, that the resonant condition is exactly satisfied. Depending on the relative positions of the levels and conduction band bottoms on the left and right, we consider three main configurations:

(1) transitions occur between levels above the conduction band bottom in both the left (input) and right (output) semiconductor (Fig. 1a);

(2) electrons go to a level below the conduction band bottom of the left semiconductor and above the conduction band bottom of the right semiconductor (Fig. 1b);

(3) transitions occur between a level below the conduction band bottom of the right semiconductor and a level above the conduction band bottom of the right semiconductor (Fig. 1c).

In the first two cases, the unperturbed electron wave function $\psi_0$ normalized to a single electron is

$$\psi_0(x) = \begin{cases} \exp(ik_0x) + D_0\exp(-ik_0x), & x<0, \\ A_0\sin kx + B_0\cos kx, & 0<x<a, \\ C_0\exp[ik_1(x-a)], & x>a, \end{cases} \tag{2}$$

where

$$k_0 = \sqrt{\frac{2m^*\varepsilon}{\hbar^2}}, \quad k = \sqrt{\frac{2m^*(\varepsilon+U)}{\hbar^2}},$$

$$k_1 = \sqrt{\frac{2m^*(\varepsilon+U_1)}{\hbar^2}}$$

are wave vectors and $\varepsilon$ is the energy of the incident electrons.

In the small-signal approximation, the correction $\psi_1$ to the wave function of the initial state is given by[3]

$$\psi_1 = \psi_{1+}(x)\exp[-i(\omega_0+\omega)t] + \psi_{1-}(x)\exp[-i(\omega_0-\omega)t], \quad \omega_0=\varepsilon/\hbar.$$

 1063-7761/97/070130-05$10.00  

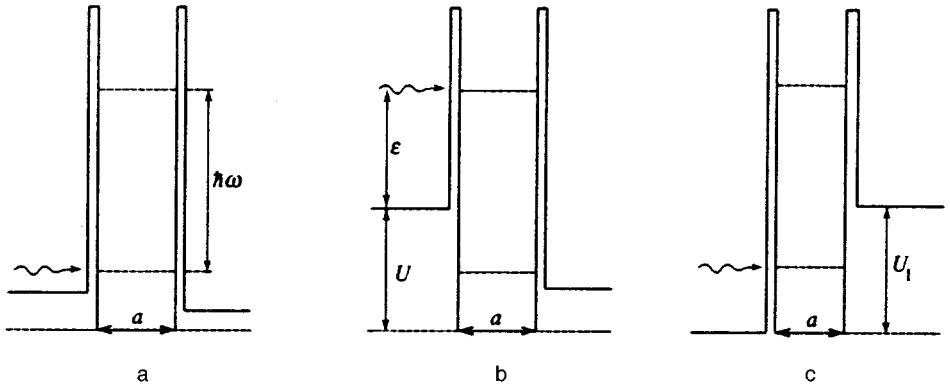



FIG. 1. Diagrams of energy bands in double-barrier structures under discussion.

In the first case[3,4]

$$\psi_{1\pm}(x)=\begin{cases} D_{1\pm}\exp(-ik_{0\pm}x), & x<0, \\ A_{1\pm}\sin k_{\pm}x+B_{1\pm}\cos k_{\pm}x+\chi_{1\pm}(x), & 0<x<a, \\ C_{1\pm}\exp[ik_{1\pm}(x-a)]+P_{1\pm}\exp[ik_1(x-a)], & x>a, \end{cases} \tag{3}$$

where

$$k_{0\pm}=\sqrt{\frac{2m^*(\varepsilon\pm\hbar\omega)}{\hbar^2}},$$

$$k_{\pm}=\sqrt{\frac{2m^*(\varepsilon+U\pm\hbar\omega)}{\hbar^2}},$$

$$k_{1\pm}=\sqrt{\frac{2m^*(\varepsilon+U_1\pm\hbar\omega)}{\hbar^2}},$$

$$\chi_{1\pm}(x)=\mp\frac{qEx\psi_0(x)}{\hbar\omega}+\frac{qE\psi_0'(x)}{m^*\omega^2},$$

$$P_{1\pm}=\mp\frac{qEa\psi_0(a)}{\hbar\omega}.$$

In Ref. 1 the coefficients $A_{1\pm},B_{1\pm},C_{1\pm},D_{1\pm}$ were calculated for the case of electron energies exactly equal to those of mid-level positions. Here we calculate these coefficients for the case of small deviations of the wave vectors from their resonant values. The equations for the coefficients in the wave function (3)[1,3] have the form

$$\begin{pmatrix} 1 & 0 & -1 & 0 \\ ik_{0\pm}-y & k_{\pm} & 0 & 0 \\ 0 & \sin(k_{\pm}a) & \cos(k_{\pm}a) & -1 \\ 0 & -k_{\pm}\cos(k_{\pm}a) & k_{\pm}\sin(k_{\pm}a) & ik_{1\pm}-\gamma y \end{pmatrix}\times\begin{pmatrix} D_{1\pm} \\ A_{1\pm} \\ B_{1\pm} \\ C_{1\pm} \end{pmatrix}=\begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \end{pmatrix}, \tag{4}$$

where

$$f_1=\chi_{1\pm}(0),\quad f_2=-\chi_{1\pm}'(0),\quad f_3=P_{1\pm}-\chi_{1\pm}(a),$$

$$f_4=(\gamma y-ik_1)P_{1\pm}+\chi_{1\pm}'(a),\quad y=2m^*\alpha/\hbar^2.$$

When the barriers are high enough, $y\gg k_{\pm}$, and at a wave vector corresponding to a resonant level (this condition is satisfied when $\tan(k_{\pm}a)=-(1+\gamma)k_{\pm}/\gamma y$), the determinant of the system (4) is small:

$$\Delta(k_{\pm})\approx ik_{\pm}(k_{1\pm}+\gamma^2k_{0\pm})(-1)^{L+1}/\gamma,$$

and in the case of nonresonant tunneling the determinant is, at least, linear in $y$. Thus, for narrow resonance levels, the transition probability will only be significant between two levels. Therefore, we consider in what follows only transitions between the initial and higher resonance levels (labeled by ''+''), or the initial and lower levels (labeled by ''−'').

Let the wave vectors of electrons transmitted across the structure and those interacting with the hf electric field be close to resonance:

$$k=k_r+\delta k,\quad k_{\pm}=k_{r\pm}+\delta k_{\pm},$$

where $k\gg\delta k$, $k_{\pm}\gg\delta k_{\pm}$. It can be shown that the determinant of the system (4) in this case is

$$\Delta(k_{r\pm}+\delta k_{\pm})\approx\left(\frac{ik_{\pm}}{\gamma}(k_{1\pm}+\gamma^2k_{0\pm})-\gamma y^2\delta k_{\pm}a\right)\times(-1)^{L+1}. \tag{5}$$

Naturally, in this approach it makes sense to take into

consideration only such deviations from resonance at which $\Delta(k_{r\pm}+\delta k_\pm)$ is only slightly larger than $\Delta(k_{r\pm})$ and

$$\frac{\Delta(k_{r\pm}+\delta k_\pm)}{\Delta(k_{r\pm})} \ll \frac{y}{k_\pm}. \tag{6}$$

Given that the matrix describing transmission of electrons through the structure without changing their energy is similar to the matrix in Eq. (4), and the determinants used in the calculation of the coefficients $A_0$, $B_0$, and $C_0$ of the time-independent wave function (2) depend weakly on $\delta k$, we obtain

$$D_0=\frac{\gamma^2k_0-k_1-is}{\gamma^2k_0+k_1+is}, \quad B_0=\beta_0=\frac{2\gamma^2k_0}{\gamma^2k_0+k_1+is},$$
$$C_0=\frac{\beta_0}{\gamma}(-1)^{N+1}, \quad A_0=\beta_0\frac{y}{k}+i\frac{2k_1k_0}{k(\gamma^2k_0+k_1+is)}, \tag{7}$$

where

$$s=\frac{\gamma^2y^2\delta k a}{k}.$$

Note that when the barriers are strong enough ($y \gg k$), the ground-state wave function on the interval $0<x<a$, as in the case of exactly resonant transmission,[1] can be expressed as

$$\psi_0 \approx \beta_0\left[\frac{y}{k}\sin(kx)+\cos(kx)\right]. \tag{8}$$

For $y \gg k_\pm$, taking into account only terms with maximal powers of $y$ (only terms with $f_1$ and $f_3$ contribute to the column-wise expansion of the determinant), we derive from Eq. (4) the coefficients of the wave function (3) for transitions in which the change in the level number is odd:

$$D_{1\pm}\approx\frac{qE}{im^*\omega^2}\frac{2\gamma^2y^2}{k_{1\pm}+\gamma^2k_{0\pm}+is_\pm}\beta_0=\beta_1,$$
$$B_{1\pm}\approx D_{1\pm}\approx(-1)^{L+1}\gamma C_{1\pm}, \quad A_{1\pm}\approx\frac{y}{k_\pm}D_{1\pm}. \tag{9}$$

Note that, when the initial and final energies of electrons are close to the centers of the levels, an important property of resonant transmission noted in Ref. 1 applies: between the barriers ($0<x<a$) the first-order correction to the ground-state wave function is similar to the wave function itself:

$$\psi_{1\pm}(x)\approx\beta_1\left[\frac{y}{k_\pm}\sin(k_\pm a)+\cos(k_\pm a)\right]. \tag{10}$$

Here we have taken into account the fact that, since $y \gg k_\pm$ and $\chi_{1\pm}(x)$ contains only terms at most linear in $y/k_\pm$, while the deviation from the resonant condition is such that condition (6) is satisfied for the determinant of the system (4), the contribution to the first-order correction to the wave function due to $\chi_{1\pm}(x)$, like the contribution to the particular solution of the equation for the second-order correction,[4] is small, as in the case of exact resonance.[1] Moreover, since $|C_{1\pm}| \gg |P_{1\pm}|$ and the function $f_4$, which includes $P_{1\pm}$, does not make a significant contribution to $\psi_{1\pm}$, here and in subsequent calculations we can neglect components like $P_{1\pm}\exp[k_1(x-a)]$. Thus, having repeated the calculation of corrections to the wave function described above, taking into account the fact that only $\psi_{1-}(x)$ contributes to the second-order correction if the initial electron energy coincides with the upper resonant level, or only $\psi_{1+}(x)$ if it coincides with the lower level, whereas the other components are small, we obtain

$$\psi_2(x)\approx\begin{cases} D_2\exp(-ik_0x), & x<0,\\ A_2\sin(kx)+B_2\cos(kx), & 0<x<a,\\ C_2\exp[ik_1(x-a)], & x>a,\end{cases} \tag{11}$$

where

$$D_2\approx-\left(\frac{qE}{m^*\omega^2}\right)^2 \times\frac{4\gamma^4y^4}{(k_1+\gamma^2k_0+is)(k_{1\pm}+\gamma^2k_{0\pm}+is_\pm)}\beta_0=-\beta_0z,$$
$$B_2\approx D_2\approx(-1)^{N+1}\gamma C_2, \quad A_2\approx\frac{y}{k}D_2,$$
$$s_\pm=\frac{\gamma^2y^2\delta k_\pm a}{k_\pm}, \tag{12}$$
$$\psi_2\approx-\beta_0z\left(\frac{y}{k}\sin(kx)+\cos(kx)\right), \quad 0<x<a,$$
$$z=\left(\frac{qE}{m^*\omega^2}\right)^2\frac{4\gamma^4y^4}{(k_1+\gamma^2k_0+is)(k_{1\pm}+\gamma^2k_{0\pm}+is_\pm)}. \tag{13}$$

Here, as in the previous case, we have taken into account only terms with the maximum power of $y/k$.

Thus, if we continue the procedure of calculation of higher-order corrections, as in Ref. 1, taking into account the smallness of $k/y$, and then add them, the wave function coefficients at each resonant level can be expressed in the form of a constant factor times an alternating series (unlike the case of an exact resonance, it is complex):

$$1-z+z^2-z^3+\ldots+(-1)^{n+1}z^n. \tag{14}$$

In its domain of convergence $|z|<1$, this is the expansion of $1/(1+z)$ in powers of $z$, and the desired electron wave function has the form

$$\psi\approx\psi_N(x)\exp(-i\omega_0t)+\psi_L(x)\exp[-i(\omega_0\pm\omega)t], \tag{15}$$

where

$$\psi_N(x)=\begin{cases}\exp(ik_0x)+D\exp(-ik_0x), & x<0,\\ A\sin(kx)+B\cos(kx), & 0<x<a,\\ C\exp[ik_1(x-a)], & x>a,\end{cases} \tag{16}$$

$$D=D_0-\frac{\beta_0z}{1+z}, \quad A=\frac{1}{1+z}A_0, \quad B=\frac{1}{1+z}B_0,$$
$$C=\frac{1}{1+z}C_0,$$

$$\psi_L(x)=\frac{1}{1+z}\begin{cases} D_{1\pm}\exp(-ik_{0\pm}x), & x<0,\\ A_{1\pm}\sin(k_{\pm}x)+B_{1\pm}\cos(k_{\pm}x), & 0<x<a,\\ C_{1\pm}\exp[ik_{1\pm}(x-a)], & x>a. \end{cases} \tag{17}$$

Interestingly enough, the larger the offset from resonance, the greater the field amplitude at which the series (14) converges.

In the second case, when the resonance level to which the electrons go is below the conduction band bottom downstream of the structure, $\psi_{1-}$ is derived from Eq. (3), and the equation system for its coefficients is derived from Eq. (4) by replacing $ik_{0-}$ with $-\kappa$, where $\kappa=[2m^*(\hbar\omega-\varepsilon)/\hbar^2]^{1/2}$. The determinant of this system is small if

$$\tan(k_-a)=-\left(\frac{1+\gamma}{\gamma}-\frac{\kappa}{y}\right)\frac{k_-}{y},$$

and at small offset from resonance,

$$\Delta(k_{r-}+\delta k_-)\approx\left(\frac{ik_-k_{1-}}{\gamma}-\gamma y^2\delta k_-a\right)(-1)^{L+1}. \tag{18}$$

Using the procedure of calculation of corrections to the wave function described above with due account of the fact that Eq. (18) applies to transitions from the higher to lower level, and Eq. (5) to transitions from the lower to higher level, one can easily show that the solution of the problem is described by the wave function (15) with $\psi_N(x)$ determined by Eq. (16), while $\psi_L(x)$ takes the form

$$\psi_L(x)=\frac{1}{1+z}\begin{cases} D_{1-}\exp(\kappa x), & x<0,\\ A_{1-}\sin(k_-x)+B_{1-}\cos(k_-x), & 0<x<a,\\ C_{1-}\exp[ik_{1-}(x-a)], & x>a, \end{cases} \tag{19}$$

where

$$D_{1-}\approx\frac{qE}{im^*\omega^2}\frac{2\gamma^2y^2}{k_{1-}+is_-}\beta_0,$$

$$B_{1-}\approx D_{1-}\approx(-1)^{L+1}\gamma C_{1-},\quad A_{1-}\approx\frac{y}{k_-}D_{1-}, \tag{20}$$

$$z=\left(\frac{qE}{m^*\omega^2}\right)^2\frac{4\gamma^4y^4}{(k_1+\gamma^2k_0+is)(k_{1-}+is_-)}. \tag{21}$$

In the third case, when electrons transfer from a level below the conductance band bottom upstream of the structure to a level above the conductance band bottom, given that the resonance conditions have the form

$$\tan(ka)=-\left(\frac{1+\gamma}{\gamma}-\frac{\kappa}{\gamma^2y}\right)\frac{k}{y},\quad \kappa=\sqrt{\frac{2m^*(U_1-\varepsilon)}{\hbar^2}},$$

and the determinant of the equation system for corrections to the wave function is

$$\Delta(k_r+\delta k)=(i\gamma k_0k-\gamma y^2\delta ka)(-1)^{N+1}, \tag{22}$$

one can easily show that the solution is described by the wave function (16), where

$$\psi_N(x)=\begin{cases} \exp(ik_0x)+D\exp(-ik_0x), & x<0,\\ A\sin(kx)+B\cos(kx), & 0<x<a,\\ C\exp[-\kappa(x-a)], & x>a, \end{cases} \tag{23}$$

$$D=\frac{k_0-is}{k_0+is}-\frac{\beta_0z}{1+z},\quad A=\frac{\beta_0}{1+z}\frac{y}{k},\quad B=\frac{\beta_0}{1+z},$$

$$C=\frac{\beta_0}{1+z}\frac{1}{\gamma}(-1)^{N+1},\quad \beta_0=\frac{2k_0}{k_0+is},\quad s=\frac{y^2\delta ka}{k},$$

$$\psi_L(x)=\frac{1}{1+z}\begin{cases} D_{1+}\exp(-ik_{0+}x), & x<0,\\ A_{1+}\sin(k_+x)+B_{1+}\cos(k_+x), & 0<x<a,\\ C_{1+}\exp[ik_{1+}(x-a)], & x>a, \end{cases} \tag{24}$$

$$D_{1+}\approx\frac{qE}{im^*\omega^2}\frac{4\gamma^2y^2}{k_{1+}+\gamma^2k_{0+}+is_+},$$

$$B_{1+}\approx D_{1+}\approx(-1)^{L+1}\gamma C_{1+},\quad A_{1+}\approx\frac{y}{k_+}D_{1+}, \tag{25}$$

$$z=\left(\frac{qE}{m^*\omega^2}\right)^2\frac{4\gamma^2y^4}{(k_{1+}+\gamma^2k_{0+}+is_+)(k_0+is)}. \tag{26}$$

These equations enable us to estimate the intensity of interaction between electrons and hf field of a finite amplitude in conditions of finite widths of resonant levels and an arbitrary energy distribution of incident electrons.

It is clear that even for the case of a fixed energy of electrons, the above expressions for the wave function coefficients are notably different from those given in Ref. 1 for the case of exact resonance. Since the range of possibilities is wide, especially when realistic electron energy distributions are considered, it is hardly possible to compare the resonant and nonresonant interaction between electrons and hf field in all possible cases. Therefore, we discuss the cases of electron transmission through the structure which are, in our opinion, the most interesting. We assume that the electrons are uniformly distributed over a certain energy range, at least when the deviation from the resonance is comparable to the main-level width.

Let the energy of incident electrons be lower than the conductance band bottom on the right, thus the third configuration of energy levels is realized (Fig. 1c). At resonance ($s=0$) and at the field amplitude corresponding to $z_r=1$, all electrons go to the upper level (the reflectivity $|D|^2=0$).[1] When the electron energy deviates from resonance by $\pm\Gamma_N/2$, where $\Gamma_N$ is the resonant-level width, the static transmission coefficient is half the resonance value. This corresponds to twice the determinant (22) squared:

$$|\Delta(k_r+\delta k)|^2=2|\Delta(k_r)|^2.$$

The upper level width $\Gamma_L$ is usually much larger than that of the lower level (for a symmetric structure $\Gamma_L/\Gamma_N=(L/N)^3$), so we can assume that all electrons absorbing photons with energy $\hbar\omega$ go to the region of allowed states around the middle of the upper level ($s_+\ll k_{1+}+\gamma^2k_{0+}$). It follows from Eqs. (23) and (26) that the dynamic reflectivity at the edges of the transmission band decreases five-fold

($D=-i/(2+i)\neq 0$, $|D|^2=0.2$), thus about 80% of incident electrons at the lower level go to the upper level.

A much more interesting process is the transfer of electron energy to the field due to the hf field. Consider the second configuration of the structure (Fig. 1b) and transitions from the upper to lower level. If the transmission coefficient is high ($k_1\approx\gamma^2 k_0$), the structure transparency cannot be increased by the hf field. Therefore the largest change in the current through the structure should occur when the static transmission coefficient is small. In the case of exact resonance of electrons with the upper level ($s=0$, $D_0$ is real) and at $z_r=D_0$, the transmission coefficient goes to zero.

Let us estimate the offset from resonance at which the transmissivity increases to half of $D_0^2$. It follows from Eqs. (16) and (21) that this occurs at $s_-/k_{1-}=1+D_0$. For $D_0$ close to unity, this offset is approximately double the width of the lower level $\Gamma_L$, where $s_-/k_{1-}=1$. Let us approximate the dynamic transmission coefficient by unity within the ''transmission band'' and by zero outside it. Thus the increase in current due to the hf field is $\Delta I \sim 2\Gamma_L D_0^2$, and the direct current via the states of the main (upper) level is proportional to the product of its width and the transmission coefficient: $I \propto \Gamma_N(1-D_0^2)$. With due account of Eq. (7), we obtain

$$\frac{\Delta I}{I}\approx\left(\frac{L}{N}\right)^2\frac{k_{1-}}{k_1}\frac{D_0^2}{1+D_0}. \tag{27}$$

It is clear that when the ratio $(L/N)^2 k_{1-}/k_1$ is close to unity (transitions occur between levels with large quantum numbers), $\Delta I/I\approx 0.5$. Detection of this increase in the current concurrently with emission of photons with energy $\hbar\omega$ is quite feasible.

In the cases discussed above, the hf field generates the greatest change in the current through the structure. From the standpoint of application of such devices as sources of radiation, the more interesting parameter is quantum efficiency, i.e., the ratio of the number of electrons that have emitted a photon to the total number of transmitted electrons.

Consider the second configuration of energy bands in a structure with a sufficiently large transmission coefficient ($k_1\approx\gamma^2 k_0$). In the case of resonant tunneling and $z_r=1$, the fraction of incident electrons incident upon the lower level is $(1+D_0)/2$. Using Eqs. (16) and (21), one can show that this ratio is a factor of two smaller at $s_-/k_{1-}=2$, i.e., in this case the transmission band is also twice as wide as the lower, narrower level. By approximating the dynamic transmission coefficient by unity within the transmission band and by zero outside it, as was done above, we calculate the currents through the upper and lower levels, which are proportional in the presence of a hf field to

$$I\propto(\Gamma_N-2\Gamma_L)(1-D_0^2)+\Gamma_L(1-D_0^2)/2,$$

$$I_\omega\propto\Gamma_L(1+D_0),$$

and, taking into account Eq. (7) we obtain their ratio:

$$\frac{I_\omega}{I}\approx\frac{L^2k_{1-}}{2k_1[N^2-3L^2k_{1-}/2(\gamma^2k_0+k_1)]}. \tag{28}$$

For $k_1\approx\gamma^2 k_0$ (large transmission coefficient), we have

$$\frac{I_\omega}{I}\approx\frac{L^3}{2N^3-3L^3/2} \tag{29}$$

if $k_{1-}/k_1\approx L/N$, and

$$\frac{I_\omega}{I}\approx\frac{L^2}{2N^2-3L^2/2} \tag{30}$$

if the ratio $k_{1-}/k_1$ is close to unity. Hence it follows that for sufficiently large quantum numbers of resonance levels $I_\omega/I\approx 2$, thus the quantum efficiency of photon emission can be up to 66%, which is much higher than the parameters obtained heretofore.[5] This discrepancy can be attributed to the fact that in experimental conditions,[5] scattering plays an important role in electron transport, whereas the estimates given above apply to the case of ballistic transport, for which the time of flight through the structure is appreciably shorter than the collision time in it. Such conditions can be realized in both symmetric (with an injector) and asymmetric double-barrier structures with thin enough (much thinner than in Ref. 5) barriers, so that the electron time of flight should be relatively short and the resonant properties of the structure persist.[6] In principle, the quantum numbers of levels at which the maximum quantum efficiency is achieved can be significantly reduced by fabricating a structure with $k_1>\gamma^2 k_0$. At first sight, it might seem that the ratio $I_\omega/I$ can be indefinitely large in this case. One must bear in mind, however, that Eq. (28) applies only when the upper level is considerably wider than the lower level, and if they are comparable, a more accurate calculation is needed.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 94-02-04449) and Scientific Council on the Physics of Solid-State Nanostructures program (Grant No. 1-050).

[1] E. I. Golant and A. B. Pashkovskii, JETP Lett. **63**, 590 (1996).
[2] N. B. Delone and V. P. Krainov, *Atoms in Intense Optical Fields* [in Russian], Energoatomizdat, Moscow (1984), p. 16.
[3] A. B. Pashkovskiĭ, Fiz. Tekh. Poluprovodn. **29**, 1712 (1995) [Semicond. **29**, 893 (1995)].
[4] A. B. Pashkovskiĭ, Pis'ma Zh. Tekh. Fiz. **21**(11), 28 (1995) [Tech. Phys. Lett. **21**, 595 (1995)].
[5] J. Faist, F. Capasso, C. Sirtoru, D. Sivco, A. L. Hutchinson, S. N. G. Chu, and A. Y. Cho, Appl. Phys. Lett. **64**, 1144 (1994).
[6] E. I. Golant, A. B. Pashkovskiĭ, and A. S. Tager, Pis'ma Zh. Tekh. Fiz. **20**(**21**), 74 (1994) [Tech. Phys. Lett. **20**, 886 (1994)].

Translation was provided by the Russian Editorial office.

# Photoinduced light absorption by $C_{60}$ films in the 0.08–4.0-eV spectral range

A. V. Bazhenov, A. V. Gorbunov, M. Yu. Maksimuk, and T. N. Fursova

*Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia*

(Submitted 20 August 1996)
Zh. Éksp. Teor. Fiz. **112**, 246–256 (July 1997)

Spectra of photoinduced light absorption in $C_{60}$ films at high and low excitations in the temperature range between 15 and 300 K have been measured. In addition to the well-known explanation of photoinduced absorption in terms of optical transitions in the system of photogenerated singlet excitons, triplet excitons, and polarons, changes in the absorption spectrum of the fullerite ground state must be considered. We suggest taking into account the effect of crystal field in explaining the features of the photoinduced absorption spectrum. A feature similar to the inverted luminescence spectrum and ascribed to optical excitation of singlet excitons, which is partially allowed owing to intermolecular interaction, has been detected in spectrum of photoinduced absorption. © *1997 American Institute of Physics.* [S1063-7761(97)02107-0]

## 1. INTRODUCTION

It follows from the analysis of published data[1] that crystalline $C_{60}$ behaves in some experiments as an almost ideal molecular crystal. In such crystals, absorption of light leads, as a rule, to generation of nonconducting electronic excitations, i.e. excitons.[2] Presently, there is no commonly accepted interpretation of fullerite luminescence in the spectral range around 1.7 eV,[1,3,4] but all authors agree that this luminescence is due to recombination of singlet excitons. At the same time, some experimental results[1] can be interpreted in terms of a one-electron semiconductor model. Investigations of spectra of photoconductivity,[5–7] persistent photoconductivity,[7,8] photovoltaic[9] and xerographic[10] effects have demonstrated that in fullerite, as in a common semiconductor, free carriers are generated by light. Wang *et al.*[1] suggested that the best model is, probably intermediate between these two extreme approaches.

Fullerite doped with alkali metals is a high-temperature superconductor (HTSC). Like semiconducting phases of a copper HTSC, it is a highly correlated system.[11] It is conceivable that it manifests some features of cuprates, in which optical generation of electron–hole pairs or $p$-doping leads to radical restructuring of the electronic spectrum, namely, the spectral weight of transitions between the valence and conducting bands decreases, and a complex absorption spectrum emerges in the low-energy range (0.1–0.5 eV).[12–18]

The photoinduced absorption technique, which enables one to study changes in the electronic spectrum at a low uniform concentration of optically generated excitations, has demonstrated its efficacy in semiconducting phases of a copper HTSC. In the fundamental absorption range of fullerite, the photoinduced absorption technique at $T=300$ K allowed detection of redistribution of the spectral weight in favor of optical transitions dipole-forbidden in the $C_{60}$ molecule,[19] also known as HOMO–LUMO transitions (highest occupied molecular orbital–lowest unoccupied molecular orbital). It is worthwhile to study photoinduced absorption in $C_{60}$ at low temperatures, which can be radically different from that reported in Ref. 19, since phase transitions occur in fullerite at $T\simeq 260$ and 80–90 K.[20]

Dick *et al.*[21] interpreted results similar to those of Ref. 19, but obtained at high levels of excitation by femtosecond laser pulses, as optical transitions in a system of photoexcited singlet excitons, whereas the spectrum of photoinduced absorption was interpreted in Ref. 19 as changes in the ground-state absorption due to optical pumping. In addition, unlike the case of pulsed laser excitation, a broad absorption band with a maximum around 1.1 eV was detected under excitation by a cw laser with a photon energy lower than the fullerite band gap.[22] Magneto-optical measurements[21] demonstrated that in this case the spectrum of photoinduced absorption in the range 0.5–2.5 eV consists of lines with energies 1.1 and 1.8 eV, due to generation of excitations with spin 1, and lines with energies of 0.8 and 2.0 eV due to excitations with spin 1/2. These pairs of lines were associated with transitions between states of triplet excitons and polarons, respectively.

In order to test this model, clarify the inconsistencies, and search for photoinduced optical transitions indicating correlation interactions in the $C_{60}$ fullerite, we have studied photoinduced absorption spectra both at high and low levels of optical pumping in the spectral range 0.08 to 4.0 eV, at temperatures between 15 and 300 K.

## 2. EXPERIMENTAL TECHNIQUES

We have studied fullerite films with thicknesses of about 0.1 and 2.0 $\mu$m fabricated from $C_{60}$ powder by vacuum sublimation and deposition on a sapphire substrate. Their quality was tested using IR absorption spectra and by viewing through an optical microscope. At a substrate temperature of about 100 °C, the film surface was specular, i.e., the crystallite size was less than 1 $\mu$m. The basic measurements were performed on films with clearly defined texture. According to x-ray diffraction measurements, the (001) axis of the fullerite was normal to the substrate surface. No absorption lines associated with either $CO_2$ ($2350\ \mathrm{cm}^{-1}$) or C–H

 1063-7761/97/070135-06$10.00  

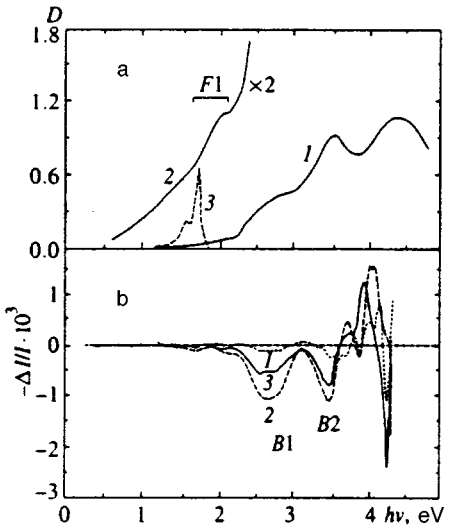



FIG. 1. (a) Spectra of optical density of $C_{60}$ films with thicknesses of (curve *1*) 0.1 and (*2*) 2.0 $\mu$m on sapphire substrates at $T=300$ K. Curve *3* is a luminescence spectrum at 15 K plotted in arbitrary units. (b) Effect of 4-°C overheating in a $C_{60}$ film with a thickness of 0.1 $\mu$m at (*1*) $T=270$ K, (*2*) 170, and (*3*) 90 K.

($2900\ \text{cm}^{-1}$) vibrations of the organic solvent were detected in IR spectra. Narrow absorption lines corresponded to previously known intramolecular vibrational modes of $C_{60}$.

Photoinduced absorption spectra in the range 1.1–4.0 eV were obtained using a prism spectrograph and an optical multichannel analyzer, and in the range 0.08–1.2 eV using a Fourier transform spectrometer. The photoexcitation was induced either by a cw $Ar^+$ laser ($h\nu=2.54$ eV) with power density $P\leqslant 1\ \text{W/cm}^2$ or by an XeCl excimer laser ($h\nu=4.03$ eV) with a pulse width of 20 ns and a repetition rate of 20 Hz (the time-averaged power density was $P=60\ \text{mW/cm}^2$, the peak power density was $P\simeq 150\ \text{kW/cm}^2$). In both cases, we measured the time-averaged signal of photoinduced optical absorption $\Delta I/I=(I-I_1)/I$, where $I_1(\nu)$ and $I(\nu)$ are the intensities of probe light transmitted through the sample in the presence and in the absence of laser pumping, respectively. The difference signal $\Delta I$ was integrated over multiple measurements of $I_1(\nu)$ and $I(\nu)$. In each cycle, the luminescence spectrum was recorded in addition to transmission spectra, and subtracted from the latter (curve *3* in Fig. 1a). As a result, we completely eliminated the luminescence contribution and greatly attenuated scattered laser light. The measurement accuracy of $\Delta I/I$ was $10^{-5}$ in the range 1.2–4.0 eV, and $10^{-4}$ in the range 0.08–1.2 eV.

Measurements of photoinduced absorption are always complicated by overheating, i.e., changes in the light transmission owing to sample heating by laser light. In our case, overheating was the more significant, as the sample was located in the evacuated volume of a helium cryostat. In order to elucidate the spectral signature of overheating, we measured changes in the optical transmission of a film, $\Delta I/I=[I(T)-I(T+\delta T)]/I(T)$, at a temperature variation $\delta T=4$ °C provided by a small heater. The optical density spectrum in Fig. 1a (curve *1*) demonstrates that at $h\nu=4.03$ eV, laser light excited almost the entire film volume. Therefore the film was heated by laser light uniformly, and this technique could be utilized in measurements of features due to the overheat by laser light. These measurements data for a $C_{60}$ film on a sapphire substrate are given in Fig. 1b. The spectral features of the heating effect at $T=300$ K and their relation to the fundamental absorption spectrum were discussed in Ref. 19. When the temperature was lowered to about 170 K, the intensities of the $B1$ and $B2$ lines increased considerably, which corresponded to a drop in the intensities of the lines at 2.7 and 3.5 eV in the fundamental absorption spectrum. Unlike overheating at 300 K, at low temperatures there was a drop in absorption in the 1.5–2.3-eV spectral range. When the temperature was further reduced to 90 K, overheating abated.

## 3. EXPERIMENTAL RESULTS

The optical density spectrum of fullerite films with thicknesses of 0.1 $\mu$m at $T=300$ K (Fig. 2a) is similar to spectra previously described in the literature[23] and consists of dipole-allowed transitions in the $C_{60}$ molecule with energies of 3.5 and 4.4 eV, and forbidden $F1$ and $F2$ transitions with red edges at about $h\nu=1.65$ and 2.25 eV. The low-energy edge of $F1$ is difficult to detect in the 0.1-$\mu$m film, but in the spectrum of the 2-$\mu$m film (curve *2* in Fig. 1a) it is clearly visible against the background absorption, which monotonically decreases with decreasing energy down to $h\nu=0.6$ eV. This absorption suggests the existence of tails in the density of states. It has been established[7,24] that fullerite films are disordered, and exhibit properties typical, for example, of amorphous Si:H.

The spectrum of photoinduced absorption due to a pulsed laser with $h\nu=4.03$ eV at 300 K (Fig. 2b) is notably different from the overheating spectrum. It does not contain the $B1$ and $B2$ features of the latter (Fig. 1b), i.e., laser overheating can be neglected at 300 K. An $S$-shaped feature can be seen around the $F1$ transition in the photoinduced absorption spectrum, i.e., photobleaching in the range 1.6 to 2.0 eV, and absorption enhancement with a maximum at $h\nu=2.2$ eV. Optical generation of electron–hole pairs leads to higher absorption around the $F2$ transition. The absorption line at $h\nu=3.5$ eV corresponds to an $S$-shaped feature in the photoinduced absorption spectrum, with the inflection point near the absorption-line maximum, i.e., the line shifts to lower energy under optical pumping. Moreover, since the inflection point is in the region of negative photoinduced absorption, photoinduced absorption at the maximum of the 3.5-eV line is reduced by optical pumping.

As the temperature is lowered to 90 K, photoinduced absorption turns into an increase in transmission (Fig. 2c). But this effect is hardly caused by the phase transition at 260 K. The spectrum clearly demonstrates the $B1$ and $B2$ features typical of overheating (Fig. 1b), i.e., the combination of overheating and of photoinduced absorption proper is measured. We could not eliminate overheating by reducing the pumping power density at 90 K. The dashed line in Fig. 2c shows an approximate spectrum of photoinduced absorption obtained by subtracting the overheating spectrum multiplied

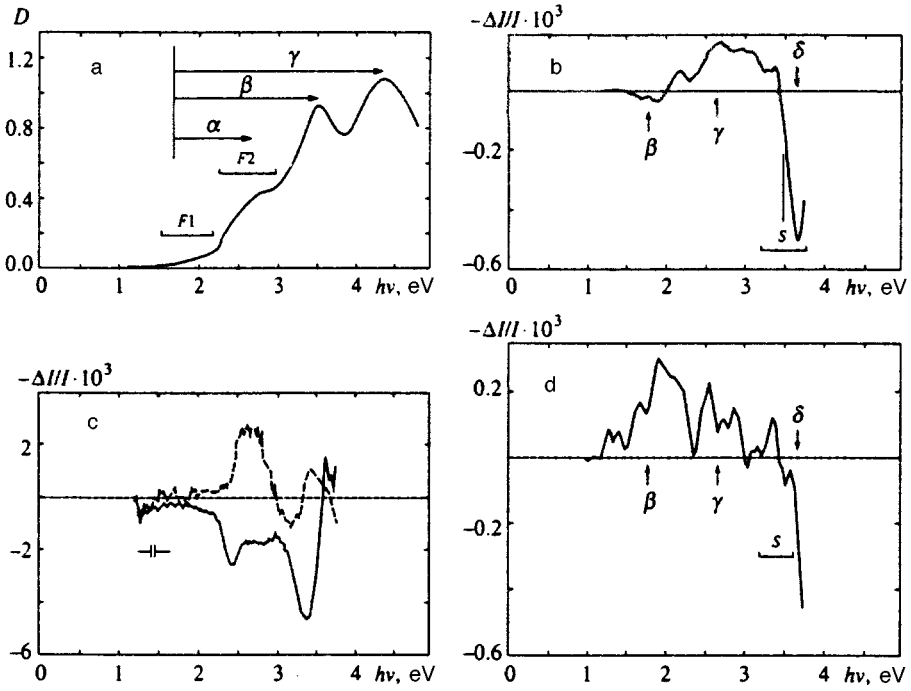



FIG. 2. Spectra of (a) optical density and (b, c, d) photoinduced absorption of a $C_{60}$ film when nonequilibrium carriers are generated by a pulsed XeCl laser with $h\nu=4.03$ eV at $T=300$ K, 90 K, and 15 K, respectively.

by a factor at which features characteristic of overheating are eliminated (this corresponds to overheating by 2 °C). Its shape is similar to that of the spectrum recorded at 300 K, but its intensity is an order of magnitude higher.

When the temperature is reduced to 15 K, the effect is again inverted (Fig. 2d), and the spectrum shape is similar to that obtained at 300 K. This results from a decrease in overheating due to the temperature reduction from 90 to 15 K. Thus, the thermal effect can be neglected at 15 K, as at 300 K.

Note that the overheating contribution to the photoinduced absorption spectra of $C_{60}$ is notably larger at higher pumping powers and larger film thicknesses. This fullerite property is substantially different from that of the copper HTSC, because its thermal conductivity (0.4 and 1.2 W/m·K at 300 and 77 K, respectively[25]) is about an order of magnitude lower than in cuprate compounds. At low temperatures this difference is even more significant, since the fullerite specific heat drops by a factor of seven as the temperature is reduced from 300 to 100 K.[26]

The photoinduced absorption spectra under cw optical pumping ($P\leqslant 1$ W/cm$^2$) were studied in the spectral range 0.08 to 2.2 eV. No changes in the transmission of a 0.1-$\mu$m film, to within $\Delta I/I=10^{-4}$, were detected in the 0.08–1.6-eV range. Since the photoinduced absorption signal should be proportional to the film thickness in sufficiently thick films, we tested a 2-$\mu$m film (optical density spectrum *2* in Fig. 1a). The analysis of photoinduced reflection and absorption of such a film based on equations given in Ref. 27 indicates that, at least at $h\nu\geqslant 0.6$ eV, reflection can be neglected, and the effect should be determined only by changes in the optical density.

As shown in Fig. 3a, pumping of fullerite by a cw $Ar^+$ laser at $T=15$ K generates a broad absorption line with a maximum at $E_1\simeq 1.1$ eV, a feature at $E_0\simeq 0.8$ eV at its low-energy edge, and an absorption line at $E_2\simeq 1.9$ eV. The shape of the *B* feature in the 1.3–1.8-eV range is essentially identical to the inverted luminescence spectrum of the film (dashed line in Fig. 3a); only the position of the luminescence line maximum is 0.03 eV lower than the minimum of the *B* line. When optical pumping is reduced, the intensities of the $E_1$, $E_2$, and *B* features decrease proportionally (Fig. 4a). The photoinduced absorption at $h\nu=1.1$ and 1.9 eV depends on the pumping power as $P^{1/2}$ (Fig. 5).

The shapes of the 1.1-eV and 1.9-eV lines are constant with temperature (Fig. 4b), but above 40 K the photoinduced

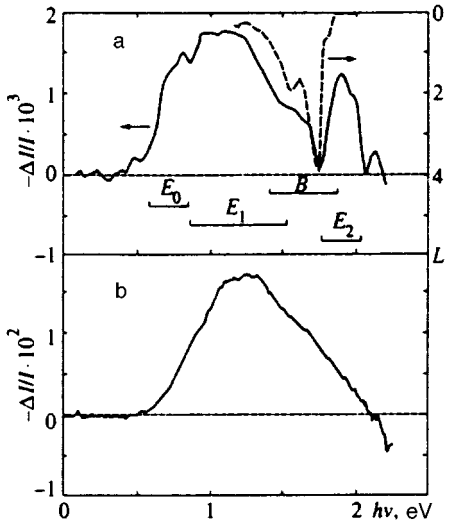



FIG. 3. Spectra of photoinduced absorption of (a) fresh and (b) aged $C_{60}$ film with a thickness of 2.0 $\mu$m under an electron–phonon pair excitation by a cw $Ar^+$ laser ($h\nu=2.54$ eV). The pumping density $P=1$ W/cm$^2$, $T=15$ K. The dashed line shows an inverted luminescence spectrum *L* plotted in relative units.

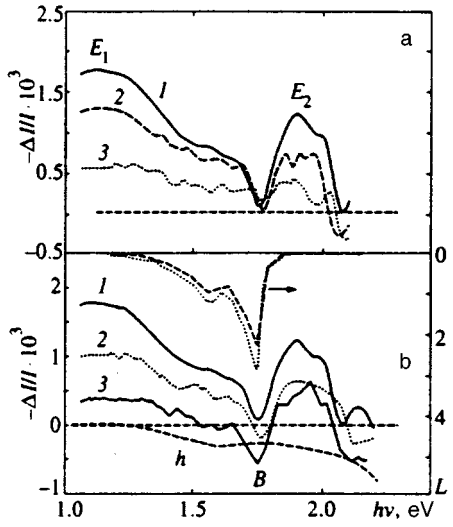



FIG. 4. (a) Spectra of photoinduced absorption of a 2.0-$\mu$m $C_{60}$ film at different pumping power densities $P=1000$, 250, and 45 mW/cm$^2$ (curves *1*, *2*, and *3*, respectively) at $T=15$ K. (b) Spectra of photoinduced absorption at $T=15$, 60, and 90 K (curves *1*, *2*, and *3*, respectively) for $P=1$ W/cm$^2$. Curves $L1-L3$ are luminescence spectra at the same temperatures; $h$ is the overheating spectrum at $T=90$ K.

absorption at $h\nu \simeq 1.7$ eV and $h\nu \geqslant 2.0$ eV becomes negative because of the larger contribution of overheating. At 15 K this contribution is small, but at 90 K the overheating spectrum exhibits an appreciable decrease in absorption for $h\nu \geqslant 1.5$ eV (dashed line $h$ in Fig. 4b). Therefore, in plotting experimental data against temperature, the amplitude of the 1.9-eV line was measured with respect to the curve $h$. Figure 6 shows that the amplitude of the 1.1-eV photoinduced absorption line is a flat function of temperature below 40 K, and drops abruptly as the temperature approaches 90 K. At the same time, the amplitude of the 1.9-eV line changes notably less in the range between 15 and 90 K. Its intensity and the amplitude of increased transmission in the $B$ feature as functions of temperature are similar to the curve of luminescence at $h\nu=1.7$ eV versus temperature, which is essentially constant between 15 and 50 K and drops by a factor of only 1.3 as the temperature reaches 90 K. According to Matus *et al.*,[3] this temperature dependence of the luminescence intensity of a $C_{60}$ film indicates a low content of impurities.

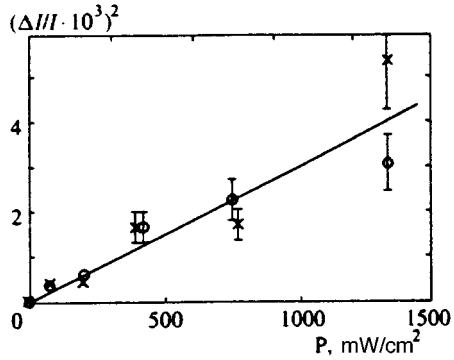



FIG. 5. Photoinduced absorption squared in a 2.0-$\mu$m fullerite film at (○) $h\nu=1.1$ eV and (×) 1.9 eV versus pumping power density at $T=15$ K.

It turned out that optical properties of fullerite films change with time. In particular, the photoinduced absorption at 0.6–2.1-eV increases considerably (Fig. 3b). The spectral position of the photoinduced absorption is constant, but its spectral shape is notably different, specifically, there is no enhanced transmission at the $B$ feature. Note that even in this film, in which the magnitude of the effect is an order of magnitude higher, no photoinduced absorption lines were detected in the 0.08–0.5-eV spectral range.

## 4. DISCUSSION

### Low pumping power

The similarities between the spectrum shapes of 1.7-eV luminescence and the increase in transmission in the same spectral range (Figs. 3a and 4b) and their temperature dependences indicate that this feature in the photoinduced absorption spectrum is due to the photoinduced decrease in light absorption by luminescent states. At low pumping power density ($P \simeq 1$ W/cm$^2$), a decrease in the excitonic absorption in the fullerite ground state is possible only at sufficiently long lifetimes of singlet excitons. At $T=300$ K and high power of femtosecond laser pumping, the characteristic luminescence decay time is in the nanosecond range, and the luminescence intensity drops with time as $t^{-0.57}$ (see Ref. 21). Given that the luminescence intensity decays as a power of time, one can expect the lifetime of a singlet exciton to increase to microseconds at low temperatures and low pumping powers, then excitonic absorption should decrease owing to partial filling of excitonic states. This mechanism is supported by the absence of a feature at $h\nu=1.73$ eV in the spectrum of photoinduced absorption of aging films. It turned out that unlike freshly fabricated films, they contain lines of $C_{60}O_x$ and $C_{60}H$ complexes in their IR spectra. It is known that fullerite oxidation leads to lower luminescence efficiency at $h\nu=1.69$ eV,[3] i.e., a shorter exciton lifetime.

Presently, several mechanisms of fullerite luminescence are under discussion, namely, recombination of autolocalized excitons,[3] free excitons in the crystal volume and excitons localized on defects (lines at 1.69 eV and a set of neighboring lines),[4] and recombination of singlet excitons.[1] Since the latter process is dipole-forbidden owing to the state symmetry, a vibronic model was invoked in interpreting it. According to this model, the red edge of excitonic absorption is at $h\nu=1.95$ eV, i.e., it is shifted with respect to the luminescence spectrum to the high-energy side by $\Delta=0.26$ eV.

In the case under discussion, $\Delta=0.03$ eV, and it can be interpreted in terms of relaxation of interatomic bonds under photoexcitation. Investigation of the luminescence excitation spectra of $C_{60}$ films[28] confirm that $\Delta$ is considerably smaller than the value suggested by the vibronic mechanism.[1] The excitation spectrum, in fact, contains lines of vibronic origin at 1.92, 2.0, and 2.2 eV, but the excitation efficiency increases as the energy of incident photons decreases to 1.8 eV. This observation cannot be interpreted in terms of the vibronic model.[1]

A model in which symmetric wave functions of singlet states responsible for the absorption at $h\nu=2.7$ eV mix with the antisymmetric function of states participating in the tran-

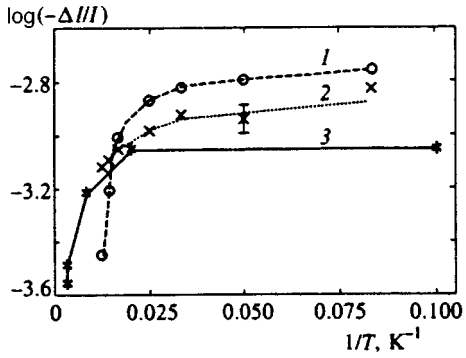



FIG. 6. Logarithm of photoinduced absorption at $h\nu=1.1$ eV and 1.9 eV (curves *1* and *2*, respectively) and of luminescence intensity at $h\nu=1.7$ eV (curve *3*) in a 2.0-$\mu$m fullerite film as a function of temperature. The pumping power density $P=1$ W/cm$^2$.

sition $h\nu=3.5$ eV was suggested in Ref. 19. This mixing is due to the crystal field and should increase with both dropping temperature and increasing pumping power. As follows from calculations,[29] the 2.7-eV line corresponds to the dipole-forbidden transition $^1A_g\rightarrow{}^1T_{2g}$, and the 3.5-eV line to the allowed transition $^1A_g\rightarrow{}^1T_{1u}$. Studies of spectra of luminescence quenching in $C_{60}$ films by electric field and of photoconductivity spectra indicate[28] that these transitions are related to intermolecular charge-transfer states, i.e., absorption of photons with $h\nu>2.3$ eV generates free carriers. The admixture of the antisymmetric wave function to charge-transfer singlet states absorbing light at $h\nu=2.7$ eV (the low-energy edge is at 2.25 eV) leads to a similar admixture to the singlet exciton state. As a result, dipole-forbidden transitions are partially allowed not only for charge-transfer states, but also for singlet excitons. Therefore the spectral positions of the excitonic luminescence and of the lowest excitonic absorption line are close.

The lines with energies 1.1 and 1.9 eV in spectra of photoinduced absorption were interpreted[21] as light absorption by photoexcited triplet excitons. In this case, their intensities should be described by similar functions of temperature and pumping power density. The latter condition is satisfied with fair accuracy (Fig. 5), but the intensity of the 1.9-eV line is a notably flatter function of temperature than the intensity of the 1.1-eV line (Fig. 6). The reason for this discrepancy in our case can be an appreciable contribution of the 1.9-eV transition in the system of singlet excitons generated by the pumping light to the photoinduced absorption.[21] This is also suggested by the similarity between the curves of the intensity of the 1.7-eV singlet exciton luminescence line and of the 1.9-eV photoinduced absorption line plotted against temperature, as well as partial saturation of absorption by singlet excitons with an energy of about 1.73 eV.

#### High pumping power

The most important difference of photoinduced absorption spectra under pulsed pumping is the absence of the 1.1-eV line (compare Figs. 2d and 3). At high excitation powers, the photoinduced absorption spectrum in the picosecond range is controlled by singlet excitons, according to conclusions by Dick *et al.*[21] Our measurements with pulsed excitation refer to a 10-ns time interval or more, but the spectrum at $T=15$ K was similar to that[21] measured in the spectral range $h\nu\leqslant 2.2$ eV at $T=300$ K. That is, singlet excitons dominate in the nanosecond range in $C_{60}$, at least at low temperatures. The reason may be, as pointed out in Ref. 21, that the conversion time of singlet excitons to triplets is substantially longer than $\tau=1.3$ ns, which is typical of $C_{60}$ molecules in solution. The low intensity of the photoinduced absorption line at 1.1 eV (Fig. 2), which is due to triplet excitons,[21] suggests that the concentration of triplet excitons generated by 20-ns laser pulses is low.

The behavior of photoinduced absorption in the spectral range 1.7–4.0 eV with changing temperature indicates that the inversion of the photoinduced absorption as the temperature is reduced from 300 to 90 K and the repeated reversal of the sign of the effect between 90 and 15 K are not associated with changes in electronic properties of the fullerite due to phase transitions at $T\simeq 260$ K and 90 K. The inversion of photoinduced absorption is due to changes in the thermal conductivity and specific heat of the film–sapphire substrate system with decreasing temperature.

Given the equal parity of the ground and lowest excited states of the $C_{60}$ molecule, Dick *et al.*[21] suggested that at high pumping powers the photoinduced absorption spectrum should be similar to the absorption spectrum of the fullerite in the ground state, but shifted by the energy of the first excited state, which corresponds to a singlet exciton. It follows from our measurements that this energy is 1.7 eV, i.e., the transitions with energies of 3.5, 4.4 eV (Fig. 2a), and 5.7 eV[27] allowed in the dipole approximation should correspond to the lines $\beta=1.8$ eV, $\gamma=2.7$ eV, and $\delta=4.0$ eV (Fig. 2a) in the photoinduced absorption spectra. The energy of latter line is essentially identical to that of the pumping laser line, and the line could not be detected, while the energies of the former two lines are close to the energies of lines detected at low temperatures (Fig. 2d). The ratio between their intensities, however, is opposite that of the respective lines in the optical density spectrum in Fig. 2a. At room temperature (Fig. 2b), the line at $\beta=1.8$ eV is much weaker than the line at $\gamma=2.7$ eV, whereas in accordance with the absorption spectrum of the fullerite ground state (Fig. 2a), they should not be so different (the 2.2-eV line observed in this spectrum must be ascribed to a transition in the system of photoexcited polarons[21]).

In our opinion, the coincidence of the 1.8-eV and 2.7-eV lines with the $F1$ and $F2$ transitions in the fullerite absorption spectrum (Fig. 2a) should be taken into account. Therefore, photoinduced changes in the ground state absorption can contribute to the photoinduced absorption, along with light absorption by photoexcited singlet excitons. An increase in the absorption due to dipole-forbidden transitions in the $C_{60}$ molecule due to the crystal field is most probable. Its dominant effect at 300 K can account for the absence of the 1.8-eV line from the photoinduced absorption spectrum and the presence of the intense 2.7-eV line.

In the $h\nu=1.6-2.25$ eV range (transition $F1$ in Fig. 2a), the absorption spectrum of the fullerite ground state is determined by a number of factors. Absorption due to tails in the

density of states can be ruled out because the spectral shape is substantially different. The high-energy section of the absorption feature in the range of 1.95 to 2.1 eV (Fig. 1a) results from vibronic transitions,[1,30] the low-energy section (1.65–1.95 eV) is due to partially dipole-allowed transitions in which singlet excitons are excited.

Another feasible mechanism of photoinduced absorption in the low-energy section is due to HOMO–LUMO optical transitions, which are forbidden owing to both system symmetry and spin conservation. It follows from the calculation[29] that the low-energy edge of such singlet-to-triplet transitions $^1A_g \rightarrow {}^3T_{2g}$ is at $h\nu = 1.6$ eV. Since their oscillator strength is a factor of about $10^8$ lower than that of the lowest dipole-allowed transition at 3.5 eV,[29] such transitions are not observed in conventional absorption spectra. But electrons and holes generated by pumping light can partly lift the restriction due to spin conservation, and the $^1A_g \rightarrow {}^3T_{2g}$ transition can be seen in the spectrum of photoinduced absorption at low temperature: the line at 1.9 eV can become more intense than the 2.7-eV line (transitions $\beta$ and $\gamma$ in Fig. 2d).

## 5. CONCLUSIONS

No radical restructuring of the fullerite energy spectrum due to optical generation of electron–hole pairs has been detected, unlike the case of semiconducting cuprate HTSC compounds. Thus, electronic properties of the fullerite can be interpreted in terms of the fixed-band model.

The model proposed by Dick *et al.*[21] to explain photoinduced absorption in fullerite films in terms of optical transitions in the system of photoexcited singlet excitons in the case of short-lived excitation and transitions in the system of triplet excitons for long-lived excitations is certainly applicable. Other mechanisms, however, must be taken into consideration, the most significant factor being the crystal field, which leads to emergence of optical transitions from the ground state (HOMO–LUMO) in photoinduced absorption, which are dipole-forbidden in the $C_{60}$ molecule. In fullerite crystals, these transitions are weakly allowed owing to interaction between molecules, and optical pumping (like a decrease in temperature[19]) intensifies this interaction, and hence increases absorption due to dipole-forbidden transitions.

Our investigation of luminescence and photoinduced absorption in $C_{60}$ films has led us to conclude that the radiative recombination of excitons with energy 1.70 eV, and the inverse absorption process at 1.73 eV can proceed not only through the vibronic absorption mechanism, but also without participation of intermolecular vibrational modes. This can also be explained in terms of the crystal field, which can also account for the absence of an analog of the 2.7-eV fullerite line in solutions of $C_{60}$, and the lower luminescence quantum efficiency in solutions, as compared to fullerite crystals.[28]

This work was part of the Fullerenes and Atomic Clusters Russian R&D program, it was supported by Grant No. 93193 from the HTSC state-sponsored program, and Grants REX000 and REX300 from the International Science Foundation and Russian Government. The authors are grateful to I. N. Kremenskaya, who supplied $C_{60}$ powder.

[1] Y. Wang, J. M. Holden, A. M. Rao *et al.*, Phys. Rev. B **51**, 4547 (1995).
[2] A. S. Davydov, *Theory of Molecular Excitons* [in Russian], Nauka, Moscow (1968).
[3] M. Matus, H. Kuzmany, and E. Sohmen, Phys. Rev. Lett. **68**, 2822 (1992).
[4] W. Guss, J. Feldmann, E. O. Göbel *et al.*, Phys. Rev. Lett. **72**, 2644 (1994).
[5] J. Mort, R. Ziolo, M. Machonkin *et al.*, Chem. Phys. Lett. **186**, 281 (1991).
[6] C. H. Lee, G. Yu, D. Moses *et al.*, Phys. Rev. B **48**, 8506 (1993).
[7] M. Hosoya, K. Ichimura, Z. H. Wang, G. Dresselhaus, M. S. Dresselhaus, and P. C. Eklund, Phys. Rev. B **49**, 4981 (1994).
[8] A. Hamed, H. Rasmussen, and P. H. Hor, Phys. Rev. B **48**, 14760 (1993).
[9] H. Yonehara and C. Pac, Appl. Phys. Lett. **61**, 575 (1992).
[10] J. Mort, M. Machonkin, R. Ziolo, and I. Chen, Appl. Phys. Lett. **61**, 1829 (1992).
[11] R. W. Lof, M. A. van Veennedaal, B. Koopmans, H. T. Jonkman, and G. A. Sawatzky, Phys. Rev. Lett. **68**, 3924 (1992).
[12] J. M. Ginder, M. G. Roe, Y. Song, R. P. McCall, J. R. Gaines, and E. Ehrenfreund Phys. Rev. B **37**, 7506 (1988).
[13] S. Uchida, T. Ido, H. Takagi, T. Arima, Y. Tokura, and S. Tajima, Phys. Rev. B **43**, 7942 (1991).
[14] G. A. Thomas, D. H. Rapkin, S. L. Cooper, S.-W. Cheong, A. S. Cooper, L. F. Schneemeyer, and J. V. Waszczak, Phys. Rev. B **45**, 2474 (1992).
[15] A. V. Bazhenov, A. V. Gorbunov, K. B. Rezchikov, T. N. Fursova, A. A. Zakharov, and M. B. Tsetlin, Physica C **208**, 197 (1993).
[16] Y. H. Kim, S.-W. Cheong, and Z. Fisk, Phys. Rev. Lett. **67**, 2227 (1991).
[17] H. J. Ye, R. P. McCall, W. E. Farneth, E. M. McCarron, and A. J. Epstein, Phys. Rev. B **43**, 10574 (1991).
[18] A. V. Bazhenov, A. V. Gorbunov, and V. B. Timofeev, Zh. Éksp. Teor. Fiz. **104**, 3193 (1993) [JETP **77**, 500 (1993)]
[19] A. V. Bazhenov, A. V. Gorbunov, and K. G. Volkodav, JETP Lett. **60**, 331 (1994).
[20] F. Gugenberger, R. Heid, C. Meingast, P. Adelmann, M. Braun, H. Wuhl, M. Haluska, and H. Kuzmany, Phys. Rev. Lett. **69**, 3774 (1992).
[21] D. Dick, X. Wei, S. Jeglinski, R. E. Benner, Z. V. Vardeny, D. Moses, V. I. Srdanov and F. Wudl, Phys. Rev. Lett. **73**, 2760 (1994).
[22] A. V. Bazhenov, A. V. Gorbunov, and V. B. Timofeev, in Proceedings of *Anharmonic Properties of High-$T_c$ Cuprates*, Slovenija, Bled, September 1994, D. Mihailovič, G. Ruani, E. Kaldis, and K. A. Müller, (eds.), World Scientific, Singapore (1994), p. 267.
[23] W. Krätschmer, L. D. Lamb, K. Fostiropoulos, and D. R. Huffman, Nature (London) **347**, 354 (1990).
[24] R. A. Cheville and N. J. Halas, Phys. Rev. B **45**, 548 (1992).
[25] R. C. Yu, N. Tea, M. B. Salamon *et al.*, Phys. Rev. Lett. **68**, 2050 (1992).
[26] T. Atake, T. Tanaka, H. Kawaji *et al.*, Physica C **185–189**, 427 (1991).
[27] A. V. Bazhenov, A. V. Gorbunov, and K. B. Rezchikov, Fiz. Tverd. Tela (Leningrad) **37**, 760 (1995) [Phys. Solid State **37**, 413 (1995)].
[28] S. Kazaoui, R. Ross, and N. Minami, Phys. Rev. B **52**, R11665 (1995).
[29] K. Yabana and G. F. Bertsch, Chem. Phys. Lett. **197**, 32 (1992).
[30] C. Reber, L. Yee, J. McKiernan *et al.*, J. Phys. Chem. **95**, 2127 (1991).

Translation was provided by the Russian Editorial office.

# Kinetics of avalanche mixing of granular materials

S. N. Dorogovtsev*)

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia*
(Submitted 17 September 1996)
Zh. Éksp. Teor. Fiz. **112**, 257–277 (July 1997)

The problem of the avalanche mixing of two fractions of granular material is solved. Mixing of the fractions takes place in a cylinder that rotates slowly about its longitudinal axis, which is positioned horizontally. The cylinder is not filled completely and at all times mixing only occurs in the surface layer of granules. It is shown that, depending on the relation of the volumes of the fractions and the volume of the empty space, mixing can take place slowly, over a large number of rotations, in a diffusive regime with convection or rapidly, by the time the cylinder has turned through a small angle. The mixing process is described analytically in terms of a purely geometrical approach and can, in a number of situations, be reduced to a sequence of discrete mappings. The characteristic mixing times are determined, including the times over which one or the other of the pure fractions no longer exists in the regions adjacent to the surface of the cylinder. Their dependence on the degree of filling of the cylinder and on the ratio of the volumes of the fractions is found.  [S1063-7761(97)02207-5]

## 1. INTRODUCTION

A new burst of interest in problems associated with the interspersion and mixing of granular materials in rotating cylinders[1–11] has arisen following the publication of pioneering work on self-organized criticality.[12–14] These problems are also related to the intensely studied crumbling of sandpiles. Because of the complexity of these systems, the overwhelming bulk of published papers on this topic have involved experimental studies and cumbersome computer simulations. In general, the physics of the mixing process has turned out to be extremely nontrivial and the experimentally observed patterns are often very striking, as, for example an illustration in Ref. 10 appeared on the cover of the March 1995 issue of *Nature*. That figure essentially initiated the present paper.

Exactly what sort of systems are we discussing? A granular material has been poured into a cylinder that is rotating about its axis, which is positioned horizontally. The cylinder lies in a gravitational field and is not filled completely, so that a free space remains on the top. (See Fig. 1a). The questions of how the material will be interspersed, how the granules will mix, and how different fractions of the material will mix, if there is more than one, are indeed very complicated and a direct computer simulation will require thousands of hours of calculations on large machines. (See Ref. 2, for example.) In fact, the problem is three dimensional; granules can also move along the longitudinal axis of the cylinder and the surface is not necessarily flat; the answers depend on the rate of rotation of the cylinder, on the characteristics of the granules, etc.

The experiment of Metcalfe *et al.*[10] is remarkable in that the system proposed there makes it possible to distinguish the essential feature of the mixing process in a rotating cylinder. Furthermore, as we shall show below, mixing in this system can be explained analytically,[11] without resorting to any kind of microscopic behavior and essentially using a purely geometric approach! The following discussion is based on the configuration introduced in Ref. 10.

Our cylinder (or drum) is flat, so that it is more or less a disk and mixing of granules along the longitudinal axis of the cylinder is negligible. The cylinder rotates adiabatically slowly and the free surface of the granular material, as in the experiment,[10] is plane and always lies at the friction angle (the angle of repose) to the horizontal. (In a real experiment the angle of inclination of the surface fluctuates slightly in time, but it turns out that these variations are small[10] and we shall neglect them.) For definiteness we shall assume that the direction of rotation of the cylinder is counterclockwise. (See Fig. 1a.)

We shall assume that while the granules are inside the volume of the material they cannot move relative to one another or slip relative to the cylinder walls. They can intersperse only when they escape to the free surface of the material. Mixing of the different fractions of material, therefore, takes place in avalanches which move continuously downward along the free surface. Thus, this kind of mixing has been called avalanche mixing,[10] although it would be more precise to refer to it as mixing in avalanches. Since the granules intersperse only at the free surface, within an inner circle of radius equal to the normal from the center to the free surface (the radius vector $OE$ in Fig. 1b), the material rotates along with the cylinder as a whole without mixing. Thus, we have only to study processes taking place in the material lying in the outer annulus adjacent to the cylinder walls. When the drum is less than half full of material, of course, all of it mixes.

We shall assume that the granules are small and we can introduce the concept of a concentration of one or another fraction at a given point in the drum. The granules of the different fractions will be distinguished only by color. The state at each point of the material will be described by the quantity $\rho$, the concentration (portion) of the black material at that point. (Where all of the material is black, $\rho=1$ and

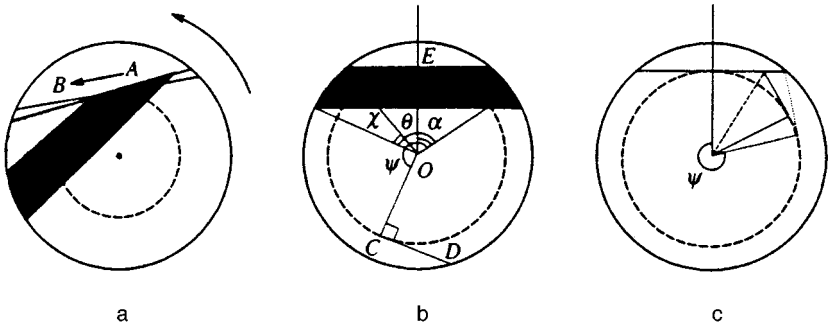



FIG. 1. (a) A sketch of avalanche mixing. For an infinitely small rotation of the drum the granules of the different fractions from sector $A$, on mixing, spill into section $B$. The free surface of the granulated material lies at the friction angle to the horizontal. The region with mixed material is shaded. Here and in the pictures to follow the different degrees of mixing are not indicated. (b) The distribution of the fractions before the cylinder starts to rotate. The position of the tangent $CD$ is fixed by the angle $\psi$. (c) One of the tangents intersects the free surface of the material $2\pi-2\theta<\psi<2\pi$. This is the case of a cylinder that is more than half full.

where it is all white, $\rho=0$.) Thus, we have to study the variable $\rho(\mathbf{x},t)$, the concentration of the black fraction at point $\mathbf{x}$ at time $t$.

Let the white fraction be below and the black, above, at the initial time. Later it will be convenient to introduce the notation shown in Fig. 1b. We shall characterize the fraction of unfilled space in the drum and the volume of the black fraction by the angles $\theta$ and $\chi$, respectively. We shall not give the trivial formulas relating the volumes of the fractions and these angles here. In the following we shall also use the angle $\alpha$, whose significance is indicated in Fig. 1b. It is easy to confirm that it obeys

$$\alpha=\arccos[\cos(\theta+\chi)/\cos\theta]. \tag{1}$$

If $\chi\tan\theta\ll 1$, then it is easy to show that $\alpha\cong\sqrt{2\chi\tan\theta}$. It is also evident that $\alpha=\pi$ for $2\theta+\chi=\pi$, i.e., if the white fraction occupies the same volume as the free space.

How fast should the cylinder rotate to satisfy the adiabaticity condition? Let $T_0$ be the rotation period of the cylinder and $\delta$ be the friction angle formed by the free surface of the material. We require that the angle $\epsilon$ through which the slowly rotating cylinder turns as a granule slides from one edge of the free surface to the other be much smaller than the friction angle $\delta$. This is the adiabaticity condition in our case. Let $t_\delta$ be the time over which a granule slides from one edge of the free surface to the other. Then

$$\epsilon=2\pi t_\delta/T_0\ll\delta. \tag{2}$$

Simple elementary estimates yield

$$t_\delta\simeq 2\sqrt{\frac{R}{g}\frac{\sin\theta}{\sin\delta}}. \tag{3}$$

Here $R$ is the cylinder radius and $g$ is the acceleration of gravity. This leads to the following condition for the rotation period of the cylinder:

$$T_0\gg 4\pi\frac{1}{\delta}\sqrt{\frac{R}{g}\frac{\sin\theta}{\sin\delta}}. \tag{4}$$

(Usually the formula $T_0\gg\sqrt{R/g}$, introduced as a justification for the dimensional arguments in Ref. 10, is employed in place of Eq. (4).)

When condition (4) is satisfied, the rotation speed of the drum does not enter in our answers, so that in the following we take the time to be simply the angle of rotation of the drum, i.e., we measure ''time'' in terms of angles. The answer does not depend on the cylinder radius, and we set that equal to unity. Finally, the friction angle does not appear in our answers, so in all the mixing schemes shown below (except that of Fig. 1a), we set it equal to zero for convenience; the horizontal surface will be depicted as a horizontal line.

We now proceed to one of the key points. We have to make a strong assumption which, as such, allows us to solve the problem avoiding any microscopic analysis. This assumption makes it possible to convert our model essentially into a problem on a circle and reduce the problem to a purely geometric one.[11] Thus, we shall assume that in the avalanches the material mixes completely, i.e., to a homogeneous state. (See Fig. 1a; if this assumption is satisfied, then the material mixes completely, spilling over from sector $A$ to sector $B$.) It is natural to choose this assumption as an initial approximation if we do not consider details associated with the structure of the granules (their clustering, meshing, etc.). Then as the cylinder is rotated, at all points on the left half of the free surface of the material the concentration $\rho$ of the black material is the same. Thus, even after the first turn (of course, with the cylinder more than half filled), for all the material in the outer annulus we can introduce $\rho(\psi,t)$, the density of the black material at time $t$ at points on the tangent of the type denoted by $CD$ in Fig. 1b ($\psi$ is the angle between the corresponding radius vector and the normal to the free surface) and the fundamental quantity with which we shall be dealing below. If, on the other hand, the cylinder is less than half full, $\rho(\psi,t)$ gives an exhaustive description of the state of the system for $t>2(\pi-\theta)$.

We note as justification that a computer simulation[10] employing this assumption gave a surprisingly good fit to the original experiment.

We shall describe the evolution of $\rho(\psi)$ in Secs. 2 and 3. Two mixing regimes will be distinguished: slow mixing, with the drum more than half full (Sec. 2), and rapid mixing, with the drum less than half full (Sec. 3). In the first case, we find a discrete linear mapping (9) which describes the transformation $\rho(\psi)$ after each new turn of the drum. (We believe that the mapping (9) is the principal result of this paper.) The answer is especially simple when the cylinder is almost half full. When we transform to a coordinate system attached to the rotating cylinder, the evolution of $\rho(\psi)$ appears as diffusion with convection on a circle: the black fraction gradually washes out into the white, while the maximum in $\rho(\psi)$ drifts in the direction of rotation of the drum. Here the drift velocity is of order $\theta$ (recall that we are measuring the dimension-

less time in terms of angles) and the characteristic mixing time is of order $\theta^{-2}$.

When the drum is less than half full, the material is already mixed over times of order $2(\pi-\theta)$. Here an exhaustive description of the mixing is provided by the linear integral equations (15) and (16) for $\rho(\psi=0,t)$. These equations can be solved analytically. We shall see that the resulting $\rho(\psi=0,t)$ contains rapidly damped oscillations similar to those observed in the experiment.[10]

The situation is unique for a half filled ($\theta=\pi/2$) cylinder. In this case, as we shall see, complete mixing does not occur in general, even though all the material is interspersed on reaching the free surface. The system rapidly goes into a state where the distribution of the fractions changes periodically. The corresponding period is half the rotation period of the cylinder. Part of the black fraction remains in its pure form, while the pure white fraction vanishes.

We shall find two other characteristic times for this system. (See Sec. 4.) It turns out that for an analytic calculation of these times it is sufficient to use a substantially weaker assumption than that employed above. We are concerned with the answers to two questions: (1) after what time $T_w$ does it first become true that $\rho>0$ everywhere, except perhaps the central region in the case $\theta<\pi/2$, i.e., when does the maximum possible volume of the white material first ''get spoiled?'' Or, when is there no more white material left in the outer annulus? (2) After what time $T_b$ does $\rho<1$ everywhere, except perhaps the central region in the case $\theta<\pi/2$, i.e., when does the maximum possible volume of the black material first get ''spoiled?'' Or, when is no more black material left in the outer annulus?

In order to find $T_w$ and $T_b$ analytically we only have to introduce the following requirements. We require that after the granules spill out over the free surface, the material should be in a mixed state at each of the points on its left half if there are granules of both types on the right half. (See Fig. 1a.) As opposed to the previous assumption, the degree of mixing at different points in the left half of the free surface can now be arbitrary. After the first turn of the drum (if, for example, $\theta<\pi/2$), there are now three possibilities for each individual tangent of the type $CD$ shown in Fig. 1b: the material at all points on the tangent is either black or white, or in a mixed state. After studying the evolution of the state for all these tangents, we find $T_w$ and $T_b$. Thus, we have again reduced our problem to a problem on a circle.

In Section 4 we construct functions $T_w(\theta)$ and $T_b(\theta)$, on each of which a sequence of nonanalytic points (discontinuities) is observed. (See Eqs. (32) and (35).)

## 2. MIXING DYNAMICS IN THE CASE OF CYLINDERS THAT ARE MORE THAN HALF FULL

We shall describe the kinetics of avalanche mixing in the case where the cylinder is more than half full, i.e., $\theta<\pi/2$). As we have already shown, the above assumption of complete mixing in the avalanches makes it possible to describe the distribution $\rho(\psi,t)$ after the first turn. Then the total amount of the black fraction outside a circle of radius $1\cdot\cos\theta$ will be

$$M=\int_0^{2\pi}d\psi\int_0^{r(\psi)}dr\ r\rho(\psi,t)$$
$$=\frac{\sin^2\theta}{2}\int_0^{2\pi-2\theta}d\psi\rho(\psi,t)$$
$$+\frac{\cos^2\theta}{2}\int_{2\pi-2\theta}^{2\pi}d\psi\rho(\psi,t)\tan^2\left(\frac{2\pi-\psi}{2}\right). \qquad (5)$$

(Of course, $M$ does not vary in time.) Here the coefficient in front of the first integral is the result of integrating

$$\int_0^{\sin\theta}dr\ r$$

along the tangent over the range $(0,2\pi-2\theta)$ of angles $\psi$ for the radius vector. When, on the other hand, these angles lie between $2\pi-2\theta$ and $2\pi$, the tangent is cut by the free surface (see Fig. 1c), its length $r(\psi)$ can easily be shown to equal $\cos\theta\tan[(2\pi-\psi)/2]$, and we obtain the second integral of Eq. (5).

Since mixing takes place only on the free surface and the granules inside the interior volume of material rotate together with the drum, for $0\leqslant\varphi<2\pi$ and $t\geqslant\varphi$, we have

$$\rho(\varphi,t)=\rho(0,t-\varphi). \qquad (6)$$

After substituting this relation in Eq. (5) and substituting the variable $\epsilon=t-\psi$, we obtain

$$\frac{2M}{\sin^2\theta}=\int_{t-2\pi+2\theta}^{t}d\epsilon\rho(0,\epsilon)$$
$$+\cot^2\theta\int_{t-2\pi}^{t-2\pi+2\theta}d\epsilon\rho(0,\epsilon)\tan^2\left(\frac{2\pi+\epsilon-t}{2}\right). \qquad (7)$$

We differentiate both parts of Eq. (7) with respect to time and use the constancy of $M$. Then,

$$0=\rho(0,t)-\rho(0,t-2\pi+2\theta)+\cot^2\theta$$
$$\times\left\{\rho(0,t-2\pi+2\theta)\tan^2\left(\frac{2\theta}{2}\right)-\int_{t-2\pi}^{t-2\pi+2\theta}d\epsilon\rho(0,\epsilon)\tan\right.$$
$$\left.\times\left(\frac{2\pi+\epsilon-t}{2}\right)\cos^{-2}\left(\frac{2\pi+\epsilon-t}{2}\right)\right\}. \qquad (8)$$

After cancelling the second and third terms on the right of Eq. (8), we obtain an equation which shows how $\rho(\varphi)$ is transformed during a rotation:

$$\rho(\varphi,t+2\pi)=\cot^2\theta\int_0^{2\theta}d\zeta\rho(\varphi-\zeta,t)\frac{\sin(\zeta/2)}{\cos^3(\zeta/2)} \qquad (9)$$

a discrete linear mapping. As an initial condition for Eq. (9) we take, for example, the distribution $\rho(\varphi,t=2\pi)$ obtained after the first turn of the drum. (For arbitrary $\chi$, describing the state of a system with the distribution $\rho(\varphi,t)$ is generally meaningful for $t\geqslant2\pi-2\theta-\chi$ only!) Finding the distribution $\rho(\varphi,t=2\pi)$ from a given $\rho(\varphi,t=0)$ (see Fig. 1b) is a very simple problem. (We, however, have not had to do this, since in the following we restrict ourselves to examining the

special case of a small amount of the black fraction, $\chi \to 0$, for which $\rho(\varphi)$ can already be specified at the initial time $t=0$.)

The discrete linear mapping (9), along with this sort of initial condition, completely describes the avalanche mixing process when the drum is more than half full. In principle, Eq. (9) is the main result of this paper.

The simplest description of the mixing process can be given for the case of an almost completely filled cylinder, i.e., for $\theta \ll \pi/2$. Then the right hand side can be expanded in $\theta$ and, on introducing the ''discrete'' time $\tilde{t} \equiv 2\pi[t/(2\pi)]$ for convenience (we are describing the slow variation of $\rho(\varphi)$, observing this distribution at times separated by integral periods), we obtain the following equation:

$$2\pi \frac{\partial \rho}{\partial \tilde{t}} = -\frac{4}{3}\theta \frac{\partial \rho}{\partial \varphi} + \theta^2 \frac{\partial^2 \rho}{\partial \varphi^2}. \tag{10}$$

(We could equally well transform to a coordinate system rotating with the cylinder.) Equation (10) has been obtained by expanding the integrand in Eq. (9) as a series in $\zeta$. This equation is valid for $\tilde{t} \gg 2\pi$. Thus, in the regime considered here avalanche mixing reduces to diffusion with convection on the circle $0 \leqslant \varphi \leqslant 2\pi$.

If we assume for simplicity that there is little of the black material ($\chi \ll \theta$), then initially it will all be concentrated at small angles. An examination of this limiting case illustrates the fundamental features of mixing when $\theta \ll \pi/2$. As an initial condition we can take $\rho(\varphi,0) = 4\chi\delta(\varphi)$. In fact, the integral

$$\int_0^{\sin\theta} dr\, r \int d\varphi 4\chi\delta(\varphi) \cong 2\chi\theta^2$$

yields the correct amount of the black fraction in the region of the outer annulus. (For $\chi \ll \theta \ll \pi/2$ it equals the product of the length $2\theta$ of the free surface and the thickness of the black fraction layer at the initial time, $\chi\theta$. See Fig. 1b.) Then we obtain the solution of Eq. (10) in the form

$$\rho(\varphi,\tilde{t}) = 4\chi\left\{\frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} \exp\left(-n^2 \frac{\theta^2}{2\pi}\tilde{t}\right)\cos\left[n\left(\varphi - \frac{2}{3\pi}\theta\tilde{t}\right)\right]\right\} = \frac{4\chi}{2\pi}\,\theta_0\left(\frac{\varphi - 2\theta\tilde{t}/3\pi}{2\pi}, \frac{\theta^2}{2\pi^3}\tilde{t}\right), \tag{11}$$

where $\theta_0$ is the theta function.[15] At late times $\tilde{t} \gg \pi^3/\theta^2$ exponential relaxation takes place, with a characteristic time $t_e = 2\pi/\theta^2$, to a uniform distribution $\rho_\infty = 2\chi\theta^2/\pi\theta^2 = 2\chi/\pi$. (Here $2\pi[1-(1-\theta^2/2)] \cong \pi\theta^2$ is the amount of material in the outer annulus for $\chi \ll \theta \ll \pi/2$.) At smaller times, however, we obtain

$$\rho(\varphi,\tilde{t}) = \frac{4\chi}{\theta\sqrt{2\tilde{t}}} \exp\left\{-\frac{\pi}{2\theta^2\tilde{t}}\left(\varphi - \frac{2\theta\tilde{t}}{3\pi}\right)^2\right\} \tag{12}$$

the standard answer for linear diffusion with convection over an infinite interval, where the center of the distribution drifts at a velocity of $2\theta/3\pi$. (Recall that we are measuring time in radians.)

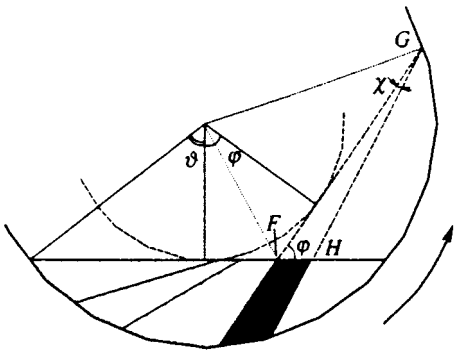



FIG. 2. One of the times $\chi < t < 2\vartheta - \chi$ ($\vartheta \equiv \pi - \theta$) during mixing with a cylinder that is less than half filled.

Here we make one further comment. The location of an individual granule in the outer annulus can be specified by the coordinates $(r,\varphi)$, i.e., the linear coordinate on a tangent of the type $CD$ in Fig. 1a and the corresponding angle. Then the granule changes its ''angle'' by $2\theta$ as it slides over the free surface. The average rate of change of this angle is, therefore, $2\theta/2\pi$. The drift velocity of the peak in the distribution obtained above is close to this value.

### 3. MIXING WHEN THE CYLINDER IS LESS THAN HALF FULL

We now consider the case $\pi/2 < \theta < \pi$, i.e., a cylinder that is less than half full (Fig. 2), where mixing takes place over the entire volume and much more rapidly than in the case examined in Sec. 2. We again use the assumption of complete mixing in avalanches (see Sec. 1), so that $\rho = \text{const}$ for $t > 2(\pi - \theta)$ along all the tangents shown in Fig. 2. Now, as opposed to the case examined above, all these tangents intersect the free surface. Thus, it is convenient for us to describe the time evolution of the mixing process using the quantity $\rho(0,t)$, the fraction of black material at the points on the tangent, which coincides with the free surface at time $t$. We use the notation $\vartheta \equiv \pi - \theta$. (See Fig. 2.)

First we find how $\rho(0,t)$ is transformed in the situation where the state at any point of the material is exhaustively described by $\rho(\zeta,t)$, i.e., when $t > 2\vartheta - \chi$. (The angle $\zeta$ characterizes the position of the corresponding tangent; see Fig. 2.) In this case the total amount of the black fraction in the drum is

$$M = \int_0^{2\vartheta} d\zeta \int_{r(\zeta)}^{\sin\vartheta} dr\, r\rho(\zeta,t) = \frac{1}{2}\int_0^{2\vartheta} d\zeta \rho(\zeta,t) \times\left(\sin^2\vartheta - \cos^2\vartheta \tan^2\frac{\zeta}{2}\right), \tag{13}$$

where $\sin\vartheta$ is the total length of the tangent (from the point of tangency to the intersection with the surface of the drum) and $r(\zeta) = \cos\vartheta \tan(\zeta/2)$ is the distance between the point of tangency and the point of intersection with the free surface for the tangent characterized by the angle $\zeta$. ($\cos\vartheta$ is the length of the normal drawn from the center of the drum to the free surface. See Fig. 2.) Using Eq. (6) and substituting the variable $\epsilon = t - \zeta$, we rewrite Eq. (13) in the form

$$\frac{2M}{\sin^2\vartheta} = -\int_t^{t-2\vartheta} d\epsilon\rho(0,\epsilon)\left[1-\cot^2\vartheta\,\tan^2\left(\frac{t-\epsilon}{2}\right)\right] \tag{14}$$

and differentiate with respect to $t$. Given the constancy of $M$, we obtain an equation for $\rho(0,t)$ when $t>2\vartheta-\chi$:

$$\rho(0,t)=\cot^2\vartheta\int_0^{2\vartheta} d\zeta\rho(0,t-\zeta)\,\frac{\sin(\zeta/2)}{\cos^3(\zeta/2)}. \tag{15}$$

Evidently, for $0\leqslant t\leqslant\chi$ we have $\rho(0,t)=1$. For $\chi<t<2\vartheta-\chi$, on the other hand, on the right hand side of the equation the integral over the region which can be described using $\rho(\zeta,t)$ must be supplemented by an additional term which accounts for the scatter in the granules belonging to the ''pure black'' fraction (see the sector of the black fraction cut off by the free surface of the material in Fig. 2):

$$\rho(0,t)=\cot^2\vartheta\int_0^{t-\chi} d\zeta\rho(0,t-\zeta)\,\frac{\sin(\zeta/2)}{\cos^3(\zeta/2)} + \frac{2}{\sin^2\vartheta}\,\frac{\partial S(t-\chi)}{\partial t}, \tag{16}$$

where $\partial S(\varphi)/\partial\varphi$ is the derivative of the area of the triangle $FGH$ (shown by the dashed lines in Fig. 2), which characterizes the additional contribution to the avalanche from granules belonging to the black fraction. The significance of the angle $\varphi$ is clear from Fig. 2. For $t=\chi$ the angle $\varphi$ goes to zero, so it is understandable why the combination $t-\chi$ appears in Eq. (16).

In order to obtain an expression for $S(\varphi)$ it is necessary to solve a ''high school'' geometry problem. (See Appendix A.) Here we limit ourselves to examining the simplest and most instructive case, $\chi\ll\vartheta\ll\pi/2$ (i.e., a black fraction much smaller than the white and an almost empty drum). The following expression for this derivative in this limit is found using the method given in Appendix A:

$$\frac{\partial S(\varphi)}{\partial\varphi}=\frac{1}{2}\,\frac{4}{3}\,\chi\left(\vartheta+\frac{\varphi}{2}\right). \tag{17}$$

(The derivation of Eq. (17) is not entirely trivial and from the content of Appendix A it is understandable why we preferred to isolate the factors 1/2 and 4/3 in Eq. (17).) As a result, for $\chi\ll\vartheta\ll\pi/2$ we have

$$\rho(0,t)=\frac{1}{\vartheta^2}\int_0^t d\zeta\,\frac{\zeta}{2}\,\rho(0,t-\zeta)+\frac{4}{3}\,\frac{\chi}{\vartheta^2}\left(\vartheta+\frac{t}{2}\right),$$

$$0<t<2\vartheta;\quad \rho(0,t)=\frac{1}{\vartheta^2}\int_0^{2\vartheta} d\zeta\,\frac{\zeta}{2}\,\rho(0,t-\zeta),\quad t>2\vartheta. \tag{18}$$

Now, introducing the ''dimensionless'' quantities $\tau\equiv t/2\vartheta$ and $\nu(\tau)\equiv\rho(0,t)/(4\chi/3\vartheta)$, we obtain

$$\nu(\tau)=2\int_0^\tau d\epsilon(\tau-\epsilon)\,\nu(\epsilon)+1+\tau,\quad 0<\tau<1,$$

$$\nu(\tau)=2\int_{\tau-1}^\tau d\epsilon(\tau-\epsilon)\,\nu(\epsilon),\quad \tau>1. \tag{19}$$

Here the solution of the first equation of the system serves as an initial condition for the second. We notice at once that no system parameters enter explicitly in Eq. (19); they are all implicit in the variable $\tau$ and the quantity $\nu$. The first equation yields $\nu(0)=1$ at the initial time. After termwise differentiation of Eq. (19), we find

$$\nu'(\tau)=2\int_0^\tau d\epsilon\nu(\epsilon)+1,\quad 0<\tau<1,$$

$$\nu'(\tau)=2\int_{\tau-1}^\tau d\epsilon\nu(\epsilon)-2\nu(\tau-1),\quad \tau>1, \tag{20}$$

where $(\ldots)'\equiv d/d\tau$. From the first of Eqs. (20) it follows at once that $\nu'(0)=1$. Differentiating Eq. (19) twice, we obtain the differential-difference equations

$$\nu''(\tau)=2\nu(\tau),\quad 0<\tau<1,$$

$$\nu''(\tau)=2\nu(\tau)-2\nu(\tau-1)-2\nu'(\tau-1),\quad \tau>1, \tag{21}$$

where $\nu(0)=\nu'(0)=1$ at the initial time. A number of books deal with this type of equation.[16–18]

The form of Eqs. (19)–(21) and the equations for the following higher order derivatives imply the following: at the point $\tau=1$ the solution $\nu(\tau)$ has a discontinuity in its value and discontinuities in all derivatives, at the point $\tau=2$ there are discontinuities in all the derivatives (the solution itself is continuous), at the point $\tau=3$ there are discontinuities in all the derivatives of the solution beginning with the second (the solution and its first derivative are continuous), and so on. For example, Eq. (19) implies that $\nu(1+0)-\nu(1-0)=-2$, Eq. (20) implies that $\nu'(2+0)-\nu'(2-0)=2[\nu(1-0)-\nu(1+0)]=4$, etc.

It now remains to solve Eq. (19) (or, for example, Eq. (21)). The solution of the first of Eqs. (21) has the form

$$\nu(\tau)=\frac{1}{2\sqrt{2}}\left[(\sqrt{2}+1)\exp(\sqrt{2}\tau)+(\sqrt{2}-1)\exp(-\sqrt{2}\tau)\right], \tag{22}$$

so that $\nu(1-0)=3.546$, $\nu'(1-0)=4.915$. The second of Eqs. (19) or (21), however, can be solved using the Laplace transform.[19] In Appendix B it is shown how to find the solution

$$\nu(\tau)=\frac{9}{4}+\sum_j\,\frac{2(z_j+1)}{z_j^2(z_j+2)}\exp[z_j(\tau-1)] \times\int_0^1 d\epsilon\nu(\epsilon)[1+z_j(1-\epsilon)-(1+z_j)\exp(-z_j\epsilon)]. \tag{23}$$

Here the sum is taken over all roots (except the zeros) of the characteristic equation $z^2-2+2(z+1)\exp(-z)=0$.[16] The roots of this equation are positioned as follows: there is a triply degenerate zero root that leads to the appearance of the term 9/4, which corresponds to uniform mixing with a density $\rho_\infty=3\chi/\vartheta$ at late times. In fact, the total amount of the black fraction in this case is equal to $2\vartheta-\chi\vartheta$ and of all the material,

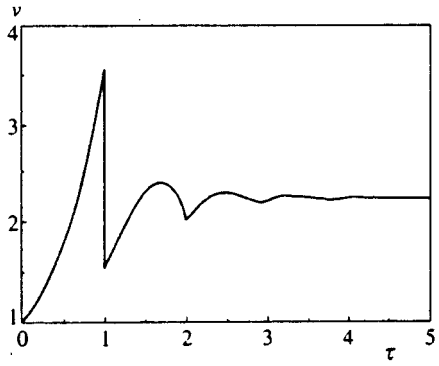



FIG. 3. The density of the black fraction on the left half of the free surface (see Fig. 2) as a function of time in the case when the total volume of material is small compared to that of the empty space and the volume of the black fraction is small compared to that of the white: $\tau \equiv t/(2\vartheta)$, $\nu(\tau) \equiv \rho(0,t)/(4\chi/(3\vartheta))$. The narrow peak at $\tau = 0$ is not indicated. (See the text.)

$$\vartheta - \frac{1}{2}\sin(2\vartheta) \cong \frac{2}{3}\vartheta^3,$$

so that

$$\rho_\infty = \frac{2\chi\vartheta^2}{2\vartheta^3/3} = \frac{3\chi}{\vartheta}.$$

There is also a set of complex conjugate roots with the asymptotic form

$$z_{\pm p} = -\ln(p\pi) \pm (2p - 3/2)\pi i + O(\ln p/p),$$

where $p$ is a positive integer greater than unity. For all $p \geqslant 2$ these values are easily refined by successively iterating the expression $z = \ln[2(z+1)/(2-z^2)]$. Taking the sum in Eq. (23), we obtain the curve shown in Fig. 3. Oscillations similar to these have been observed experimentally.[10] In Fig. 3 one can see the discontinuities in $\nu(\tau)$ and in the derivatives at integral values of $\tau$ of which we spoke above.

Beginning with $\tau \approx 2.5$, the function $\nu(\tau)$ is well fitted when the two lowest nonzero $(p=2)$ roots of the characteristic equation, $z_{\pm 2} = -1.392 \pm 7.553i$, are taken into account. This expression appears to be extremely compact if it is written down directly with the numerical coefficients

$$\nu(\tau) = 9/4 + 0.496\exp[-1.392(\tau - 1)] \times \cos[7.553(\tau - 1) - 1.336]. \quad (24)$$

We note that the period of the cosine oscillations was less than unity here. In the region $1 < \tau < 2$, on the other hand, the answer can be converted to the following analytic expression (see Appendix C):

$$\nu(\tau) = [\nu(1-0) - 2 - 3(\tau-1)/2]\cosh[\sqrt{2}(\tau-1)] + [(\nu'(1-0) - 3/2)/\sqrt{2} - \sqrt{2}(\tau-1)]\sinh[\sqrt{2}(\tau-1)], \quad (25)$$

where the numerical values of $\nu(1-0)$ and $\nu'(1-0)$ have already been given after Eq. (22).

Up to now we have only considered cylinders that are less than half full. It is easy to see, however, that Eq. (16) can be used for the case of a cylinder that is exactly half full. This case is unique in that complete mixing does not occur in this case at all! This can be explained as follows: when the cylinder is half full, the tangents mentioned above become radius vectors of the cylinder. Thus, if the concentration $\rho = \text{const}$ at all points of some radius vector, then even after spilling over the entire left half of the free surface, the same value of the concentration $\rho$ will appear on the entire left half of the free surface. Here it is clear that the distribution pattern of the fractions will repeat itself with a period of $\pi$, i.e., with a period equal to half the rotation period of the drum. Of the two pure fractions, only the black one ''survives,'' since the pure white fraction vanishes by the time $t = \pi - \chi$. (Although we have not shown the mixing schemes specially for this case, it is easy to understand from Fig. 2 how the material begins to mix in this situation.)

It is easy to see that on the left half of the free surface,

$$\rho(0, 0 \leqslant t \leqslant \chi) = \rho(0, \pi - \chi \leqslant t \leqslant \pi) = 1. \quad (26)$$

At times $\chi \leqslant t \leqslant \pi - \chi$ the density $\rho(0,t)$ is given by Eq. (16), in which the first term on the right vanishes since $\theta = \pi/2$:

$$\rho(0,t) = 2\frac{\partial S(t-\chi)}{\partial t} = \frac{\sin^2\chi}{\sin^2 t}. \quad (27)$$

Here we have used an expression for the area

$$S(\varphi) = \frac{1}{2}\frac{\sin\varphi\,\sin\chi}{\sin(\varphi+\chi)},$$

which follows from Eq. (A1) of Appendix A. This set of values of $\rho(0,t)$ together with the periodicity condition $\rho(0,t+\pi) = \rho(0,t)$ provides an exhaustive description of the mixing process when the cylinder is half full.

## 4. LOSS TIMES FOR THE PURE FRACTIONS

We now determine the time when one or the other pure fraction no longer remains in the outer annulus (i.e., outside the circle of radius $\cos\theta$). We denote the loss times for the white and black fractions by $T_w$ and $T_b$. We shall demonstrate how to obtain expressions for these times on the figures.

As noted above, in this section it is no longer necessary to introduce the assumption of complete mixing of the fractions in avalanches, which had to be used previously (Secs. 2 and 3). It will be sufficient to assume that if granules of both types appeared on the right half of the free surface (i.e., in sector $A$ of Fig. 1a), then at each point of the left surface (i.e., at each point of sector $B$ of Fig. 1a) the fractions will exist only in a mixed state. Here the degrees of mixing of the fractions at different places in sector B can be utterly different, but the fractions should not exist anywhere in sector $B$ in their pure forms (if only there is no single pure fraction in sector $A$).

Again, as in the previous sections, the material inside the circle indicated by a dashed curve in Figs. 1, 2, 4, and 5 rotates along with the drum without mixing. As for the region outside the circle of radius $\cos\theta$, after even the first turn of the drum (here we have in mind the case $\theta < \pi/2$) the above assumption implies the following regarding each of the tangents of the type $CD$ indicated in Fig. 1b: either all points on this tangent are in a mixed state or there is a pure

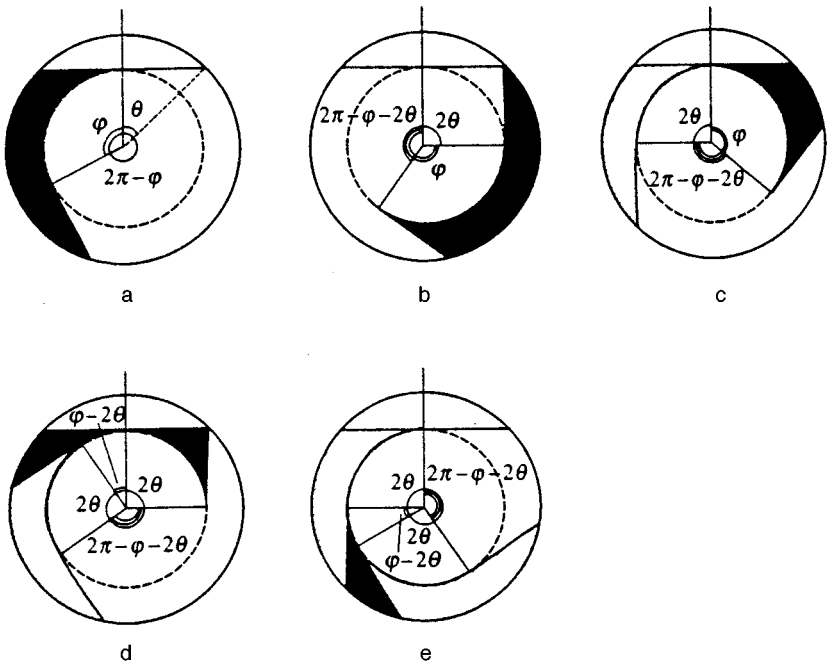



FIG. 4. Sketches of the distribution of the fractions in a cylinder at different times if the fractions are distributed as shown in sketch *a* at time $t_0$: (a) $t=t_0$, (b) $t=t_0+2\pi-2\theta-\varphi$, (c) $t=t_0+2\pi-\varphi$, (d) $t=t_0+2\pi-2\theta$, (e) $t=t_0+2\pi$.

material (white or black) at all its points. In order to find $T_b$ or $T_w$ it is necessary to know whether the material at a given point is in a pure or a mixed state. The degree of mixing of the material at a given place is of no significance to us now, so it is unimportant how the degree of mixing varies along these tangents. The only important thing is on which of these tangents a given point lies. Thus, for our purposes it is again sufficient to compare just the angle between the corresponding radius vector of type $OC$ and the normal $OE$ to the free surface (see Fig. 1b), i.e., the angle $COE$, at each point of the material in the outer annulus. The sizes of the regions in the outer annulus occupied by the different fractions can be measured in terms of the corresponding aperture angles.

Some preliminary discussion is needed. For example, let the situation depicted in Fig. 4a exist at time $t_0$. (The black material in the inner region of the cylinder is not shown, but the ''angle'' $\varphi$ occupied by the black material in the outer annulus is.) Figures 4b–e show how events develop subsequently.

It is easy to see by comparing Figs. 4a and 4d the pure black material vanishes in the outer annulus at the time

$$T_b=t_0+(2\pi-2\theta)\left[\frac{\varphi}{2\theta}\right]+2\pi, \tag{28}$$

i.e., $2\theta$ of the black material vanishes in each rotation by an angle $2\pi-2\theta$ and in order for a residue smaller than $2\theta$ of the black material to be lost in the outer annulus, yet another complete turn of the cylinder is necessary. (Here [...] denotes the integer part.) Note that instead of the white fraction in Fig. 4a, there could be mixed material or any combination of the white fraction and the mixed material; all the same, the answer (28) for $T_b$ remains unchanged.

On comparing Figs. 4a and 4e, we see that the pure white material vanishes from the outer annulus at the time

$$T_w=t_0+2\pi\left[\frac{2\pi-\varphi}{2\theta}\right]+(2\pi-\varphi)-2\theta\left[\frac{2\pi-\varphi}{2\theta}\right]. \tag{29}$$

In this case, $2\theta$ of the white material (which occupies an angle $2\pi-\varphi$ in Fig. 4a) vanishes during each complete turn of the cylinder. The second term on the right hand side of Eq. (29) corresponds to the maximum number of complete turns of the drum for which the white fraction still exists in the outer annulus. The next two terms on the right hand side of Eq. (29) correspond to the angle (less than $2\theta$) through which the drum must be turned in order for the last traces of the white fraction to vanish from the outer annulus. Equation (29) does not change if the black fraction in Fig. 4a is replaced by mixed material or an arbitrary combination of the black fraction and the mixture.

Now we can use Eqs. (28) and (29) to find the final result for the loss time for the pure fractions in the outer annulus. We shall determine the loss time $T_w$ for the white fraction. It will be convenient to start with the simplest limiting case $\chi\to 0$, i.e., with a negligibly small fraction of black material. The material is positioned as shown in Fig. 5a at the initial time $t=0$. Evidently, for $\pi/2<\theta<\pi$, i.e., when the drum is less than half full, $T_w=2(\pi-\theta)$. If, however, the drum is more than half full ($\theta<\pi/2$), we can simply set $\varphi\to 0$ and $t_0=0$ in Eq. (29) (see Fig. 4a for $\varphi\to 0$) and finally get

$$\frac{T_w}{2\pi}=1+\left(1-\frac{\theta}{\pi}\right)\left[\frac{\pi}{\theta}\right] \tag{30}$$

a function containing a sequence of discontinuities. (See Fig. 6).

We now turn to the general case, examined above, of an arbitrary $\chi$, i.e., an arbitrarily thick black layer. First of all, however, it is obvious that we must have $\theta<\pi-\chi$. Next, if $\theta>(\pi-\chi)/2$ (i.e., if there is more of the black material in

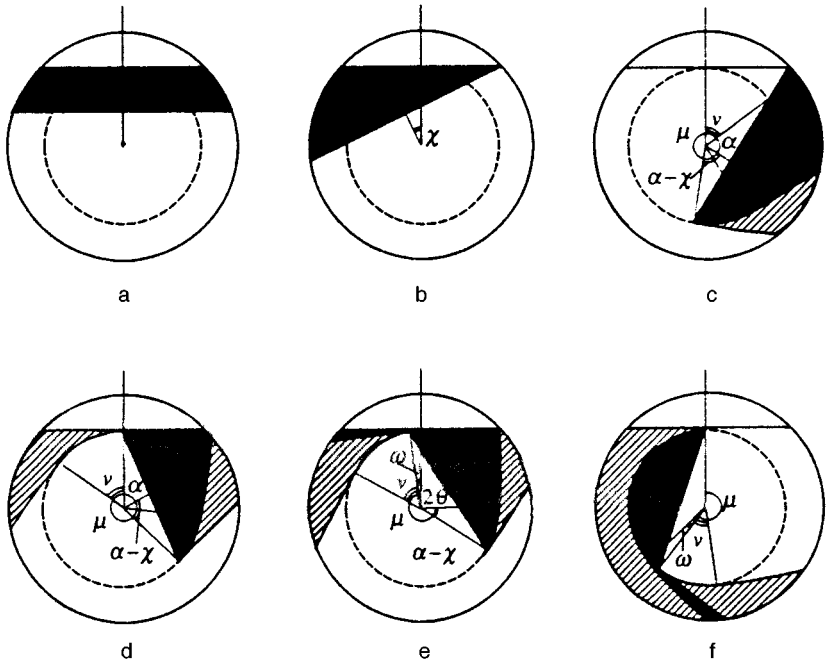



FIG. 5. The form of the mixed material at some times for $\theta<(\pi-\chi)/2$: (a) $t=0$, (b) $t=\chi$, (c) $t=2(\pi-\theta)-\chi$, (d) $t=2/\pi-\alpha$, (e) $t=2(\pi-\theta)+\chi$, (f) $t=2\pi+\alpha$.

the upper half of the drum than in the lower), then it is easy to see that $T_w=2\pi-2\theta-\chi$. If, on the other hand, $\theta>(\pi-\chi)/2$, then $T_w(\theta)$ becomes nontrivial. Some successive times in the initial stage of mixing in this situation are shown in Fig. 5. At the time $2\pi+\alpha$ shown in Fig. 5f the white material in the cylinder is in the same position as in Fig. 4a. (The angle $\alpha$ is introduced in Fig. 1b.) Thus, we can use Eq. (29) if, in it, we set $t_0=2\pi+\alpha$ and the ''angle'' of the white material $2\pi-\varphi=2\pi-(2\theta+\alpha+\chi)$. Ultimately, we have

$$T_w=\alpha+2\pi+2\pi\left[\frac{2\pi-(2\theta+\alpha+\chi)}{2\theta}\right]+2\pi-(2\theta+\alpha+\chi)-2\theta\left[\frac{2\pi-(2\theta+\alpha+\chi)}{2\theta}\right]. \tag{31}$$

Therefore,

$$\frac{T_w}{2\pi}=1-\frac{\chi}{2\pi}+\left(1-\frac{\theta}{\pi}\right)\left[\frac{\pi-(\chi+\alpha)/2}{\theta}\right], \tag{32}$$

where the angle $\alpha$ is expressed in terms of $\theta$ and $\chi$ using Eq. (1). Since we also have $\alpha\to 0$ everywhere except at $\theta=\pi/2$ when $\chi\to 0$ (see Fig. 1b or Eq. (1)), in this limit Eq. (32) transforms to Eq. (30). The set of resulting curves is shown in Fig. 6.

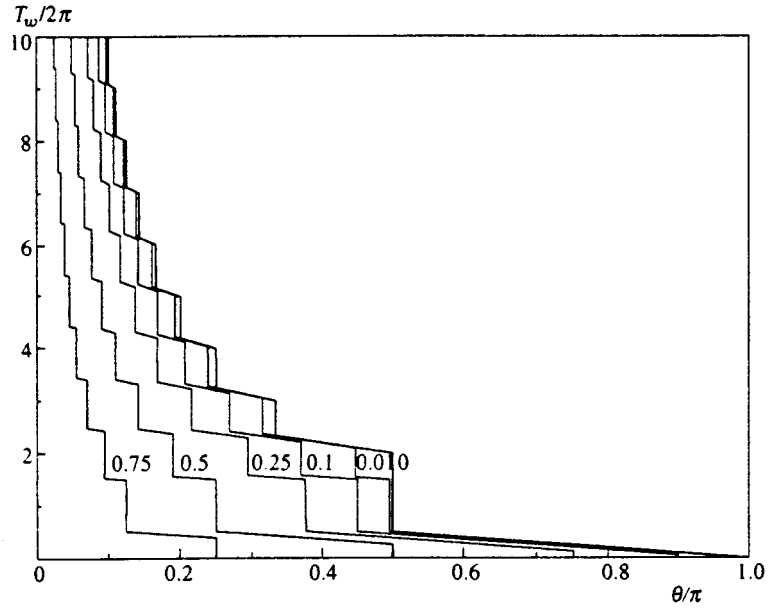



FIG. 6. The time after which none of the pure white fraction remains anywhere in the drum (except, perhaps, in the central region when the drum is more than half filled) as a function of the degree of filling of the drum. Curves are shown for the following values of the parameter $\chi/\pi$ which characterizes the volume of the black fraction: 0, 0.01, 0.1, 0.25, 0.5, 0.75.

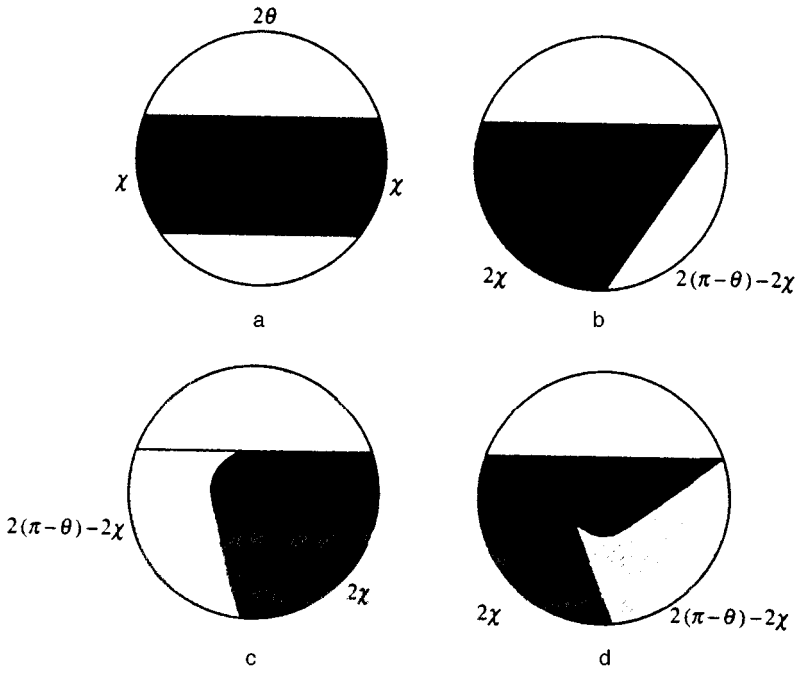



FIG. 7. Some times during mixing for the case $\pi/2-\chi/2\leqslant\theta<\pi/2$: (a) $t=0$, (b) $t=\chi$, (c) $t=2(\pi-\theta)-\chi$, (d) $t=2(\pi-\theta)+\chi$.

We now consider the additional jump to $T_w=3\pi$ which occurs when $\chi\neq 0$. (Compare with the curve for $\chi=0$ in Fig. 6.) This jump takes place at $\theta=\pi/2-\chi/2$, i.e., when the dashed circle in Fig. 1b is tangent to both the upper and lower boundaries of the black fraction. Then the angle $\alpha$ is equal to $\pi/2$, which, as is easily shown, means that $T_w(\theta=\pi/2-\chi/2)=3\pi$.

We can now turn to the question of when the black fraction vanishes everywhere except the inner circle. It is easy to show that for $\theta>\pi/2$, i.e., when the drum is less than half full,

$$T_b=2\pi-2\theta+\chi. \tag{33}$$

Now let $\pi/2-\chi/2\leqslant\theta<\pi/2$. (See Fig. 7). As can be seen from Fig. 7d, at the time $2(\pi-\theta)+\chi$ the fractions are positioned such that Eq. (29) can be used on setting $t_0=2(\pi-\theta)+\chi$ and $\varphi=2\chi$ in it (cf. Fig. 4a). As a result, for these values of $\theta$ we have

$$\frac{T_b}{2\pi}=2-\frac{\theta}{\pi}-\frac{\chi}{2\pi}+\left(1-\frac{\theta}{\pi}\right)\left[\frac{\chi}{\theta}\right]. \tag{34}$$

Finally, we turn to the case $\theta\leqslant\pi/2-\chi/2$. (See Fig. 5.) Comparing Figs. 5e and 4a, we see that it is again possible to use Eq. (29) on setting $t_o=2(\pi-\theta)+\chi$ and $\varphi=\chi+\alpha-2\theta$ in it. Thus,

$$\frac{T_b}{2\pi}=2-\frac{\theta}{\pi}-\frac{\chi}{2\pi}+\left(1-\frac{\theta}{\pi}\right)\left[\frac{\chi+\alpha-2\theta}{2\theta}\right]. \tag{35}$$

Equations (34) and (35) agree with one another at the point $\theta=\pi/2-\chi/2$. In fact, at this point $\alpha=\pi$ and, therefore, Eq. (35) transforms to Eq. (34).

Equations (33)–(35), together with Eq. (1), provide a description of the curves shown in Fig. 8. In the limit $\chi\to 0$ Eq. (35) transforms to $T_b=2\pi$ for angles $\theta<\pi/2$, since then $\alpha\to 0$. In the same limit Eq. (34) yields $T_b=3\pi$ for $\theta=\pi/2-0$. We see, therefore, that the $T_b(\theta)$ curves have a minimum for small $\chi$. (See Fig. 8.) It can be verified that this minimum exists for $\chi<\pi/6$. It vanishes for $\chi=\pi/6$ at the point $\theta=\pi/3$. Equation (35) does not give a unique result when the drum is exactly half filled. This case, in which complete mixing does not occur, requires special treatment. It was described at the end of the previous section.

## 5. CONCLUSION

We have, therefore, explained the nature of avalanche mixing. It turned out to be sufficient to apply an essentially simple geometric approach for an analytic description of the model that was used. It turned out that in a number of cases the mixing process can be described as a sequence of discrete linear mappings. We have shown that the problem has an extremely nontrivial solution, although it does not involve complicated nonlinear mappings.[20,21] Two regimes of avalanche mixing have been described: a diffusive regime with convection when the cylinder is more than half full and a rapid mixing regime when it is less than half full.

We have calculated the characteristic mixing times, including the times over which the pure fractions vanish in the regions adjacent to the surface of the cylinder. These times are a discontinuous function of the degree of filling of the center and consist of a series of jumps. Singularities of this type are generally characteristic of a wide range of systems with phase slippage centers, which include, in particular, Josephson structures, thin films of second order superconductors, etc.[22–24] This concept can also be introduced for the system examined here. We have actually used only one quantity for describing the distribution of the granules—the angle between the radius vector of the corresponding tangent and the normal to the free surface. (See Fig. 1b.) This angle, the ''phase'' of the granules, varies discontinuously by $2\theta$

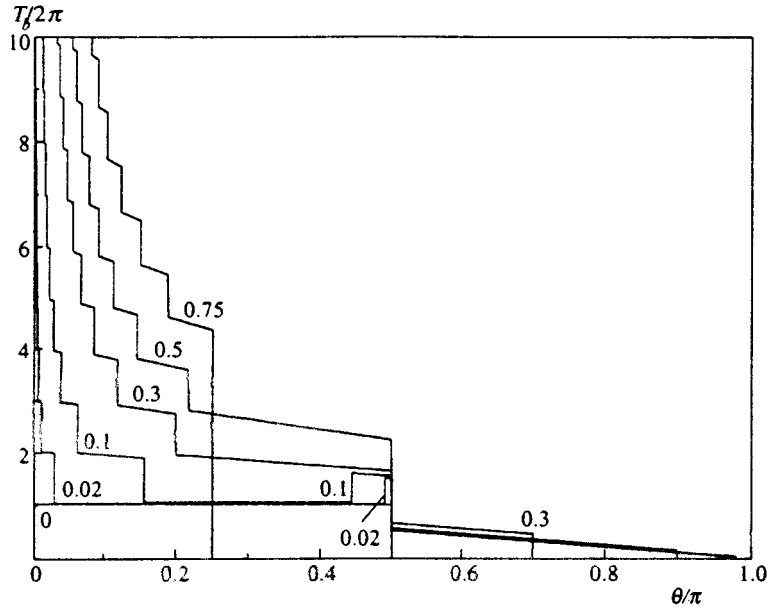



FIG. 8. The time after which none of the pure black fraction remains anywhere in the drum (except, perhaps, in the central region when the drum is more than half filled) as a function of the degree of filling of the drum. Curves are shown for the following values of the parameter $\chi/\pi$ which characterizes the volume of the black fraction: 0, 0.02, 0.1, 0.3, 0.5, 0.75. The value of $T_b(\theta=\pi/2)=\infty$ is not shown.

every time the granules slide along the free surface. Thus, in our problem the free surface of the granular material serves as a unique ''phase slippage center'' in our problem.

I am grateful to E. N. Antonov, V. V. Bryksin, S. A. Ktitorov, E. K. Kudinov, A. M. Monakhov, A. N. Samukhin, and Yu. A. Firsov for numerous, helpful discussions. Part of the results presented here were obtained while the author was at the School on Problems in Self-Organized Criticality at Turin financed by the INTAS program.

## APPENDIX A

### The form of the source in Eq. (16)

We shall find the derivative $\partial S(\varphi)/\partial\varphi$, which appears in the left hand side of Eq. (16). Here $S(\varphi)$ is the area of the triangle $FGH$ indicated by the dashed lines in Fig. 2. The significance of the angle $\varphi=t-\chi$ is also clear from this figure.

Evidently, the side of the triangle is $FG=\sin\vartheta+\tan(\varphi/2)\cos\vartheta$. Using the standard formulas for the area of a triangle, we obtain

$$S(\varphi)=\frac{1}{2}\frac{\sin\varphi(FG)^2\sin\chi}{\sin(\pi-\varphi-\chi)}=\frac{1}{2}\left[\sin\vartheta+\tan\frac{\varphi}{2}\cos\vartheta\right]^2\frac{\sin\varphi\sin\chi}{\sin(\pi-\varphi-\chi)}\cong\frac{1}{2}\left(\vartheta+\frac{\varphi}{2}\right)^2\frac{\varphi\chi}{\varphi+\chi}. \tag{A1}$$

The last relation was obtained after a transition to the limit $\chi,\varphi<\vartheta\ll 1$. In this case, we have

$$\frac{\partial S(\varphi)}{\partial\varphi}=\frac{\chi}{2}\frac{\vartheta+\varphi/2}{\varphi+\chi}\left[\varphi+\chi\frac{\vartheta+\varphi/2}{\varphi+\chi}\right]. \tag{A2}$$

This last function has a narrow (if $\chi\ll\vartheta$) peak of width $\chi$ at the point $\varphi=0$ (the maximum of $\partial S(0)/\partial\varphi=\vartheta^2/2$) and a minimum equal to $\chi\vartheta/2$ at the point $\varphi=\chi$ and is varies linearly as $\chi(\vartheta+\varphi/2)/2$ for $\varphi\gg\chi$.

Thus, if there is little of the black material (i.e., if $\chi\ll\vartheta\ll 1$) then at times $t<2\vartheta$ the mixing pattern is as follows: a black ''leader'' with an angular width of $\chi$ moves in front, a very dark narrow ''tail,'' also of width $\chi$, stretches out behind it, and this is followed by a much less dark region. In the limit $\chi\to 0$ we can join the narrow ''tail'' to the black leader and for $\varphi>0$ deal only with the linear portion of $\partial S/\partial\varphi$. Then, in order to obtain the correct answer for the amount of the black fraction

$$S(\varphi=2\vartheta)=\int_0^{2\vartheta}d\varphi\,\frac{\partial S}{\partial\varphi}=2\vartheta^2\chi,$$

it is necessary to renormalize $\chi$ with a factor of 4/3. As a result, we obtain Eq. (17) for $\partial S(\varphi)/\partial\varphi$, which is valid for $\chi\ll\varphi<\vartheta\ll 1$.

## APPENDIX B

### Solution of Eq. (19)

Since the standard formulas for this type of equation from the classical handbook of Pinii[16] lead to incorrect results, we must briefly describe in general the entirely conventional procedure for solving Eq. (19) using Laplace transforms.

We shall solve the second of the system of Eqs. (19) over the interval $1<\tau<\infty$, assuming that the solution $\nu(\tau)$ is already known in the interval (0,1). (See Eq. (22), which is obtained from the first equation of this system.) We apply the Laplace transform

$$y(z)=\int_0^\infty d\gamma e^{-z\gamma}\nu(\gamma+1) \tag{B1}$$

to both parts of Eq. (19):

$$\nu(\gamma+1)=2\int_{\gamma-1}^{\gamma}d\epsilon(\gamma-\epsilon)\nu(\epsilon+1)=2\int_{0}^{\gamma}d\epsilon(\gamma-\epsilon)\nu(\epsilon+1)-2\int_{0}^{\gamma-1}d\epsilon(\gamma-1-\epsilon)\nu(\epsilon+1)+2\int_{0}^{\gamma-1}d\epsilon\nu(\epsilon+1), \tag{B2}$$

where $\gamma$ belongs in the interval $(0,\infty)$. Now it is necessary to use the standard formulas

$$\frac{y(z)}{z}=\int_{0}^{\infty}d\gamma e^{-z\gamma}\int_{0}^{\gamma}d\epsilon\nu(\epsilon+1),$$

and

$$\frac{y(z)}{z^2}=\int_{0}^{\infty}d\gamma e^{-z\gamma}\int_{0}^{\gamma}d\epsilon(\gamma-\epsilon)\nu(\epsilon+1)$$

and carry out successive integrations by parts, separating and collecting the terms with $y(z)$. We then obtain

$$y(z)=\frac{2}{z^2-2+2(z+1)\exp(-z)}\int_{0}^{1}d\epsilon\nu(\epsilon)[1+z(1-\epsilon)-(1+z)e^{-z\epsilon}]. \tag{B3}$$

It follows immediately from the limiting value theorem that

$$\nu(\tau\to\infty)=\lim_{z\to 0}zy(z)=9/4.$$

Taking the inverse Laplace transform[19]

$$\nu(\tau+1)=\frac{1}{2\pi i}\int_{C+i\infty}^{C-i\infty}dze^{z\tau}y(z)=\sum_{j}\ \mathrm{Res}_j[e^{z\tau}y(z)] \tag{B4}$$

(the sum of the residues is taken over all poles of the function (3); since there are none of these in the right half plane $\{z\}$, $C$ can be taken to be any positive real constant) and then separating a factor of 9/4 from the third-order pole at the point $z=0$ yields the final formula (23).

## APPENDIX C

### Analytic representation of the solution of Eqs. (19) in the region $1<\tau<2$

We have pointed out that the solution $\nu(\tau)$ of these equations can be written in a very simple analytic form within the interval $1<\tau<2$. We now show how this can be done.

We calculate all the derivatives $\nu^{(m)}(\tau)$ at the point $\tau=1+0$. Equations (19) immediately imply that $\nu(1+0)=\nu(1-0)-2$, Eq. (20) implies that $\nu'(1+0)=\nu'(1-0)-2\nu(0)-1=\nu'(1-0)-3$, Eq. (21) implies that $\nu''(1+0)=2[\nu(1+0)-\nu(0)-\nu'(0)]=2[\nu(1-0)-4]$, and so on. Evidently, all the derivatives $\nu^{(m)}(1+0)$ are expressed only in terms of $\nu(0)=\nu'(0)=1$ and the known (see Eq. (22)) $\nu(1-0)$ and $\nu'(1-0)$. The following general results for the derivatives can be obtained for the even ones ($k=0,1,2,\ldots$):

$$\nu^{(2k)}(1+0)=2^k[\nu(1-0)-2(k+1)], \tag{C1}$$

and for the odd ($k=0,1,2,\ldots$):

$$\nu^{(2k+1)}(1+0)=2^k[\nu'(1-0)-3(k+1)]. \tag{C2}$$

Finally, on the interval $1<\tau<2$ we have

$$\nu(\tau)=\sum_{k=0}^{\infty}\frac{2^k}{(2k)!}[\nu(1-0)-2(k+1)](\tau-1)^{2k}+\sum_{k=0}^{\infty}\frac{2^k}{(2k+1)!}[\nu'(1-0)-3(k+1)](\tau-1)^{2k+1}. \tag{C3}$$

These series immediately sum to the final analytic expression with sinh and cosh in Eq. (25).

*e-mail: sn@dor.ioffe.rssi.ru


[1] J. Rajchenbach, Phys. Rev. Lett. **65**, 2221 (1990).
[2] V. Buchholtz, Th. Pöschel, and H.-J. Tillemans, Physica A **216**, 199 (1995).
[3] G. H. Ristow, Europhys. Lett. **28**, 97 (1994).
[4] G. A. Kohring, J. de Phys. I (France) **5**, 1551 (1995).
[5] G. Baumann, I. M. Jánosi, and D. E. Wolf, Phys. Rev. E **51**, 1879 (1995).
[6] F. Cantelaube, Y. Limon-Duparcmeur, D. Bideau, and G. H. Ristow, J. de Phys. I (France) **5**, 581 (1995).
[7] O. Zik, D. Levine, S. G. Lipson, S. Shrikman, and J. Stavans, Phys. Rev. Lett. **73**, 644 (1994).
[8] K. M. Hill and J. Kakalios, Phys. Rev. E **49**, R3610 (1993).
[9] Th. Pöschel and V. Buchholtz, Phys. Rev. Lett. **71**, 3963 (1993).
[10] G. Metcalfe, T. Shinbrot, J. J. McCarthy, and J. M. Ottino, Nature (London) **374**, 39 (1995).
[11] S. N. Dorogovtsev, JETP Lett. **62**, 262 (1995).
[12] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett. **59**, 381 (1987).
[13] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. A **38**, 364 (1988).
[14] C. Tang and P. Bak, Phys. Rev. Lett. **60**, 2347 (1988).
[15] N. I. Akhiezer, *Elements of the Theory of Elliptic Functions*, Amer. Math. Soc., Providence, 1990.
[16] E. Pinney, *Ordinary Differential-Difference Equations*, Univ. of California, Berkeley, 1958.
[17] A. A. Mirolyubov and M. A. Soldatov, *Linear Homogeneous Difference Equations* [in Russian], Nauka, Moscow (1981).
[18] A. A. Mirolyubov and M. A. Soldatov, *Linear Inhomogeneous Difference Equations* [in Russian], Nauka, Moscow (1986).
[19] V. A. Ditkin and A. P. Prudnikov, *Integral Transforms and Operational Calculus*, Pergamon, Oxford, 1966.
[20] J. D. Meiss, Rev. Mod. Phys. **64**, 795 (1992).
[21] R. Z. Sagdeev, D. A. Usikov, and G. M. Zaslavsky, *Nonlinear Physics: from the Pendulum to Turbulence and Chaos*, Harwood, Chur, 1988.
[22] A. A. Abrikosov, *Fundamentals of the Theory of Metals*, North-Holland, Amsterdam, 1988.
[23] V. V. Bryksin, A. V. Gol'tsev, and S. N. Dorogovtsev, JETP Lett. **49**, 503 (1989).
[24] V. V. Bryksin, A. V. Goltsev, and S. N. Dorogovtsev, Physica C **160**, 103 (1989).


Translated by D. H. McNeill

# The mass of a large polaron

A. É. Myasnikova and É. N. Myasnikov

*Rostov Teachers' Training University, 344092 Rostov-on-Don, Russia*
(Submitted 1 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 278–283 (July 1997)

We show that, beginning with the works of L. D. Landau and S. I. Pekar, the effective mass of a large polaron has been determined with a crucial error. Since all such research ignored the spatial dispersion of the lattice polarizability, the maximum group velocity of phonons is found to be zero, so that the phonon ''cloud'' of a polaron is unable to follow the polaron. We allow for the spatial dispersion of the lattice polarizability and derive an expression for the effective polaron mass valid over the entire velocity range in which a polaron can exist: from zero to the maximum group velocity of phonons. According to this expression, the polaron mass depends not only on the phonon frequency, reciprocal effective dielectric constant, and the carrier mass but also on the maximum group velocity of phonons interacting with the carrier and on the polaron velocity. © *1997 American Institute of Physics.* [S1063-7761(97)02307-X]

## 1. INTRODUCTION

Polaron motion is the coordinated motion of two wave packets: of the charge carrier and of lattice polarization. Naturally, this is possible if the group velocities of these wave packets coincide and are nonzero. However, if one assumes that the frequency of the lattice polarization waves is independent of their wave vector, a common approximation in polaron theory beginning with the fundamental work of Landau and Pekar,[1,2] the group velocity of the polarization wave is found to be zero, which means that it is impossible in this approximation to study polaron motion. If we assume that $\Omega=0$ and calculate the polarization induced in the medium by a moving localized charge, then, as shown in Ref. 3, the polarization, rather than being localized, corresponds to real polarization waves of the Cherenkov-radiation type emitted by the charge. This fact is in full agreement with the results of Landau's theory of a quantum Bose liquid,[4,5] where the motion of the Bose vacuum with a velocity higher than the minimum phase velocity with respect to the subsystem linked to the coordinate system is accompanied by spontaneous emission. The destruction of the polaron's polarization ''cloud'' due to fact that the polaron velocity is higher than the minimum phase velocity of the phonons comprising the cloud has been demonstrated in Ref. 6 via a model of a medium with two phonon branches interacting with the carrier and having different minimum phonon phase velocities. In such a medium a polaron at rest has two polarization clouds formed by phonons of two types. If the branch with the higher minimum phonon phase velocity is capable of ensuring the localization of the carrier, when the two-cloud polaron reaches the lower of the two velocities it loses the cloud formed by the lower-velocity phonons, as shown in Ref. 6. Here, although the carrier remains in the autolocalized state, the polaron energy, mass, and total polarization charge change. When the polaron reaches the higher minimum phonon phase velocity, the carrier becomes delocalized, since the equations of motion of carrier and polarization have no autolocalized stationary solutions at such velocities. Thus, a polaron can exist only within a limited interval of velocities, $v<u$, where $u$ is the minimum phase velocity of the phonons participating in the localization of the carrier (or the higher of the two minimum phonon phase velocities if the medium has several phonon branches capable of autolocalizing the carrier).

An important aspect of the theory of a moving polaron is the effective mass of the polaron. On the basis of the above reasoning we conclude that calculating the effective polaron mass correctly is possible only if we allow for the dependence of the polarization wave frequency on the wave vector, or for the spatial dispersion of lattice polarizability. Landau and Pekar[1] obtained their formula for the effective polaron mass without taking this dependence into account, i.e., assuming that the group velocity of the polarization waves is zero. In this case the motion of a polaron as a whole is impossible, but there was no way Landau and Pekar[1] could have noticed this since they used an impermissible expansion of the pole function $c(\omega)=\text{const}/(\omega^2-\Omega^2)$ in a Taylor series in powers of $\omega^2/\Omega^2$ ($\Omega$ is the frequency of longitudinal optical phonons).

Davydov and Énol'skiĭ[7] were the first to study the problem of the effective polaron mass with allowance for the spatial dispersion of lattice polarizability. However, with the approximation they adopted they were able to derive only a formula for the effective polaron mass valid for low phonon dispersion (characterized by a maximum phonon group velocity $u$) and phonon velocity $v$, i.e., $u,v\ll a\Omega$ ($a$ is the lattice constant), or $u,v\ll 25\times 10^4$ cm s$^{-1}$ (see Ref. 7). Nevertheless, their theory contained no internal contradictions, since the polaron velocity $v$ was less than $u$.

Below we rigorously allow for the dependence of the polarization wave frequency on the wave vector and derive formulas for calculating the effective polaron mass valid for arbitrary phonon dispersion (i.e., arbitrary $u$) and an arbitrary polaron velocity $v<u$.

## 2. POLARON ENERGY WITH ALLOWANCE FOR PHONON DISPERSION, AND THE EFFECTIVE POLARON MASS

The Hamiltonian of a system consisting of the medium and the carrier has the following form:[6,7]

 1063-7761/97/070152-04$10.00  

$$H_0+H_{\rm pol}+H_{\rm int}=\int d^3\mathbf{r}\left\{\nabla_{\mathbf{r}}\psi^2+\frac{2\pi}{c\Omega^2}\left[\Omega^2P^2+\left(\frac{\partial\mathbf{P}}{\partial t}\right)^2-u^2\mathbf{P}\cdot\nabla_r^2\mathbf{P}\right]-\mathbf{P}\cdot\mathbf{D}\right\},\tag{1}$$

$$\mathbf{D}=-e\nabla_{\mathbf{r}}\int\psi^2(\mathbf{r},t)\frac{d^3\mathbf{r}}{|\mathbf{r}-\mathbf{r}'|},$$

where $\psi(\mathbf{r},t)$ is the carrier wave function (assumed real), $\mathbf{P}$ is the polarization vector, $\Omega$ is the frequency of longitudinal optical phonons at the center of the Brillouin zone, $c$ is the reciprocal effective dielectric constant ($c=1/\varepsilon_\infty-1/\varepsilon_0$), and $u$ is a quantity to which the phase and group velocity of the phonons participating in the localization of the carrier asymptotically converge. We examine the variation of the polarization self−energy and the energy of interaction of the carrier and the polarization field, $H_{\rm pol}+H_{\rm int}$, caused by variation of the polaron velocity. If we allow for the equation of motion for $\mathbf{P}$, the corresponding part of the Hamiltonian (1) is

$$H_{\rm pol}+H_{\rm int}=-\frac{1}{2}\int\mathbf{P}\cdot\mathbf{D}d^3\mathbf{r}+\frac{2\pi}{c\Omega^2}\times\int d^3\mathbf{r}\left[\left(\frac{\partial\mathbf{P}}{\partial t}\right)^2-\mathbf{P}\cdot\frac{\partial^2\mathbf{P}}{\partial t^2}\right].\tag{2}$$

The polarization charge density can be expressed in terms of the free charge density at time zero, $\psi^2(r)$, via the Green's function $G(\mathbf{r}-\mathbf{r}',t)$ (see Ref. 6),

$$\rho(\mathbf{r},t)=ec\Omega^2\int G(\mathbf{r}-\mathbf{r}',t)\psi^2(\mathbf{r}')d^3\mathbf{r}',$$

so that the polarization vector has the form

$$\mathbf{P}(\mathbf{r},t)=\frac{ec\Omega^2}{4\pi}\nabla_{\mathbf{r}}\int\frac{G(\mathbf{r}_1-\mathbf{r}_2,t)\psi^2(\mathbf{r}_2)d^3\mathbf{r}_1d^3\mathbf{r}_2}{|\mathbf{r}-\mathbf{r}_1|}.\tag{3}$$

We select a system of cylindrical coordinates with the $z$ axis parallel to the polaron velocity $v$. Then the Green's function $G(r,t)$ obtained in Ref. 3 has the following form:

$$G_i(\mathbf{r},t)=\begin{cases}\dfrac{\exp(-\Omega_i[(z-vt)^2/\beta_{1i}^2+r^2]^{1/2}/u_i)}{4\pi u_i^2\beta_{1i}[(z-vt)^2/\beta_{1i}^2+r^2]^{1/2}},\\ v<u_i,\quad \beta_{1i}^2=1-\dfrac{v^2}{u_i^2},\\ \dfrac{\cos(\Omega_i[(z-vt)^2/\beta_{2i}^2-r^2]^{1/2}/u_i)}{2\pi u_i^2\beta_{2i}[(z-vt)^2/\beta_{2i}^2-r^2]^{1/2}},\\ v>u_i,\quad \begin{cases}z-vt<0,\\ r<|z-vt|/\beta_{2i},\end{cases}\\ 0,\quad v>u_i,\quad \begin{cases}|z-vt|<0,\\ r>|z-vt|/\beta_{2i},\end{cases}\\ |z-vt|>0,\quad \beta_{2i}^2=\dfrac{v^2}{u_i^2}-1,\end{cases}\tag{4}$$

where $i=1,2$. Plugging (3) and the expression for $\mathbf{D}(\mathbf{r},t)$ in terms of $\psi^2(\mathbf{r},t)$ into (2) and transforming the result, we arrive at the following formula:

$$H_{\rm pol}+H_{\rm int}=\frac{e^2c\Omega^2}{2}\int\frac{d^3\mathbf{r}d^3\mathbf{r}_1d^3\mathbf{r}_2}{|\mathbf{r}-\mathbf{r}_1|}\left\{\psi^2(\mathbf{r}_2)\psi^2(\mathbf{r})G(\mathbf{r}_1-\mathbf{r}_2,t)+\int d^3\mathbf{r}_3\psi^2(\mathbf{r}_2)\psi^2(\mathbf{r}_3)\left[\frac{\partial}{\partial t}G(\mathbf{r}_1-\mathbf{r}_2,t)\frac{\partial}{\partial t}G(\mathbf{r}-\mathbf{r}_3,t)-G(\mathbf{r}_1-\mathbf{r}_2,t)\frac{\partial^2}{\partial t^2}G(\mathbf{r}-\mathbf{r}_3,t)\right]\right\}.\tag{5}$$

We seek the effective polaron mass in the form $m^{**}=2(E(v)-E(0))/v^2$ (Ref. 7), where $E$ is the polaron energy. Bearing in mind that in (5) only the Green's function $G(\mathbf{r},t)$ depends on the polaron velocity $v$, we can easily obtain the following expression for $m^{**}$:

$$m^{**}=\frac{e^2c\Omega^2}{v^2}\int\frac{d^3\mathbf{r}d^3\mathbf{r}_1d^3\mathbf{r}_2}{|\mathbf{r}-\mathbf{r}_1|}\times\left\{\psi^2(\mathbf{r}_2)\psi^2(\mathbf{r})[G(\mathbf{r}_1-\mathbf{r}_2,v,t)-G(\mathbf{r}_1-\mathbf{r}_2,0,t)]+\int d^3\mathbf{r}_3\psi^2(\mathbf{r}_2)\psi^2(\mathbf{r}_3)\left[\frac{\partial}{\partial t}G(\mathbf{r}_1-\mathbf{r}_2,v,t)\frac{\partial}{\partial t}G\times(\mathbf{r}-\mathbf{r}_3,v,t)-\frac{\partial}{\partial t}G(\mathbf{r}_1-\mathbf{r}_2,0,t)\frac{\partial}{\partial t}G(\mathbf{r}-\mathbf{r}_3,0,t)-G(\mathbf{r}_1-\mathbf{r}_2,v,t)\frac{\partial^2}{\partial t^2}G(\mathbf{r}-\mathbf{r}_3,v,t)+G(\mathbf{r}_1-\mathbf{r}_2,0,t)\frac{\partial^2}{\partial t^2}G(\mathbf{r}-\mathbf{r}_3,0,t)\right]\right\}.\tag{6}$$

Clearly, calculating the numerical values of the effective polaron mass by (6) is difficult. Note, however, that the integral in (6) is of the convolution type, which can be simplifying by introducing Fourier transforms. It is more convenient to introduce Fourier transforms in Eq. (2). The integrands in both integrals on the right−hand side of Eq. (2) are products of two functions of the radius vectors. By expanding each in a Fourier series and integrating over the radius vector we can transform (2) to the following form:

$$H_{\rm pol}+H_{\rm int}=\frac{1}{2}\int\mathbf{P}_k\cdot\mathbf{D}_k\frac{d^3\mathbf{k}}{(2\pi)^3}+\frac{4\pi}{c\Omega^2}\int\frac{d^3\mathbf{k}}{(2\pi)^3}k_z^2v^2P_k^2,\tag{7}$$

where $\mathbf{P}_k$ is the Fourier transform of the polarization vector (we have also allowed for the fact that $\mathbf{D}_{-k}=-\mathbf{D}_k$ and $\mathbf{P}_{-k}=-\mathbf{P}_k$). Clearly, (3) is a convolution in the variables $\mathbf{r}_1$ and $\mathbf{r}_2$. Using (3) and the property of Fourier transforms of convolutions, we can write the Fourier transform $\mathbf{P}_k$ of the polarization vector as a product of the Fourier transforms of the functions $\nabla_{\mathbf{r}}(1/|\mathbf{r}|')G(\mathbf{r},t)$ and $\psi^2(\mathbf{r},t)$:

$$\mathbf{P}_k=\frac{ec\Omega^2}{4\pi}\frac{4\pi i\mathbf{k}}{k^2+k_z^2}\frac{1}{k_z^2(v^2-u^2)-k^2u^2-\Omega^2}\psi_k^2,\tag{8}$$

where $\psi_k^2$ is the Fourier transform of the square of the carrier's wave function. Plugging $\mathbf{P}_k$ specified by (8) and $\mathbf{D}_k$ found in a similar manner into (7), we arrive at a formula for the dependence of the energy considered here on the polaron velocity:

$$H_{\rm pol}+H_{\rm int}=\frac{e^2c\Omega^2}{\pi u^2}\int_0^\infty \frac{k\,dk\,dk_z}{k^2+k_z^2}\frac{(\psi_k^2)^2}{k_z^2(1-v^2/u^2)+k^2\Omega^2/u^2}\times\left(1-\frac{2k_z^2v^2/u^2}{k_z^2(1-v^2/u^2)+k^2+\Omega^2/u^2}\right). \quad (9)$$

The expression for the effective polaron mass obtained in the usual way via (9) is

$$m^{**}=\frac{2e^2c\Omega^2}{\pi u^2v^2}\int_0^\infty \frac{k\,dk\,dk_z}{k^2+k_z^2}(\psi_k^2)^2\left\{\frac{1}{k_z^2+k^2+\Omega^2/u^2}-\frac{k_z^2(1-3v^2/u^2)+k^2+\Omega^2/u^2}{(k_z^2(1-v^2/u^2)+k^2+\Omega^2/u^2)^2}\right\}. \quad (10)$$

For polaron velocities low compared to the minimum phonon phase velocity $u$ ($v\ll u$) the expression for the polaron mass is simpler:

$$m^{**}_{v\ll u}=\frac{2e^2c\Omega^2}{\pi u^4}\int_0^\infty \frac{k\,dk\,k_z^2dk_z}{k^2+k_z^2}\frac{(\psi_k^2)^2}{k_z^2+k^2+\Omega^2/u^2}. \quad (11)$$

### 3. DISCUSSION

Equation (10) shows that the effective mass of a large polaron depends not only on such medium parameters as the optical phonon frequency $\Omega$, the reciprocal effective dielectric constant $c$, and the ratio $m^*/m_e$, as follows from the formula of Landau and Pekar,[1] or the coupling constant (determined by the same parameters), as, say, stated in Refs. 8 and 9, but also on the minimum phase velocity $u$ of the optical phonons participating in the localization of the carrier and on the polaron velocity. More than that, the energy of a polaron at rest proves to be dependent on the minimum phase velocity of the phonons, which is quite natural (see (9)) since, as the expression (4) for the Green's function implies, the spread of the distribution of the polarization charge is proportional to $u/\Omega$.

Figure 1 depicts the dependence of the effective polaron mass on the polaron velocity $v$ for different values of the minimum phonon phase velocity (the ratio $v/u$ is laid off on the horizontal axis) in the limited velocity interval $v<u$ in which a polaron can exist. The polaron mass was calculated by formula (10) for the following values of the medium parameters: $\Omega=360\ \text{cm}^{-1}$, $c=0.27$, and $m^*/m_e=1$. The wave function was taken in the form of the Pekar wave function[1,2]

$$\psi(\mathbf{r},t)=\frac{\alpha^{3/2}}{\sqrt{7\pi}}(1+\alpha r)\exp(-\alpha r),$$

where the values of $\alpha$ (the localization parameter) were obtained by minimizing the polaron's energy functional. Curves *1*, *2*, and *3* correspond to the following values of the minimum phonon phase velocity: $u_1=2\times10^6$ cm/s, $u_2=10^6$ cm/s, and $u_3=5\times10^5$ cm/s. Clearly, the effective

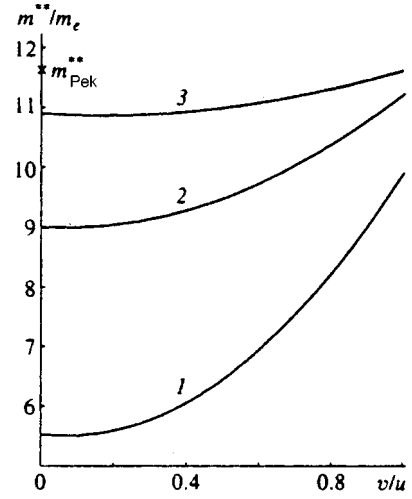



FIG. 1. The polaron-velocity dependence of the effective polaron mass calculated by Eq. (10) with the following values of the medium parameters: $\Omega=360\ \text{cm}^{-1}$, $c=0.27$, and $m^*/m_e=1$. Curves *1*, *2*, and *3* correspond to the following values of the minimum phonon phase velocity: $u_1=2\times10^6$ cm/s, $u_2=10^6$ cm/s$^{-1}$, and $u_3=5\times10^5$ cm/s. The point on the vertical axis ($v=0$) denoted by $m^{**}_{\rm Pek}$ stands for the value of the effective mass obtained by Landau and Pekar's formula[1,2] for the given values of the medium parameters.

polaron mass grows as the polaron velocity approaches its critical value equal to the minimum phonon phase velocity. This increase, apparently, is due to the deformation of the polarization cloud of the polaron, as the form of the Green's function $G(\mathbf{r},\mathbf{r}',t)$ implies (see Eq. (4)). The decrease in the phonon part of the effective polaron mass with the increase in the minimum phonon phase velocity $u$ can easily be explained by the growing ability of polarization to move independently.

Such behavior of the effective polaron mass considered as a function of $u$ and $v$ agrees with the result obtained earlier by Davydov and Énol'skiĭ[7] for the region of low phonon dispersion. They derived an approximate expression for $m^{**}$ of the form

$$m^{**}\approx m^*+\frac{4e^2\gamma^2}{(9\pi)^2a^3\Omega^2}\left(1+\frac{\gamma^2}{5\pi}\left[\frac{181}{21}\frac{v^2}{a^2\Omega^2}-18\frac{u^2}{a^2\Omega^2}\right]\right), \quad (12)$$

with $\gamma=(e^2c/a)(m^*a^3/\hbar^2)$, by replacing the exact formula (4) for the Green's function $G(\mathbf{r},,t)$ with an approximation expression on the assumption that $u,v\ll a\Omega$. Their results in the region $v,u\ll a\Omega$ (for typical ionic crystals this corresponds[7] to $v,u\ll2.5\times10^5$ cm/s, and for the values of the parameters used in constructing Fig. 1 the region is $v,u\ll5\times10^5$ cm/s) are close to those obtained in the present paper: formula (12) demonstrates that the effective polaron mass decreases with increasing $u$ and increases with $v$. However, outside the region where $u\ll a\Omega$ formula (12) may yield negative values of $m^{**}$ for $v$ close to zero.

The point on the vertical axis ($v=0$) denoted by $m^{**}_{\rm Pek}$ stands for the value of the effective mass obtained by Landau

and Pekar's formula[1,2] for the given values of the medium parameters. Clearly, the limit of (10) with $u,v\to 0$, which is

$$\lim_{u\to 0,v<u} m^{**}=\frac{2e^2c}{\pi\Omega^2}\int \frac{k\,dk\,k_z^2dk_z}{k^2+k_z^2}(\psi_k^2)^2, \tag{13}$$

coincides, as it does for any physical problem, with the value in the limit $u=0$, $v\to 0$ obtained by Landau and Pekar.[1,2]


[1] L. D. Landau and S. I. Pekar, Zh. Éksp. Teor. Fiz. **18**, 419 (1948).
[2] S. I. Pekar, *Studies in the Electron Crystal Theory*, Rept. AEC-tr-5575, U.S. Atomic Energy Commission (1963).
[3] É. N. Myasnikov and A. P. Popov, Dokl. Akad. Nauk UkSSR **5**, 73 (1980).
[4] L. D. Landau, Zh. Éksp. Teor. Fiz. **11**, 592 (1941).
[5] L. D. Landau, J. Phys. USSR **5**, 71 (1941); reprinted in I. M. Khalatnikov, *An Introduction to the Theory of Superfluidity*, W. A. Benjamin, New York (1965), pp. 185–204.
[6] A. É. Myasnikova, Phys. Rev. B **52**, 10 457 (1995).
[7] A. S. Davydov and V. Z. Énol'skiĭ, Phys. Status Solidi B **143**, 167 (1987); Zh. Éksp. Teor. Fiz. **94**, No. 2, 177 (1988) [Sov. Phys. JETP **67**, 313 (1988)].
[8] R. P. Feynman, Phys. Rev. **97**, 660 (1955).
[9] J. Appel, ''Polarons,'' in *Solid State Physics: Advances in Research and Applications*, H. Ehrenreich, F. Seitz, and D. Turnbull (eds.), Vol. 21, Academic Press, New York (1968).



Translated by Eugene Yankovsky

# Magnetic and structural correlations in EuMnO and BiMnO crystals in the paramagnetic temperature range

E. I. Golovenchits,*) V. A. Sanin, and A. V. Babinskiĭ

*A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia*
(Submitted 24 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 284–295 (July 1997)

The magnetic and dielectric properties of $EuMn_2O_5$ and $BiMn_2O_5$ crystals are investigated over a wide range of temperatures 4.2–250 K, including the range $T \gg T_N \simeq 40$ K. Significant departures from the Curie–Weiss law are observed in both crystals for the magnetic susceptibility in the paramagnetic range; they are attributed to the presence of correlated domains of magnetic order over a wide range of temperatures. Anomalies in the dielectric properties of the crystals are observed in the same temperature range $T > T_N$ and, as in the case $T < T_N$, are correlated with the magnetic properties. © *1997 American Institute of Physics.* [S1063-7761(97)02407-4]

## 1. INTRODUCTION

Crystals of the series $RMn_2O_5$ ($R^{3+}Mn^{3+}Mn^{4+}O_5^{2-}$, where $R^{3+}$ denotes a rare-earth ion or $Bi^{3+}$) belong to the class of magnetoelectric crystals, for which magnetic and structural phase transitions take place at the same temperature $T \simeq 40$ K (Refs. 1–6). At room temperature these crystals have orthorhombic symmetry described by the space group *Pbam*.[7,8] Neutron-diffraction studies of several $RMn_2O_5$ crystal[9–13] have shown that magnetic ordering with an incommensurate propagation vector $\mathbf{k}=(0.5,\ 0,\ \kappa)$ sets in at a temperature $T \leqslant 40$ K, where $0.16 < \kappa < 0.37$ for crystals with various rare-earth ions. Magnetic studies of $RMn_2O_5$ at temperatures of 16–22 K have often disclosed one other magnetic phase transition. A linear magnetoelectric effect and pyroelectric effect have been observed in a number of $RMn_2O_5$ crystals at temperatures conducive to the existence of long-range magnetic order.[3–6]

Here we investigate the magnetic and dielectric properties of $EuMn_2O_5$ and $BiMn_2O_5$ crystals in the paramagnetic temperature range. It has been shown[1,2] that the magnetic phase transitions in $EuMn_2O_5$ at temperatures of 18 K and 40 K are associated with structural phase transitions. It has also been shown[3] that a linear magnetoelectric effect takes place in $EuMn_2O_5$ at temperatures $T < 40$ K, implying that the structural phase transition at $T \simeq 35$ K is a transition to a noncentrosymmetric space group.

According to neutron-diffraction date,[13] a magnetic incommensurate structure appears in $EuMn_2O_5$ for $T < T_N$ with the propagation vector $\mathbf{k}=(0.5,\ 0,\ 0.33)$, and at a temperature $T < 18$ K a magnetic superstructure emerges with an incommensurate propagation vector close to $\mathbf{k}$. A comparative analysis of data in Refs. 2 and 13 suggests that the onset of a complex magnetic structure and the associated change in crystal structure could be attributable to the presence of numerous, essentially equivalent magnetic exchange couplings of $Mn^{3+}$ and $Mn^{4+}$ ions through oxygen ions.

In the present study we show that correlation between the magnetic and dielectric properties of the investigated crystals is also observed in the paramagnetic temperature range ($T_N < T < 250$ K), probably because of the presence of bounded, correlated magnetic domains at these temperatures.

We carry out a comparative analysis of the magnetic and dielectric properties of $EuMn_2O_5$ and $BiMn_2O_5$ crystals. The $EuMn_2O_5$ crystal has two magnetic subsystems: the subsystem of manganese ions ($Mn^{3+}$ and $Mn^{4+}$) and the subsystem of $Eu^{3+}$ ions. The $Eu^{3+}$ ions are nonmagnetic in the ground state $^7F_0$. However, the first excited state $^7F_1$ of these ions, which is separated from the ground state by an energy $\sim 300$ cm$^{-1}$, is magnetic (i.e., the $Eu^{3+}$ ions form a Van Vleck paramagnetic subsystem). We assume that the magnetic subsystems of $Mn^{3+}$ and $Mn^{4+}$ ions are similar for both of the crystals investigated. Consequently, the magnetic properties of the $BiMn_2O_5$ crystal provide a gauge for assessing the magnetic properties of the subsystem of manganese ions in the $EuMn_2O_5$ crystal as well (since $Bi^{3+}$ ions are nonmagnetic).

In the paramagnetic range correlated anomalies of the magnetic and dielectric susceptibilities are observed both in $BiMn_2O_5$ and in $EuMn_2O_5$. However, the dielectric properties of $EuMn_2O_5$ have certain distinctive features of their own, which are clearly attributable to the influence of thermally excited $Eu^{3+}$ ions.

## 2. MAGNETIC PROPERTIES OF $BiMn_2O_5$ AND $EuMn_2O_5$

The static magnetic susceptibility was investigated on a vibration magnetometer in the temperature range 4.2–250 K along the three principal axes of the crystals. The $BiMn_2O_5$ and $EuMn_2O_5$ single crystals were grown by the technology described in Ref. 1 and had the appearance of brilliant, well-faceted pyramids with dimensions $5\times4\times4$ mm. X-ray phase analysis was performed, confirming the orthorhombic symmetry of the crystals with space group *Pbam*. The crystals were oriented radiographically to within angles of order 1°.

The low-temperature magnetic properties of $EuMn_2O_5$ have been investigated previously,[2] as have those of $BiMn_2O_5$.[14] Here we are concerned primarily with the magnetic properties of these crystals in the paramagnetic temperature range. Certain prominent features of the temperature curves of the reciprocal susceptibility at temperatures close

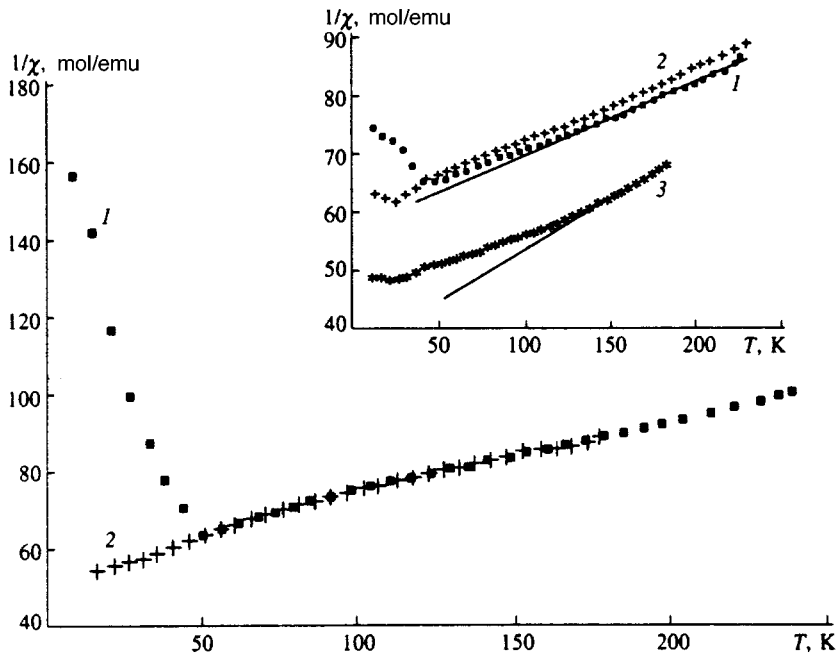



FIG. 1. Temperature dependence of the reciprocal magnetic susceptibility for a $BiMn_2O_5$ crystal. *1)* Orientation of the external static magnetic field along the **a** axis in the crystal (**H**∥**a**); *2)* **H**∥**c**. The inset shows the same dependence for $EuMn_2O_5$: *1)* **H**∥**a**; *2)* **H**∥**b**; *3)* **H**∥**c**.

to 40 K and 20 K (the latter for $EuMn_2O_5$) are visible in Fig. 1. As mentioned in the Introduction, the anomaly near 40 K in all crystals of the $RMn_2O_5$ class is usually caused by the onset of long-range magnetic order, which has been confirmed by neutron-diffraction examinations for several crystals, including $EuMn_2O_5$ (Ref. 13) and $BiMn_2O_5$ (Ref. 14).

The paramagnetic susceptibility of $BiMn_2O_5$ is isotropic, whereas the susceptibility of $EuMn_2O_5$ along the **c** axis ($\chi_c$) differs from the two susceptibilities along the **a** and **b** axes ($\chi_{a,b}$), which are close to each other (see Fig. 1). It is also apparent from Fig. 1 that the values of the paramagnetic susceptibility for $T>40$ K are close for both crystals and, therefore mainly characterize the contribution of the $Mn^{3+}$ and $Mn^{4+}$ ions. Usually the Van Vleck susceptibility for paramagnets containing $Eu^{3+}$ ions is of order $10^{-2}$ emu/mol (see Ref. 15), i.e., are generally of the same order as our measured total paramagnetic susceptibility of the $EuMn_2O_5$ crystal. Moreover, the Van Vleck paramagnetic susceptibility can be anisotropic. According to Ref. 16, rare-earth ions with an integer-valued total momentum $J$ ($J=1$ of the state $^7F_1$), as a rule, have Ising anisotropy, which is precisely the type of anisotropy we should expect of the Van Vleck paramagnetic susceptibility of $Eu^{3+}$ ions. As a result, we can infer from an analysis of Fig. 1 that the contribution of the Van Vleck susceptibility of $Eu^{3+}$ ions imparts anisotropy to the paramagnetic susceptibility of the $EuMn_2O_5$ crystal and is responsible for the deviation of $\chi_c$ from the values of $\chi_{a,b}$.

Furthermore, it is evident from Fig. 1 that the paramagnetic susceptibilities of both crystals deviate significantly from the Curie–Weiss law. The slope of the temperature curve of the reciprocal susceptibility changes for both crystals at a temperature $T\approx 130$ K, a feature that is especially conspicuous for $EuMn_2O_5$ along the **c** axis of the crystal. If we attempt to process the temperature curve of the reciprocal susceptibility according to the paramagnetic law $\chi=C/(T+\Theta_N)$ for antiferromagnets in the two temperature intervals 40 K$<T<$130 K and $T>130$ K, we find that the paramagnetic Néel temperature $\Theta_N$ is much higher than the true value $T_N\simeq 40$ K for $EuMn_2O_5$ along the **c** axis, having values of $-430$ K and $-270$ K, respectively. The effective moments for the two temperature intervals are $\mu^c_{\text{eff}}=8.7\mu_B$ and $\mu^c_{\text{eff}}=7.3\mu_B$, respectively. These moments are approximately equal to the sum of the effective moments of the two ions $Mn^{3+}$ and $Mn^{4+}$ (see, e.g., Ref. 17). We note that similar ratios between $\Theta_N$ and $T_N$ and close values of the moments $\mu_{\text{eff}}$ are obtained for orientations along the **a** and **b** axis in the $EuMn_2O_5$ crystal and for all orientations in the $BiMn_2O_5$ crystal.

### 3. DIELECTRIC PROPERTIES OF $BiMn_2O_5$ AND $EuMn_2O_5$

Wafers of thickness 150–200 $\mu$m oriented perpendicular to the principal (**a**, **b**, **c**) axes were cut from $BiMn_2O_5$ and $EuMn_2O_5$ single crystals. Capacitors were fabricated from these wafers by the spray deposition of gold contacts having an area of $2\times 2$ mm. The real part of the dielectric susceptibility ($\varepsilon'$) and the dielectric loss tangent ($\tan\delta$) were measured in the temperature range 4.2–250 K at frequencies of 100 Hz to 20 kHz.

Dielectric investigations of $EuMn_2O_5$ with an electric field applied along the **a** and **b** axes and along the [110] direction (**e**∥**a**,**b**,[110]) have been carried out previously.[1] Here we report detailed investigations for $BiMn_2O_5$ along all the principal axes of the crystal and for $EuMn_2O_5$ in the case **e**∥**c**. For the paramagnetic range dielectric investigations along the **c** axes are the most informative, because dielectric anomalies are observed only for the orientation **e**∥**c** in the temperature range $T_N<T<250$ K. We note that the kink on the temperature curve of the reciprocal magnetic susceptibility (see Fig. 1) is most conspicuous for orientation along the **c** axis. In the low-temperature range ($T<T_N$), on the other

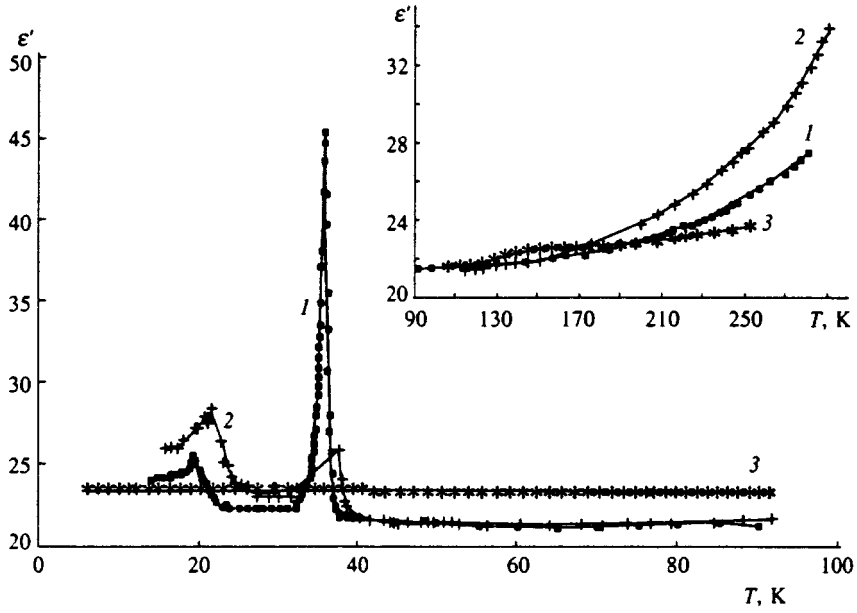



FIG. 2. Temperature dependence of the real part of the dielectric susceptibility for a $EuMn_2O_5$ crystal for three orientations of the alternating electric field relative to the principal axes of the crystal. *1)* **e**$\|$**a**); *2)* **e**$\|$**b**; *3)* **e**$\|$**c**. The frequency is 1 kHz. The inset shows the same dependence for a different temperature interval and a larger $\varepsilon'$ scale.

hand, the anomalies of the magnetic susceptibility are accompanied by prominent features of the dielectric susceptibility and loss tangent for directions in the *ab* plane (see Fig. 2 for $EuMn_2O_5$).

It is evident from the inset to Fig. 2 that a somewhat weak anomaly of the $\varepsilon'(T)$ curve is observed at $T\sim 140-150$ K for $EuMn_2O_5$ with orientation **e**$\|$**c**. The value of $\varepsilon'$ begins to increase rapidly at a higher temperature ($T>250$ K). This increase has been investigated previously[1] for orientations **e**$\|$**a**,**b**,[110], where it was shown to be caused by structurally correlated states associated with thermal population of the first excited level by $Eu^{3+}$ ions.

Figures 3 and 4 show the dielectric anomalies observed in a $EuMn_2O_5$ crystal for orientation **e**$\|$**c**, along with their frequency dispersion. Clearly, the step anomaly $\Delta(\varepsilon')/\varepsilon'$ has a relative maximum $\approx 5\%$, but is readily distinguishable.

The observed low-frequency dispersion of the dielectric properties of $EuMn_2O_5$ is typical of structural glasses[18] and is usually attributed to the coexistence of relaxors with a broad spectrum of relaxation times $\tau_{\min}\ll\tau\ll\tau_{\max}$. Assuming that the structural-glass state occurs for $EuMn_2O_5$, we analyze the dielectric susceptibility in the anomalous dispersion range with a view toward finding characteristic quantities of this state.

Let the temperature $T=T_{\max}$ correspond to the temperature where the loss tangent is a maximum at a fixed frequency $\omega_1$ (Fig. 4). It can be verified that the dependence of the relaxation time $\tau$ on the temperature corresponding to $T=T_{\max}$ satisfies the Arrhenius law, and $\tau=1/\omega_1=\tau_0\exp(E_A/k_BT)$. It is evident from Fig. 5 that $\ln\tau$ as a function of $T$ for $EuMn_2O_5$ has two linear segments with a kink at $T_{\max}=T_{cr}\simeq 130$ K and $\tau=\tau_{cr}\simeq 1.5\cdot 10^{-4}$ s. The values of the characteristic quantities $E_A$ and $\tau_0$ of the acti-

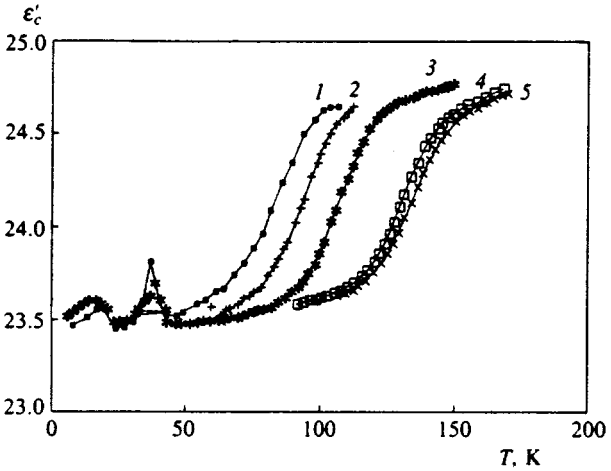



FIG. 3. Temperature dependence of the real part of the dielectric susceptibility of a $EuMn_2O_5$ crystals at several frequencies with the alternating electric field oriented along the **c** axis of the crystal. *1)* 230 Hz; *2)* 500 Hz; *3)* 1 kHz; *4)* 5 kHz; *5)* 20 kHz.

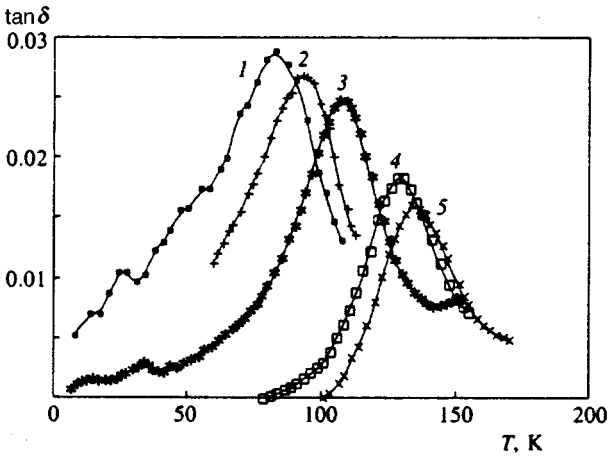



FIG. 4. Temperature dependence of the dielectric loss tangent of a $EuMn_2O_5$ crystal at several frequencies with the alternating electric field oriented along the **c** axis of the crystal. *1)* 200 Hz; *2)* 500 Hz; *3)* 1 kHz; *4)* 10 kHz; *5)* 20 kHz.

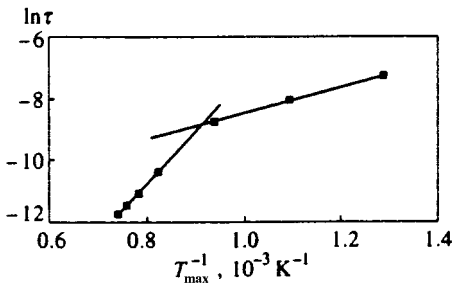



FIG. 5. Graph of $\ln\tau$ vs $T_{\max}^{-1}$ for $EuMn_2O_5$. See the text for further explanations.

vation barrier for the high-temperature (and simultaneously high-frequency) linear segment differ significantly from the values of these quantities for the low-temperature (and low-frequency) segment. In the temperature range $T<130$ K we have $E_A\simeq 0.04$ eV and $\tau_0\simeq 4\times 10^{-6}$ s. In the temperature range $T>130$ K we have $E_A\simeq 0.17$ eV and $\tau_0\simeq 10^{-11}$ s. It is evident that the values of the activation barriers differ considerably. Here a lower temperature corresponds to a smaller value of $E_A$. The values of the characteristic relaxation times $\tau_0$, on the other hand, differ by five orders of magnitude.

Thus, the structural-glass state in $EuMn_2O_5$ is characterized by the critical temperature $T\simeq 130$ K and the critical relaxation time $\tau=\tau_{cr}\simeq 1.5\times 10^{-4}$ s (i.e., the critical scale of the range of correlated states), at which the properties of the glass state suddenly change. In the temperature interval $T_N<T\leqslant 130$ K strong dispersion is encountered in the $EuMn_2O_5$ crystal due to the presence of a broad set of states with a low activation barrier and long characteristic relaxation times, i.e., large domains that are structurally correlated but weakly coupled with the lattice exist in this temperature interval. On the other hand, at temperatures $T\geqslant 130$ K there are smaller-scale (with small characteristic relaxation times $\tau_0$) structurally correlated domains having a high activation barrier. The nature of the structurally correlated domains in the crystal change quite abruptly at $T\simeq 130$ K. We recall that the magnetic properties of the $EuMn_2O_5$ crystal are also observed to change at this temperature. Consequently, correlations of the magnetic and structural properties of $EuMn_2O_5$ crystals are observed experimentally at temperatures $T>T_N$, for which long-range magnetic order is nonexistent.

As mentioned, the contribution of thermally excited $Eu^{3+}$ ions leads to cooperative structural correlations in the $EuMn_2O_5$ crystal only for $T>250$ K (Ref. 1). The observed anomalies of the dielectric susceptibility and the correlations in the magnetic and structural properties for $T_N<T<250$ K are therefore naturally linked to the subsystem of $Mn^{3+}$ and $Mn^{4+}$ ions. It was deemed important in this light to also investigate the dielectric susceptibility in $BiMn_2O_5$ crystals (Figs. 6 and 7). It is evident from Fig. 6 that at temperatures below 120–130 K the susceptibility $\varepsilon_c'$ does not depend on the temperature and is close to the background level of $\varepsilon_c'$ for $EuMn_2O_5$ (Fig. 3). The most noticeable maximum of $\tan\delta$ and the growth of $\varepsilon_c'$ are observed for $BiMn_2O_5$ at $T\sim 130$ K (superposed on the high-temperature growth of $\varepsilon_c'$ and $\tan\delta$), preceding the main high-temperature rise (Figs. 6 and 7). Unfortunately, the high-temperature growth of the dielectric susceptibility in $BiMn_2O_5$ begins at lower temperatures ($T>180$ K) than in $EuMn_2O_5$.

It is evident from Fig. 7 that $BiMn_2O_5$ does not exhibit the strong frequency dispersion observed in the $EuMn_2O_5$ crystal at temperatures $T_N<T<130$ K (Figs. 3 and 4). There is only slight dispersion near 130 K, and it can be identified with ordinary relaxation processes. A comparison of Fig. 4 with Figs. 6 and 7 reveals that the dielectric losses in $BiMn_2O_5$ are much higher than in $EuMn_2O_5$. The relaxation processes in $BiMn_2O_5$ have been studied in Ref. 19, where the estimate $E_A\sim 0.15$ eV is given for the characteristic re-

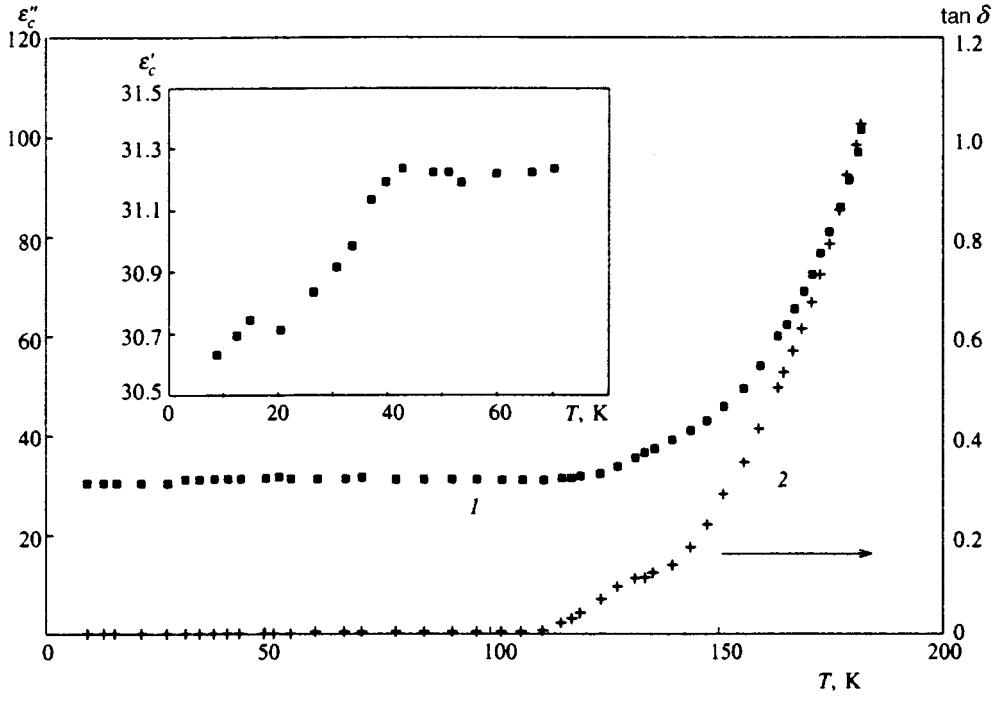



FIG. 6. Temperature dependence of the real part of the dielectric susceptibility (graph *1*) and the dielectric loss tangent (graph *2*) for a $BiMn_2O_5$ crystal at a frequency of 140 Hz with the alternating electric field oriented along the **c** axis of the crystal. The inset shows the dependence of $\varepsilon'$ on $T$ under the same conditions but in the low-temperature range.

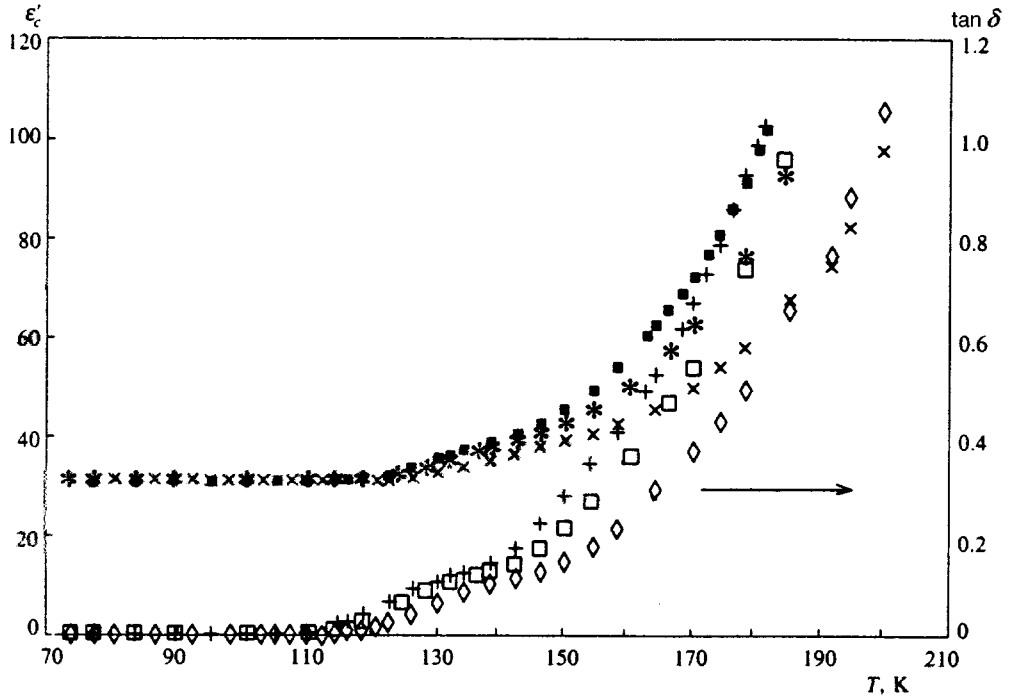



FIG. 7. Temperature dependence of the real part of the dielectric susceptibility and the dielectric loss tangent of $BiMn_2O_5$ at several frequencies with the alternating electric field oriented along the **c** axis of the crystal. ● and + denote $\varepsilon'$ and $\tan\delta$ (respectively) at 140 Hz; * and □ at 230 Hz; × and ◇ at 700 Hz.

laxor activation barrier. We have obtained a similar value in our experiments. We also note that the value of $E_A$ for $BiMn_2O_5$ is close to the value of $E_A$ for the high-temperature (high-frequency) segment in the case of $EuMn_2O_5$.

Consequently, the dielectric and magnetic properties of $BiMn_2O_5$ also have correlated anomalies at a temperature $T\sim 130$ K well above $T_N\simeq 40$ K. These anomalies are not as clearly pronounced as in $EuMn_2O_5$, but they show up experimentally at the same temperature at which the properties of $EuMn_2O_5$ change. We can therefore assume that the indicated anomalies in the magnetic and structural properties of $EuMn_2O_5$ and $BiMn_2O_5$ crystals are mainly associated with manganese ions.

#### 4. MAGNETIC AND STRUCTURAL CORRELATIONS: INTERPRETATION OF THE EXPERIMENTAL DATA

The investigated crystals contain the Jahn-Teller magnetic ion $Mn^{3+}$. We assume that the observed correlation of the magnetic and structural properties in these crystals are in fact attributable to the Jahn–Teller nature of these ions.

It is noteworthy that the anomalies of the structural-glass state in $EuMn_2O_5$ are analogous to those of the orbital-glass state observed earlier in a $Eu_2CuO_4$ crystal containing the Jahn–Teller ion $Cu^{2+}$ (Ref. 20). Anomalous dispersion of the dielectric susceptibility with a low activation barrier and long relaxation times, which suddenly changes in character at $T>250$ K, is also observed in $EuMn_2O_5$ in the temperature interval 120 K$\leqslant T<250$ K. The anomalies of the dielectric susceptibility of $EuMn_2O_5$, as in the crystals analyzed by us, are correlated with distinctive characteristics of the antiferromagnetic state.

It is postulated in Ref. 20 that the anomalous dielectric properties of $EuMn_2O_5$ are attributable to the orbital-glass state induced by two-dimensional antiferromagnetic spin fluctuations with large correlation lengths. Here we attempt to apply the orbital-glass model[20] to the investigated crystals and to find out why an analogous glass state occurs in $EuMn_2O_5$ but is not observed in the related $BiMn_2O_5$ crystal.

The orbital-glass model[20] is based on two principal assumptions: 1) The $Cu^{2+}$ ions have a degenerate orbital ground state (tetragonal doublet $l_z=\pm 1$ with wave functions $d_{xz}$ and $d_{yz}$); 2) orbital–orbital interaction through spin fluctuations exceeds orbital–orbital interaction through phonons. Reference 20 was concerned primarily with the Jahn–Teller effect, in which the role of vibronic interactions is played by interaction through spin fluctuations. Orbital–orbital interaction through spin fluctuations in $Eu_2CuO_4$ is distinguished by its long-range character and by a change of sign at each lattice constant. The long-range character of the interaction is determined by the correlation length of the spin fluctuations and significantly intensifies the interaction.

The orbital ground state of the $M^{3+}$ ion in the investigated crystals has not been established at the present time. The $Mn^{3+}$ ion has a nonstandard nearest-neighbor environment: It exists inside a distorted square pyramid (see Refs. 2 and 8). We assume that the orbital ground state of the $Mn^{3+}$ ions in $EuMn_2O_5$ and $BiMn_2O_5$ crystals is a tetragonal doublet crated by splitting of the principal orbital triplet $\Gamma_5^t(t_{2g})$ in a cubic lattice (see Ref. 21 for the state notation).

It follows from an analysis of the magnetic properties of the $EuMn_2O_5$ and $BiMn_2O_5$ crystals (see Sec. 2) that in the paramagnetic range there is a significant difference between the paramagnetic temperature $\Theta_N$ and the true temperature $T_N$. Such a difference can be attributed to the existence of correlated, bounded, magnetic-order domains for $T>T_N$. The investigated crystals are known to have several ex-

change integrals with different values and signs between neighboring pairs of magnetic ions.[2,9]

Exchange coupling between $Mn^{3+}$ and $Mn^{3+}$ ions, between $Mn^{4+}$ and $Mn^{4+}$ ions, and between $Mn^{3+}$ and $Mn^{4+}$ ions through $O^{2-}$ ions in the lattice of $RMn_2O_5$ crystals has been investigated previously.[2] The subsystem of $Mn^{4+}$ ions forms quasi-one-dimensional chains running along the **c** axis. For the $Mn^{3+}$ ions there are closely spaced pairs in the *ab* planes with the coordinate $c=1/2$. Clearly, $Mn^{3+}$–$Mn^{4+}$ coupling also determines the Néel temperature of the crystal. However, the correlated magnetic-order domains due to $Mn^{3+}$–$Mn^{3+}$ and $Mn^{4+}$–$Mn^{4+}$ exchange coupling can be preserved at temperatures $T>T_N\simeq 40$ K. Furthermore, we note that thermal population of the low magnetic excited level $^7F_1$ by $Eu^{3+}$ ions can also be conducive to the onset of (or the consolidation of existing) correlated magnetic-order domains in $EuMn_2O_5$ crystals as a result of the inclusion of $Eu^{3+}$–$Mn^{3+}$, $Mn^{4+}$ exchange couplings. This possibility does not exist in $BiMn_2O_5$ crystals, so that the correlated magnetic-order domains in $EuMn_2O_5$ can exceed the same domains in $BiMn_2O_5$ crystals.

At temperatures $T>T_N$, therefore, $EuMn_2O_5$ and $BiMn_2O_5$ crystals have native (not attributable to defects or impurities and not frozen in by them) bounded domains of magnetic order, i.e., spin fluctuations with correlation lengths comparable to the size of these domains exist at $T>T_N$. The change of slope of the reciprocal magnetic susceptibility as a function of the temperature at $T\simeq 130$ K for both the $EuMn_2O_5$ and the $BiMn_2O_5$ crystal evinces a change of magnetic state in the subsystem of $Mn^{3+}$ and $Mn^{4+}$ ions at this temperature. It is natural to assume that the correlation length of the magnetic fluctuations will decrease at this temperature.

Consequently, the postulated degenerate orbital state for $Mn^{3+}$ ions and the presence of spin fluctuations at temperatures $T>T_N$ set the stage for orbital–orbital interaction through spin fluctuations in $EuMn_2O_5$ and $BiMn_2O_5$ crystals. This interaction exists together with the usual vibronic Jahn–Teller interaction for $Mn^{3+}$ ions (orbital–orbital interaction through phonons). If interaction through spin fluctuations is dominant, we encounter a magnon analog of the Jahn–Teller effect and the emergence of an orbital-glass state. But if vibronic Jahn–Teller interaction is dominant, we have the customary vibronic Jahn–Teller effect for $Mn^{3+}$ ions. When the Jahn–Teller effect has a magnon character, the tetragonal doublet splits, primarily by virtue of spin fluctuations with low lattice distortion. For the ordinary vibronic Jahn–Teller effect the splitting of the orbital doublet is accompanied mainly by local distortion of the lattice.

As mentioned above (see also Ref. 20), orbital–orbital interaction through spin fluctuations is a long-range phenomenon, and its magnitude depends on the correlation lengths of the spin fluctuations. Consequently, in $EuMn_2O_5$ crystals, in which spin fluctuations can exist with greater correlation lengths than in $BiMn_2O_5$, orbital–orbital interaction through spin fluctuations can exceed the same interaction for $BiMn_2O_5$. In this light we can now understand the observed difference in the dielectric properties in $EuMn_2O_5$ and $BiMn_2O_5$ crystals if we assume that orbital–orbital interaction through spin fluctuations is dominant at temperatures $T_N<T<T_{cr}\sim 130$ K, while for $T>T_{cr}$ for $BiMn_2O_5$ orbital–orbital interaction through phonons is dominant over the entire temperature range. This pattern is consistent with the experimentally determined activation barriers and characteristic lifetimes, and it accounts for the rather abrupt changes of these parameters at $T=T_{cr}$ and $\tau=\tau_{cr}$. For correlated domains of critical size the magnitudes of orbital–orbital interaction through magnons and through phonons are equal at all temperatures. At $T=T_{cr}\simeq 130$ K the most probable size of a correlated domain (corresponding to the maximum density of states) is simultaneously the critical size.

Consequently, the correlations of the magnetic and structural properties of both investigated samples are attributable to the Jahn–Teller nature of the $Mn^{3+}$ ions. The difference in the dielectric properties of $EuMn_2O_5$ and $BiMn_2O_5$ crystals is associated with the contribution of thermally excited $Eu^{3+}$ ions. In $EuMn_2O_5$, owing to the presence of magnetic, thermally excited $Eu^{3+}$ ions, magnetic fluctuations with large correlation lengths occur at temperatures $T>T_N$. This phenomenon is responsible for the anomalous dielectric dispersion observed in $EuMn_2O_5$ in the temperature interval $40\ \text{K}<T\leqslant 130$ K. The upper and lower limits of this interval are dictated by the following considerations. For $T>130$ K, although the density of thermally excited $Eu^{3+}$ ions increases, the exchange coupling in the manganese subsystem, which accounts for the existence of magnetic fluctuations with sufficiently large correlation lengths, is disrupted. At temperatures $T\leqslant 40$ K, on the other hand, in the domain of long-range magnetic order, first of all, magnetic fluctuations are suppressed by the molecular mean field and, second, the symmetry of the crystal is lowered as a result of phase transition,[1,3] and the degeneracy of the orbital ground state cannot be maintained.

Thus, correlation of the magnetic and structural properties exists in the paramagnetic range ($T\gg T_N$) in magnetoelectric $EuMn_2O_5$ and $BiMn_2O_5$ crystals. A particularly interesting physical situation arises in the $EuMn_2O_5$ crystal, in which we assume that the Jahn–Teller effect occurs at sufficiently low temperatures by virtue of orbital–orbital interaction through spin fluctuations, and an orbital-glass state sets in. The detection of such a state in a $Eu_2CuO_4$ crystal has been reported earlier.[20]

This work has received support from the Russian Fund for Fundamental Research (Grants 94-02-05025 and 97-02-18061).

*)E-mail: EGOL@golov.ioffe.rssi.ru

[1] V. A. Sanina, L. M. Salozhnikova, E. I. Golovenchits, and N. V. Morozov, Fiz. Tverd. Tela (Leningrad) **30**, 3015 (1988) [Sov. Phys. Solid State **30**, 1736 (1988)].
[2] E. I. Golovenchits, N. V. Morozov, V. A. Sanina, and L. M. Salozhnikova, Fiz. Tverd. Tela (Leningrad) **34**, 108 (1992) [Sov. Phys. Solid State **34**, 56 (1992)].
[3] T. Doi and K. Kohn, Phase Transit. **38**, 273 (1992).
[4] H. Tsujino and K. Kohn, Solid State Commun. **83**, 639 (1992).
[5] Y. Tanaka, K. Saito, H. Tsujino, and K. Kohn, Ferroelectrics **161**, 125 (1994).
[6] K. Saito and K. Kohn, J. Phys.: Condens. Matter **7**, 2855 (1995).
[7] V. A. Bokov, G. A. Smolenskiĭ, S. A. Kizhaev, and I. E. Myl'nikova, Fiz.

Tverd. Tela (Leningrad) **5**, 3607 (1963) [Sov. Phys. Solid State **5**, 2646 (1963)].
[8] N. Niizeki and M. Wachi, Z. Kristallogr. **127**, 173 (1968).
[9] G. Buisson, Phys. Status Solidi A **16**, 533 (1973).
[10] G. Buisson, Phys. Status Solidi A **17**, 191 (1973).
[11] P. P. Gardner, C. Wilkinson, J. B. Forsyth, and B. M. Wanklyn, J. Phys. C **21**, 5653 (1968).
[12] C. Wilkinson, F. Sinclair, P. Gardner *et al.*, J. Phys. C **14**, 1671 (1981).
[13] I. A. Zobkalo, V. A. Polyakov, O. P. Smirnov *et al.*, Fiz. Tverd. Tela (St. Petersburg) **38**, 1307 (1996) [Phys. Solid State **38**, 725 (1996)].
[14] E. F. Bertraut, G. Buisson, S. Quezel-Ambrunaz, and G. Quezel, Solid State Commun. **5**, 25 (1967).
[15] A. V. Babinskiĭ, E. I. Golovenchits, N. V. Morozov, *et al.*, Fiz. Tverd. Tela (Leningrad) **34**, 60 (1992) [Sov. Phys. Solid State **34**, 31 (1992)].
[16] F. K. Zvezdin, V. M. Matveev, A. A. Mukhin, and A. I. Popov, *Rare-Earth Ions in Magnetically Ordered Crystals* [in Russian], Nauka, Moscow (1985).
[17] J. S. Smart, *Effective Field Theories of Magnetism*, Saunders, London (1966) [Russ. transl., Mir, Moscow (1968)].
[18] S. L. Ginzburg, *Irreversible Phenomena in Spin Glasses* [in Russian], Nauka, Moscow (1989).
[19] I. D. Zhitomirskiĭ, N. E. Skorokhodov, A. A. Bush *et al.*, Fiz. Tverd. Tela (Leningrad) **25**, 953 (1983) [Sov. Phys. Solid State **25**, 550 (1983)].
[20] A. V. Babinskiĭ, S. L. Ginzburg, E. I. Golovenchits, and V. A. Sanina, JETP Lett. **57**, 299 (1993).
[21] A. Abragam and B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions*, Clarendon Press, Oxford (1970) [Russ. Transl., Vols. 1 and 2, Mir, Moscow (1972, 1973)].

Translated by James S. Wood

# Superconductivity fluctuations in a one-dimensional two-band electron–phonon model with strong repulsive interactions

S. I. Matveenko

*Landau Institute for Theoretical Physics, 117940 Moscow, Russia*

A. R. Bishop

*Theoretical Division, Los Alamos National Laboratory, 87545 Los Alamos, New Mexico, USA*

A. V. Balatsky

*Landau Institute for Theoretical Physics, 117940 Moscow, Russia; Theoretical Division, Los Alamos National Laboratory, 87545 Los Alamos, New Mexico, USA*

(Submitted 31 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 296–303 (July 1997)

We study a one-dimensional, two-band model with short-range electron–electron repulsions (onsite $U$ and nearest-neighbor $V$ terms) and electron–phonon coupling. We show that there is a region of $U$, $V$ and band filling in which singlet superconductivity fluctuations are dominant. This region is absent without electron–phonon interactions and includes large values of $U$ and $V$.
 [S1063-7761(97)02507-9]

## 1. INTRODUCTION

The physics of low-dimensional strongly correlated fermion systems with repulsive interactions is a topic of active interest, largely because the origin of high-$T_c$ superconductivity in cuprate oxides and the role of phonons in these correlated systems are not clearly understood. Using a simple one-dimensional (1D) Cu–O chain model,[1] we investigate the effects of both short range electron–electron ($e-e$) repulsive interactions (onsite $U$ and nearest-neighbor Cu–O repulsion $V$) and electron–phonon ($e-p$) coupling in the ground state of the system. We show that superconducting (SC) correlations are absent in the model if we take into account $e-e$ interactions only. The inclusion of $e-p$ interactions leads to the appearance of a $(U,V,\rho)$ region (where $\rho$ is the band-filling) in which superconducting fluctuations are dominant. On the other hand, the ground state of the system in the absence of $e-e$ repulsion is a state with a charge-density wave (CDW) or spin-density wave (SDW) state without a divergent SC response. Thus, the region with dominant SC response results from the combined effect of $e-e$ and $e-p$ interactions for this model.

We use a renormalization-group (RG) two-cutoff approach developed in earlier works.[2,3] With some assumptions on the model parameters, our analysis is valid in the limit of large $U$ and $V$. The possibility of SC fluctuations in quasi-1D systems with strong repulsive $e-e$ interaction and $e-p$ coupling was first raised in work of Zimanyi *et al.*,[3] where results are obtained for a massive Thirring model. The two-band model without $e-p$ interaction was considered in Ref. 1, where numerical results are presented, pointing out the possible existence of SC fluctuations in the strong-coupling limit. This statement is based on the numerical results, which point to the presence of phase separation in the strong-coupling limit, and an intuitive assumption that results obtained for the Luttinger liquid in the weak-coupling limit are valid qualitatively in the strong-coupling limit for the quantum lattice model. Then in some vicinity of the phase separation (where the correlation exponent satisfies $K_\rho \to \infty$) one has a divergent SC response if $K_\rho > 1$ holds (see Eq. (24) below). Here we investigate RG weak-coupling solutions. Therefore we do not consider phase separation. We show that in a region of sufficiently large $e-e$ repulsion, where the RG approach is still applicable, a ground state with a dominant SC response can be achieved due to the interplay of $e-e$ and $e-p$ interactions. We will show that, due to two-band hybridization, the repulsion amplitudes $U$ and $V$ are multiplied by small parameters in the RG equations. This allows us to consider larger $U$ and $V$ values, which exceed, for example, the width of the upper band.

The plan of this article is as follows. In Sec. 2 we define our Hamiltonian and calculate the band structure in the absence of $e-e$ and $e-p$ terms. In Sec. 3 we take into account $e-e$ and $e-p$ terms, and construct the ground-state phase diagram on the basis of our RG analysis. In the Conclusion we discuss our results and their implications.

## 2. THE HAMILTONIAN

We consider a chain consisting of two types of atoms: Cu on odd sites with $d$-orbitals and O on even lattice sites with $p$-orbitals. The Hamiltonian of the system is

$$H = H_0 + H_{ee} + H_{ep}, \tag{1}$$

$$H_0 = -t\sum_{\langle i,j\rangle} c^+_{p,i}c_{d,j} + \text{h.c.} + \sum_i \Delta(c^+_{p,i}c_{p,i} - c^+_{d,i}c_{d,i}), \tag{2}$$

$$H_{ee} = \sum_{\alpha=d,p}\sum_i U_\alpha c^+_{\alpha,i,\uparrow}c_{\alpha,i,\uparrow}c^+_{\alpha,i,\downarrow}c_{\alpha,i,\downarrow} + V\sum_{\langle i,j\rangle} c^+_{d,i}c_{d,i}c^+_{p,j}c_{p,j}, \tag{3}$$

where $t$ is the hopping integral, $\langle i,j\rangle$ are nearest-neighbor sites, $\Delta=(E_p-E_d)/2$, $E_p$ and $E_d$ are site energies, $U_d$ and $U_p$ are Hubbard onsite repulsive energies, and $V$ is the repulsion amplitude between nearest-neighbor sites. Direct antiferromagnetic coupling between Cu sites is omitted. Also

$$H_{ep}=H_{ep,1}+H_{ep,2}, \tag{4}$$

and we consider two models of electron–phonon coupling: the molecular crystal (MC) model with the Hamiltonian $H_{ep,1}$ in which optic phonons couple to the electron site energy, and the Su–Schrieffer–Heeger (SSH) model with the Hamiltonian $H_{ep,2}$ in which the lattice distortions modulate the electron–hopping matrix element $t$. The Hamiltonian $H_{ep,1}$ consists of two parts: $H_{ep,1}=H_{ep,d}+H_{ep,p}$, where each part has the form

$$H_{ep}=\sum_i \frac{P_i^2}{2M}+\frac{1}{2}\kappa q_i^2+\lambda q_i\rho_i=\sum_k \omega_0\left(d_k^+ d_k+\frac{1}{2}\right)$$

$$+\frac{g}{\sqrt{N}}(d_k+d_k^+)\rho_k \tag{5}$$

with $\omega_0=\sqrt{\kappa/M}$, $g=\lambda/\sqrt{2M\omega_0}$, $\rho_k=\Sigma c_{k+q}^+ c_q$. Here $M$ is the ion mass, $\omega_0$ is the optic-phonon frequency, $\kappa$ is the elasticity constant, and $\lambda$ is the $e-p$ coupling constant. All terms in (5) have labels $d$ or $p$, and the sum is over odd or even sites for $H_{ep,d}$ or $H_{ep,p}$, respectively. The Hamiltonian $H_{ep,2}$ takes into account intermolecular phonon modes:

$$H_{ep,2}=\sum_i \frac{P_i^2}{2M}+\frac{1}{2}\kappa(q_{i+1}-q_i)^2-\sum_{\langle i,j\rangle}\delta t_{i,j}c_{d,i}^+ c_{p,j}$$

$$=\sum_k \omega_k\left(f_k^+ f_k+\frac{1}{2}\right)+\frac{1}{\sqrt{N}}\sum_{k,q} g(k,q)$$

$$\times(f_q+f_{-q}^+)c_{d,k,+q}^+ c_{p,k}, \tag{6}$$

where $\delta t_{i,j}=\lambda(q_i-q_j)$, $\omega_q=2\sqrt{\kappa/M}\,\sin(qa/2)$ is the acoustic–phonon frequency, $g(k,q)=4i\lambda\,\sin(qa/2)\cos(ka+qa/2)/\sqrt{2M\omega_q}$, and $a$ is the Cu–O lattice constant.

First we consider the noninteracting Hamiltonian $H_0$. Diagonalization gives

$$H_0=\sqrt{4t^2\,\cos^2 ka+\Delta^2}[c_2^+(k)c_2(k)-c_1(k)^+c_1(k)], \tag{7}$$

where

$$c_d(k)=\cos\,\theta_k c_1(k)+\sin\,\theta_k c_2(k),$$
$$c_p(k)=-\sin\,\theta_k c_1(k)+\cos\,\theta_k c_2(k) \tag{8}$$

with $\tan(2\theta_k)=-2t\cos(ka)/\Delta$, $-\pi/2<2\theta_k<\pi/2$. Now we have a two-band electronic structure and consider the case of an entirely filled lower band. The filling factor of the upper band is $0<\rho<2$ (empty for $\rho=0$ and filled for $\rho=2$). With unit cell $2a$, the quasimomenta $k$ and $k+\pi/a$ are equivalent, and we may assume that the states in the lower band have quasimomenta in the interval $-\pi/2a<k<\pi/2a$ and that in the upper band in the interval $\pi/2a<|k|<\pi/a$; then $k_F a=\pi/2+\pi\rho/4$. The Fermi velocity is

$$v_F=-\frac{2at^2\,\sin(2k_F a)}{\sqrt{4t^2\,\cos^2(k_F a)+\Delta^2}}. \tag{9}$$

### 3. RG TREATMENT

Since we will use an RG approach we take into account only states in the upper band in the vicinity of $E_F$ which are described by the operators $c_2$. Then $H_0$ has the form, in the $x$-representation,

$$H_0=v_F\Psi_{2,+}^+\left(-i\frac{\partial}{\partial x}\right)\Psi_{2,+}+v_F\Psi_{2,-}^+\left(i\frac{\partial}{\partial x}\right)\Psi_{2,-}, \tag{10}$$

where $\Psi_{2,\pm}$ include momenta near $\pm k_F$, respectively. Below we will omit the subscript 2 and also terms in $H$ with $\Psi_1$. (Taking into account the terms with $\Psi_1$ can produce a shift of the chemical potential and some renormalization of the Fermi velocity.) Therefore in the Hamiltonian we can make the replacements

$$\Psi_d\to\sin\,\theta_F\Psi(x),\quad \Psi_p\to\cos\,\theta_F\Psi(x). \tag{11}$$

Note that in the case $t/\Delta\ll 1$ or $\rho\ll 1$ we have

$$\sin\,\theta_F\approx\theta_F\approx\frac{t\,\sin(\pi\rho/4)}{\Delta}. \tag{12}$$

First we consider $e-e$ interaction effects. For the Cu–O case it is appropriate to consider $U_d\gg U_p$. Let us study the case $U_p=0$. The effect of small $U_p$ is easily taken into account and will be discussed below. In terms of a «g-ology» model,[4] the Hamiltonian $H_{ee}$ gives the scattering amplitudes

$$g_1=\frac{Ua}{2}\sin^4\,\theta_F+2Va\,\sin^2\,\theta_F\,\cos(2k_F a)=g_3.$$
$$g_2=\frac{Ua}{2}\sin^4\,\theta_F+2Va\,\sin^2\,\theta_F=g_4, \tag{13}$$

where $g_1$ is the backscattering amplitude, and $g_2$ and $g_4$ are forward scattering amplitudes. The «Umklapp» part $g_3$ exists only for the half-filled case $\rho=1$: for simplicity we will not consider this case. Since we use a RG approach below, we consider $g_i/\pi v_F\ll 1$, i.e., $Ua$, $Va\ll\pi v_F$ or $\sin\theta\ll 1$ for large $U$ and $V$. We have a spin-rotation invariance, i.e., $g_\perp=g_\parallel$. Therefore, when they are not essential, we will omit spin indices. The effect of the $g_4$ term is taken into account separately: it simply produces a shift in the velocity of the spin and charge degrees of freedom: $v_\sigma=v_F(1+g_4)$, $v_\rho=v_F(1-g_4)$.

The familiar RG equations defining the scaling behavior of the system are[4]

$$g_1'=\frac{1}{\pi v_\sigma}g_1^2, \tag{14}$$

$$g_c\equiv g_1-2g_2=\text{const}. \tag{15}$$

For $g_1\geqslant 0$ the excitation spectrum is gapless, $g_1\to g_1^*=0$, while there is a gap if $g_1<0$. The charge excitation spectrum is gapless for $g_c\geqslant 0$ and has a gap $\Delta_\rho$ if $g_c<0$. The ground state has the most divergent singlet (triplet) SC response for $g_c\geqslant 0$ and $g_1<0$ (or $g_c\geqslant 0$ and $g_1\geqslant 0$). In our case

$$g_c = -\frac{Ua}{2}\sin^4\theta_F + 2Va\sin^2\theta_F\cos^2\theta_F \times[\cos(2k_F)-2] < 0. \quad (16)$$

Therefore there is no region in $(U,V)$ with divergent SC fluctuations, in accordance with results[1] for the weak-coupling case. The possible ground states are a CDW or SDW, depending on the sign of $g_1$. (This sign can vary due to the $\cos k_F a$ term.) We see that in order to obtain SC correlations it is necessary to have large positive $g_1^*$ or negative $g_2^*$ terms. As we will see below, this condition can be achieved by taking into account an appropriate $e-p$ interaction.

Second-order perturbation theory in $e-p$ interaction produces a retarded $e-e$ interaction[2] for $\omega$ less than the Debye frequency, $\omega < \omega_D \sim \sqrt{\kappa/M}$. (We consider the case $\omega_D < E_F$.) The effective $e-e$ interaction can be described in «g-ology» terminology: $g_{1,ph} = -2g^2(k_F,2k_F)/\omega_{2k_F}$, $g_{2,ph} = -2g^2(k_F,0)/\omega_0$, $g_{3,ph} = g_{1,ph}$ (half-filled band only). In the case of the MC model (5) we have

$$g_{1,ph} = g_{2,ph} = g_{3,ph} = -\frac{\lambda^2}{4\kappa}, \quad (17)$$

whereas the SSH model (6) gives

$$g_{1,ph} = g_{3,ph} = -4\frac{\lambda^2}{\kappa}(\sin^2\theta_F\cos^2\theta_F). \quad (18)$$

The parameters $\kappa$ and $\lambda$ in (17) and (18) are, of course, different, as well as the other parameters in the Hamiltonians $H_{ep,1}$, $H_{ep,2}$. Note that all terms are negative, and $g_{2'ph}$ is due solely to onsite $e-p$ coupling and does not contain renormalization terms $\sin\theta_F$ and $\cos\theta_F$. In the case $\theta \ll 1$, the onsite $e-p$ interaction is dominant. Now we have two types of $e-e$ interaction with cutoffs $E_F$ and $\omega_D$. Thus we use the RG procedure[2,3] for a two-cutoff model. The one-loop scaling equations (14) and (15) for $g_i$ are unaffected by the presence of retarded interaction. The equations for the $g_{i,ph}$, taking into account the cross terms $g_i g_{j,ph}$, were derived in Ref. 3:

$$g'_{1,ph} = \frac{1}{\pi v_F}\left(\frac{3}{2}g_1 + \frac{1}{2}g_c + g_{1,ph}\right)g_{1,ph}, \quad (19)$$

$$g'_{2,ph} = 0. \quad (20)$$

We shall consider the case $g'_{3,ph} = 0$. The integration in (19) and (20) is taken from $E_F$ to $\omega_0 \sim \omega_D(\omega_0)$, where $\omega_D(\omega_0)$ is the renormalized value of $\omega_D$.[3] As a result, the combined action of different scattering processes is described by

$$g_i^T = g_i^* + g_{i,ph}^*. \quad (21)$$

The properties of the system at energies small compared to $\omega_0$ are derived from a model with single interactions $g_i^T$ and bandwidth $\omega_0$.

Now we examine the solutions of Eqs. (14), (15), (19), and (20). The initial conditions for (14) and (15) are defined by (13). The initial conditions for (19) and (20) are defined by (17) and (18). We write $g_{1,ph}^{(0)} = -\gamma$, $g_{2,ph}^{(0)} = -\tilde{\gamma}$. If $g_1^{(0)} \geqslant 0$ holds (we shall see that this is the situation in the interesting region), from (12) we find that $g_1$ scales toward small positive values $g_1^* \ll g_1^{(0)}$. Note that in the case $\theta \ll 1$ we have $g_{1,ph}^{(0)} \approx g_{2,ph}^{(0)}$, i.e., $\tilde{\gamma} \approx \gamma$. From (13) it follows that $g_1^* - 2g_2^* = g_1^{(0)} - 2g_{(2)}^0$. A positive derivative in (19) implies that $g_{1,ph}$ scales toward large negative values. We consider the opposite case $g'_{1,ph} < 0$. Then, at least initially, $g_{1,ph}$ will scale toward a small negative value. Therefore we demand that

$$\frac{3}{2}g_1^{(0)} + \frac{1}{2}g_c^{(0)} + g_{1,ph}^{(0)} > 0, \quad (22)$$

since $g_{1,ph} < 0$. The inequality (22) can not be valid throughout the scaling process, since $g_1$ scales to small values. Therefore the value $g_{1,ph}^*$ may not be very small. We do not require $|g_{1,ph}^*| \ll \gamma$; for our purposes it is sufficient that $g_{1,ph}^* > \gamma - 2\tilde{\gamma}$ holds, as we show below. The value $g_{2,ph}$ is not scaled, as follows from (20), i.e., $g_{2,ph}^* = -\tilde{\gamma}$. This value does not contain the renormalization coefficient $\sin\theta_F$. As a result of scaling we have the state with $g_i^T = g_i^* + g_{i,ph}^*$. The ground state of the system with the new scaling amplitudes has dominant divergent SC susceptibility if

$$g_c^T = g_1^T - 2g_2^T = g_1^{(0)} - 2g_2^{(0)} + g_{1,ph}^* + 2\tilde{\gamma} > 0. \quad (23)$$

Since we assume that $g_{1,ph}^T \approx g_{1,ph}^* < 0$, we have a state with spin gap $\Delta_\sigma$. Therefore the dominant singularity is the singlet SC response with SC correlation function

$$R(x) \sim x^{-1/K_\rho}, \quad K_\rho = \sqrt{\frac{1+g_c^T/2\pi v_\rho}{1-g_c^T/2\pi v_\rho}} > 1. \quad (24)$$

In this case the CDW response can be divergent with a correlation function $\propto x^{-K_\rho}$. The inequalities (22) and (23) define the region in which the singlet SC correlations are dominant. In terms of $u = (Ua/2)\sin^4\theta$, $v = Va\sin^2\theta\cos^2\theta$ we rewrite (22) and (23) as

$$\gamma + 2v\left(1 + 2\cos\frac{\pi\rho}{2}\right) < u < 2\gamma^* - 2v\left(2 + \cos\frac{\pi\rho}{2}\right), \quad (25)$$

where $2\gamma^* = 2\tilde{\gamma} + g_{1,ph}^*$. It is easy to obtain the solution of (25). This is the region $ABCD$ in Fig. 1 bounded by the lines $u=0$, $v=0$, $u=\gamma-2v$, $u=2\tilde{\gamma}+g_{1,ph}^*-2v$. Recall that in the limit $\theta \ll 1$, Eq. (12), the bare repulsive energies satisfy $U \sim u/\theta^4$, $V \sim v/\theta^2 \gg \gamma$. Thus our model includes the case of strong electron repulsion. For any point $(u,v)$ in the region $ABCD$ the inequality (25) is valid for

$$\rho > \frac{2}{\pi}\cos^{-1}\left(\max\left\{\frac{u-\gamma-2v}{4v}, \frac{2\gamma^*-4v-u}{2v}\right\}\right). \quad (26)$$

In the limit $t/\Delta \ll 1$ we can obtain the phase diagram in terms of the bare values $U$ and $V$. Then the coordinates of the points $A$, $B$, $C$, and $D$ are $A=\{0,(4\gamma^*)(\Delta/t)^4\}$, $B=\{0,2\gamma(\Delta/t)^4\}$, $C=\{(\gamma/2)(\Delta/t)^2,0\}$, and $D=\{(\gamma^*/2)\times(\Delta/t)^2,0\}$. The SC region is deformed to include region II due to the $\sin(\pi\rho/4)$ term. The equation of the curve $EF$ is

$$V = \frac{\Delta^2}{t^2}\frac{((2\gamma^*-\gamma)k+4\gamma+16\gamma^*)^2}{72(k+2)(\gamma+2\gamma^*)}, \quad U = 2kV\frac{\Delta^2}{t^2}. \quad (27)$$

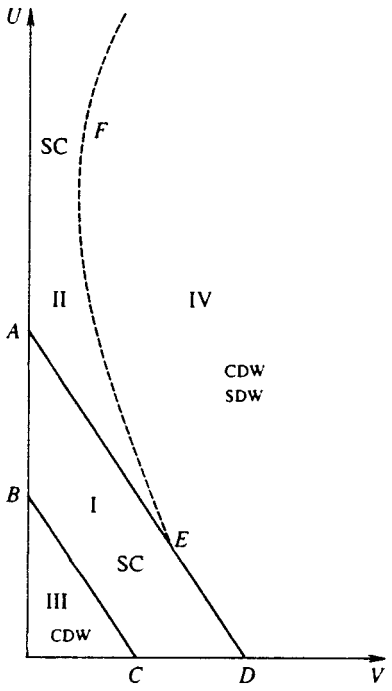



FIG. 1. Phase diagram obtained by two-cutoff RG scaling. The divergent SC response regions are I and II.

In the limit $k\to\infty$ we have $U\propto V^2$, but in this region $\rho\sim 1/U^{1/4}\to 0$. The inequality (25) with $t/\Delta\ll 1$ becomes

$$(U_1+8V_1)y^2-6V_1y-\gamma>0,$$

$$(4V_1-U)y^2-6V_1y+2\gamma^*>0, \tag{28}$$

where $V_1=V(\Delta/t)^2$, $U_1=(U/2)(\Delta/t)^4$, $y=\sin^2(\pi\rho/4)$. In region I the solution of (28) is $\rho_0<\rho<2$, where $\sin^2(\pi\rho_0/4)=y_0$ is the largest root of Eq. (28). In region II we have $\rho_1(U,V)<\rho<\rho_2(U,V)$, where $\rho_1$ and $\rho_2$ can be easily obtained from (28). If $V=0$ holds, the solution is

$$\frac{4}{\pi}\sin^{-1}\left(\frac{\gamma}{U_1}\right)^{1/4}<\rho<\frac{4}{\pi}\sin^{-1}\left(\frac{2\gamma^*}{U_1}\right)^{1/4} \tag{29}$$

for $U_1>\gamma$; if $U=0$ holds, then in the region $\gamma/2<V_1<\gamma^*$ the solution is $\sin^2(\pi\rho/4)>y_0$, where $y_0$ is the largest root of Eq. (28) for $U=0$.

In using the RG approach, we supposed as usual that $g_i/\pi v_F<1$. For small $t/\Delta$ we have the initial value $v_F^{(0)}\sim(t^2/\Delta)\sin(\pi\rho/2)$. Recalling that $g_i\sim V[t\sin(\pi\rho/4)/\Delta]^2$ or $U[t\sin(\pi\rho/4)/\Delta]^4$, $g_{2,ph}=\text{const}$, we can regard our results as reasonable if we are not too close to band edges, where $v_F\to 0$, i.e., $\epsilon_1<\rho<2-\epsilon_2$ and $\rho\neq 1$ ($g_3=0$). It follows from our analysis that in region III we have the large spin and charge gaps, so that there is only a CDW divergent response. In region IV we have $g_c^T<0$, small $g_1^T<0$ and thus divergent CDW and SDW (in the limit $g_1^*\to 0$) responses.

We considered the effect of the lower band only through the renormalization of the bare values $U$ and $V$. Thus, we did not take into account the terms $(V\Theta^2+U)\Psi_1^+\Psi_1\Psi_1^+\Psi_1+(V+U\Theta^2)\Psi_1^+\Psi_1\Psi_2^+\Psi_2$. Therefore our results are valid in the region $U$, $V\leqslant E_{\text{gap}}\sim\Delta$. In order to estimate the effect of the cross term, we can rewrite our two-band model in terms of the two-chain model and use the results of RG investigation.[5] It is easily to see that electron–hole pair interchain hopping is irrelevant (scales to small values), if $U>V$. Thus we can consider the region of large values $U$ and $V$ ($U>V$) in comparison with the upper-band width $\sim t^2/\Delta\ll\Delta$.

In this treatment we have not taken into account the effects of $U_p$ repulsion. This is easily achieved by substituting $u$ into (20) in the form $u=(Ua\sin^4\theta_F+U_pa\cos^4\theta_F)/2$. For small values of $U_p$ the RG approach remains valid, and all results continue to hold in terms of the new $u$ and $v$. For $t/\Delta\ll 1$ we have $\cos\theta_F\sim 1$, so that we cannot consider the large-$U_p$ limit in our approach.

#### 4. CONCLUSIONS

In conclusion, using a two-cutoff RG approach we have studied a two-band, 1D tight-binding model with $e-e$ and $e-p$ interactions. We included onsite $U$ and nearest-neighbor $V$ electron repulsions, as well as intra- and intermolecular $e-p$ coupling. We have shown (in accordance with Ref. 1) that there is no $U$, $V$, $\rho$, $t$, $\Delta$ parameter region with dominant divergent SC response in the absence of $e-p$ interaction. In the lowest-order RG approach we found that such a region does occur if we include $e-p$ coupling with optical intra-molecular modes. Only this form of $e-p$ interaction produces an effective renormalized $g_{2,ph}$ term. We have found that the singlet SC region includes large values of the $U$ and $V$ repulsive interactions if $t\sin(\pi\rho/4)/\Delta\ll 1$. Note that a similar behavior is possible in a one-band model, for which $\Delta=0$. Then, instead of (25), we have

$$\gamma+2V[1-2\cos(\pi\rho)]<U<2\tilde{\gamma}+g_{1,ph}^*-2V[2-\cos(\pi\rho)], \tag{30}$$

where $\gamma=-g_{1,ph}^{(0)}$, $\tilde{\gamma}=-g_{2,ph}^{(0)}$, $0<\rho<2$. The solution of (30) is the same region $ABCD$ in Fig. 1 provided that $2\tilde{\gamma}+g_{1,ph}^*>\gamma$. However, the bare values $U$ and $V$ must be small, of the order of phonon scattering strengths. Note also that we have used a RG approach. Therefore we did not consider the strong-coupling limit ($V$, $U\gg t$, $\Delta$), where a phase separation instability could take place.[1]

The main results of our treatment are the following:

1) Using a 1D two-band model, we have taken into account both $e-p$ coupling and $e-e$ repulsion and have shown that there is a region of parameters with dominant divergent SC response. This effect is absent in the model without $e-p$ coupling and is a result of the interplay of $e-e$ and $e-p$ interactions.

2) We have found that weak $e-p$ interactions and relatively strong $e-e$ repulsions can result in effective electron pairing and a divergent SC response. This is possible in the limit $t/\Delta\ll 1$, where we can take into account large $U$ and $V$ values, since effective $e-e$ interactions are scaled by a factor $t/\Delta$. As a result we find that dominant divergent SC fluctuations are possible in the region $V(t/\Delta)^2\sim U(t/\Delta)^4\sim g_{ph}$, $V<U<\Delta$, as shown in Fig. 1.

3) We have found that dominant SC fluctuation states are possible only in a some interval $\rho_1 < \rho < \rho_2$ of band filling.

4) We have found that only $e-p$ interaction with optical intramolecule phonon modes can result in SC

5) Our conclusions are valid also beyond the limit $t/\Delta \ll 1$ for the two-band model and for the one-band model $(\Delta = 0)$. But in these cases the SC fluctuation ground state is possible in the region of relatively small repulsive constants $(U,\ V \sim g_{ph})$.

This model without $e-p$ coupling was studied in Ref. 1 where some indications of SC fluctuations in the strong-coupling limit were obtained. We have considered a substantially another region.

We have proposed one possible scenario for the origin of dominant SC fluctuations in quasi-one-dimensional systems as a result of the combined effect of repulsive $e-e$ and attractive $e-p$ interactions in a two-band situation. We suggest that features of this picture will survive in analogous two-dimensional models of high-$T_c$ superconductors, in particular in three-band Peierls–Hubbard models.[6] However, the orbital structure of the order parameter in this case ($s$-wave vs $d$-wave) is unclear without detailed calculations.

One of us (S.M) wishes to thank Los Alamos National Laboratory for support and hospitality. The work at Los Alamos was performed under the auspices of the U.S. DoE.

[1] C. M. Varma, S. Schmitt-Rink, and E. Abrahams, Solid State Commun. **62**, 681 (1987); A. Sudbø, S. Schmitt-Rink, and C. M. Varma, Phys. Rev. B **46**, 5548 (1992); A. Sudbø, C. M. Varma, T. Giamarchi, E. B. Stechel, and R. T. Scalettar, Phys. Rev. Lett. **70**, 978 (1993).
[2] L. G. Caron and C. Bourbonnuis, Phys. Rev. B **29**, 4230 (1984); G. S. Grest, E. Abrahams, S.-T. Chui, P. A. Lee, and A. Zawadowski, Phys. Rev. B **14**, 1225 (1976).
[3] G. T. Zimanyi, S. A. Kivelson, and A. Luther, Phys. Rev. Lett. **60**, 2089 (1988).
[4] J. Solyom, Adv. Phys. **28**, 201 (1979).
[5] F. V. Kusmartsev, A. Luther, and A. Nersesyan, JETP Lett. **55**, 724 (1992); V. M. Yakovenko, JETP Lett. **56**, 510 (1992).
[6] Z. Tesanovic, A. R. Bishop, and R. L. Martin, Solid State Commun. **68**, 337 (1988).

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

# Influence of an isolated magnetic impurity on an unconventional superconducting state

Yu. S. Barash and A. G. Grishin

*P. N. Lebedev Physics Institute of Russian Academy of Sciences, 117924 Moscow, Russia*

M. Sigrist

*Theoretische Physik, ETH-Hönggerberg, 8093 Zürich, Switzerland*
(Submitted 6 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 304–312 (July 1997)

The effect of the moment of a magnetic impurity on the order parameter of an unconventional superconductor is examined. The coupling of the magnetic moment to the order parameter induces a locally time-reversal symmetry-breaking state which generates a magnetic field distribution in the vicinity of the impurity. The magnetic field can cause precession of the magnetic moment. The case of a spin polarized muon injected into the superconductor is discussed. © *1997 American Institute of Physics.* [S1063-7761(97)02607-3]

## 1. INTRODUCTION

Some heavy-fermion superconductors possess complex phase diagrams with various superconducting phases.[1] These phase diagrams provide strong evidence for unconventional superconductivity, because the different phases should be distinguished by symmetry. The two examples of such heavy Fermion superconductors are $U_{1-x}Th_xBe_{13}$ and $UPt_3$, which both show two consecutive transitions with high- and low-temperature superconducting states. The minimal requirement for such behavior is that the order parameter have more than one component. Considerable effort on the theoretical and experimental sides has been invested in determining the symmetry of the order parameter in both systems. So far no unambiguous identification of their order parameter symmetry has been achieved. Nevertheless, there is convincing evidence that the low-temperature states in both systems break the time-reversal symmetry $\mathscr{T}$. This fact arises very naturally in most of the phenomenological theories explaining the phase diagram. $\mathscr{T}$-violating states have particular magnetic properties which can be observed in experiment. The zero-field relaxation rate of injected muons shows an increase when the material enters the low-temperature state,[2,3] though the magnitude of this increase may depend sensitively on the sample quality.[4] This rate is a measure of the internal field distribution and its increase indicates additional magnetization occurring in connection with the lower transition.

The additional magnetic fields are due to spontaneous supercurrents flowing in the vicinity of inhomogeneities of the time reversal symmetry-breaking superconducting order parameter, for example, around (nonmagnetic) impurities.[5–9] The net magnetization of an isolated impurity vanishes. There are two length scales involved, the London penetration depth $\lambda$ and the coherence length $\xi$. While screening currents usually affect the magnetic field over a length $\lambda$, the spatial modulation of the currents can lead to an effective canceling of the magnetization on a shorter length comparable with $\xi$ rather than $\lambda$. At the same time, the possible existence of a supercurrent decreasing over a characteristic scale greater than $\xi(T)$ with distance from the impurity, may be associated in this approach with the continuous degeneracy of the superconducting state, which is lifted due to the interaction with an impurity.[5,8]

In this work we consider the problem of spontaneous currents for the time-reversal symmetric phase above the lower transition in the presence of a static magnetic ''impurity.'' This impurity could be an injected muon whose spin can be regarded as static on the relevant time scales of the superconductor. The magnetic moment of the impurity couples to the superconducting order parameter. As we will show, the basic effect is the appearance of a locally $\mathscr{T}$-violating order parameter. By analogy with the case mentioned above, spontaneous supercurrents will be generated. The aim of this paper is to investigate the spatial distribution of these currents and the field pattern. Of particular interest is the magnetic field generated at the impurity site, as it would cause precession of the impurity spin. The essential coupling between impurity and order parameter originates from the combined scattering from the hyperfine and nonmagnetic (and/or spin–orbit and spin–spin) impurity potentials.

## 2. GINZBURG–LANDAU THEORY

Our discussion is based on a generalized Ginzburg–Landau (GL) functional. To be concrete, we use the example of a two-component order parameter as introduced in theories of the phase diagram of $UPt_3$. Thus the order parameter $\boldsymbol{\eta}=(\eta_1,\eta_1)$ belongs either to the irreducible representation $E_1$ or $E_2$ of either parity (singlet or triplet pairing).[10] The general free energy functional is identical for both cases and has the following form

$$
\begin{aligned}
F=\int dV\Big\{ & a_1|\eta_1|^2+a_2|\eta_2|^2+\beta_1(|\eta_1|^2+|\eta_2|^2)^2 \\
& +\beta_2|\eta_1^2+\eta_2^2|^2+K_{123}(|p_x\eta_1|^2+|p_y\eta_2|^2) \\
& +K_1(|p_x\eta_2|^2+|p_y\eta_1|^2)+K_2(p_x^*\eta_1^* p_y\eta_2 \\
& +p_x\eta_1 p_y^*\eta_2^*)+K_3(p_x^*\eta_2^* p_y\eta_1+p_x\eta_2 p_y^*\eta_1^*) \\
& +K_4(|p_z\eta_1|^2+|p_z\eta_2|^2)+\frac{\nabla\times\mathbf{A}^2}{8\pi}\Big\}, \qquad (1)
\end{aligned}
$$

 1063-7761/97/070168-05$10.00  

where $\mathbf{p}=-i\nabla-(2e/c)\mathbf{A}$ ($\mathbf{A}$ is the vector potential), $a_j=\alpha(T-T_{cj})$ and the coefficients are real numbers in the standard notation. We assume $T_{c1}>T_{c2}$ so that in the temperature range $T_{c1}>T>T^*$ only the $\eta_1$-component of the order parameter is finite. $T^*$ denotes the low-temperature transition point below which $\eta_2$ appears,

$$T^*=\frac{T_{c1}+T_{c2}}{2}-\frac{\beta_1}{2\beta_2}(T_{c1}-T_{c2}). \tag{2}$$

In order to have a $\mathscr{T}$-violating low-temperature phase, it is necessary that $\beta_2>0$.

We introduce now the coupling to an impurity located at the origin. The terms in lowest order are

$$F_{\rm imp}=\int dV[\kappa(|\eta_1|^2+|\eta_2|^2)+\nu(|\eta_1|^2-|\eta_2|^2) +\gamma(\eta_1\eta_2^*+\eta_1^*\eta_2)+i\mu(\eta_1\eta_2^*-\eta_1^*\eta_2)]\delta(\mathbf{r}). \tag{3}$$

From the invariance of this expression under the spatial symmetry group for the system of the crystal and the isolated impurity and from its time-reversal symmetry, it follows that $\nu$, $\gamma$, and $\mu$ are not scalar quantities. The coefficients $\nu$ and $\gamma$ differ from zero for a hexagonal crystal only for impurity states breaking the symmetry with respect to rotations around the hexagonal axis through the angle $\pi/3$. The last term describes the linear (or odd order) coupling of the magnetic moment to the order parameter. Note that $\eta_1\eta_2^*-\eta_1^*\eta_2$ is finite only if the order parameter $\boldsymbol{\eta}$ breaks time-reversal symmetry, i.e., the relative phase between the two components is not 0 or $\pi$. Hence, the coefficient $\mu$ differs from zero only for the time-reversal-breaking state of the impurity.

We consider now the effect of the impurity on the order parameter in the high-temperature phase where $\boldsymbol{\eta}=\boldsymbol{\eta}_0=\eta_0(1,0)$, choosing $\eta_0$ real with

$$\eta_0^2=\frac{-a_1(T)}{2(\beta_1+\beta_2)}. \tag{4}$$

For simplicity we assume that the coupling is weak so that the distortion of the order parameter is small. We consider $\boldsymbol{\eta}=\boldsymbol{\eta}_0+\boldsymbol{\psi}$, where $\boldsymbol{\psi}=\boldsymbol{\psi}'+i\boldsymbol{\psi}''$ is small compared with $\eta_0$. Since for the homogeneous phase the vector potential vanishes, we can also assume $\mathbf{A}$ to be small. Therefore we analyze the GL equations linearized in $\boldsymbol{\psi}$ and $\mathbf{A}$. This leads to seven coupled equations, obtained by varying $F+F_{\rm imp}$ with respect to the order parameter,

$$\begin{aligned}
&2a_1\psi_1'+K_{123}\partial_{xx}^2\psi_1'+K_1\partial_{yy}^2\psi_1'+K_4\partial_{zz}^2\psi_1'+K_{23}\partial_{xy}^2\psi_2' =(\kappa+\nu)\eta_0\delta(\mathbf{r}),\\
&\Delta a\psi_2'-K_{123}\partial_{yy}^2\psi_2'-K_1\partial_{xx}^2\psi_2'-K_4\partial_{zz}^2\psi_2'-K_{23}\partial_{xy}^2\psi_1' =-\gamma\eta_0\delta(\mathbf{r});\\
&K_{123}\partial_x\left(\partial_x\psi_1''-\frac{2e\eta_0}{c}A_x\right)+K_1\partial_y\left(\partial_y\psi_1''-\frac{2e\eta_0}{c}A_y\right) +K_4\partial_z\left(\partial_z\psi_1''-\frac{2e\eta_0}{c}A_z\right)+K_{23}\partial_{xy}^2\psi_2''=0,
\end{aligned} \tag{5}$$

$$\begin{aligned}
&a_*\psi_2''-K_{123}\partial_{yy}^2\psi_2''-K_1\partial_{xx}^2\psi_2''-K_4\partial_{zz}^2\psi_2''-K_2\partial_y \times\left(\partial_x\psi_1''-\frac{2e\eta_0}{c}A_x\right)-K_3\partial_x\left(\partial_y\psi_1''-\frac{2e\eta_0}{c}A_y\right) =-\mu\eta_0\delta(\mathbf{r});
\end{aligned} \tag{6}$$

and with respect to the vector potential

$$\begin{aligned}
\partial_x\ \mathrm{div}\ \mathbf{A}-\Delta A_x&=\frac{16\pi e\eta_0}{c}K_{123}\left(\partial_x\psi_1''-\frac{2e\eta_0}{c}A_x\right)+\frac{16\pi e\eta_0}{c}K_2\partial_y\psi_2'',\\
\partial_y\ \mathrm{div}\ \mathbf{A}-\Delta A_y&=\frac{16\pi e\eta_0}{c}K_1\left(\partial_y\psi_1''-\frac{2e\eta_0}{c}A_y\right)+\frac{16\pi e\eta_0}{c}K_3\partial_x\psi_2'',\\
\partial_z\ \mathrm{div}\ \mathbf{A}-\Delta A_z&=\frac{16\pi e\eta_0}{c}K_4\left(\partial_z\psi_1''-\frac{2e\eta_0}{c}A_z\right).
\end{aligned} \tag{7}$$

Here the following abbreviations were used: $\Delta a=a_2-a_1$, $a_*=2\beta\alpha(T-T_*)/(1+\beta)$, $\beta=\beta_2/\beta_1$. Note that the first two equations do not couple to the remaining five. The $\kappa$, $\nu$ and $\gamma$ terms in $F_{\rm imp}$ act only on the real part of the order parameter, inducing a finite real $\eta_2$-component in the vicinity of the impurity. We will not discuss these two equations further here, since they cause distortion of the order parameter without interesting effects involving the magnetic properties.

Clearly the imaginary part of the order parameter and the vector potential couple in Eqs. (6) and (7). The right-hand sides of the last three equations correspond essentially to the components of the supercurrents $4\pi\mathbf{j}/c$. It is only the last term of $F_{\rm imp}$ which enters into these equations. Obviously, the presence of a magnetic moment drives the imaginary order parameter components. This leads immediately to finite supercurrents and a magnetic field distribution.

A simplification occurs if we take the gauge freedom of the order parameter phase into account. In first order the quantity $\psi_1''/\eta_0$ is in fact a common phase of the order parameter $((\eta_0+\psi_1'+i\psi_1'',\psi_2'+i\psi_2'')\simeq(\eta_0+\psi_1',\psi_2'+i\psi_2'')\exp(i\psi_1''/\eta_0))$, whose value is directly associated with a gauge for the vector potential $\mathbf{A}$. Therefore we can choose

$$\psi_1''=0 \tag{8}$$

as a gauge condition. Furthermore, one can see easily that the first equation in Eqs. (6) is equivalent to the condition $\mathrm{div}\,\mathbf{j}=0$, and the same condition obviously follows from the Maxwell equations. Therefore, this equation may be omitted and we reduce the problem with the aid of Eq. (8) to the following four equations for the unknown quantities $\mathbf{A}$, $\psi_2''$:

$$\begin{aligned}
&\partial_{xy}^2A_y+\partial_{xz}^2A_z+(\lambda_{123}^{-2}-\partial_{yy}^2-\partial_{zz}^2)A_x=\frac{16\pi}{c}e\eta_0K_2\partial_y\psi_2'',\\
&\partial_{xy}^2A_x+\partial_{yz}^2A_z+(\lambda_1^{-2}-\partial_{xx}^2-\partial_{zz}^2)A_y=\frac{16\pi}{c}e\eta_0K_3\partial_x\psi_2'',
\end{aligned}$$

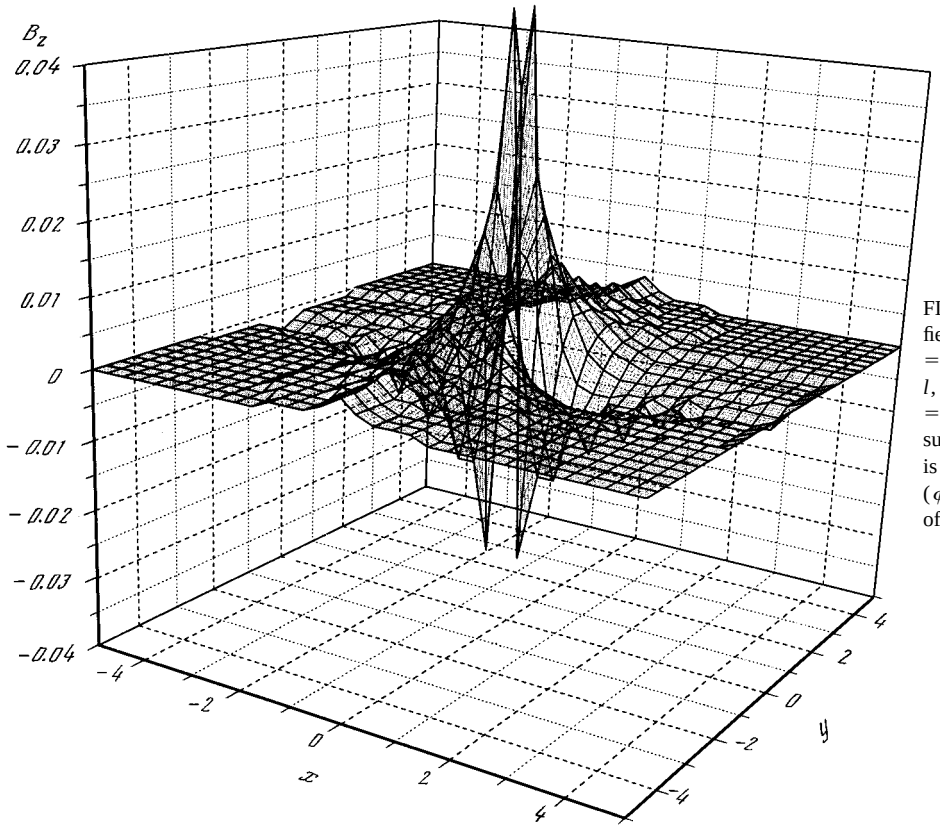



FIG. 1. The spatial distribution of the magnetic field at $z=0$ for $K_3=1.5K_2\ll K_1=K_4$, $l=10$, $l_2=0.258$, $l_3=0.316$, $l_4=1$. The parameters $l$, $l_2$, $l_3$, $l_4$ are defined as follows: $\lambda_1=l_2\lambda_2=l_3\lambda_3=l_4\lambda_4=l\xi_1(T)$. All distances are measured in units of $\xi_1(T)$, while the magnetic field is given in units of $(K_2)/(K_1)$ $(\mu/\pi^3 a^* \xi_1^3(T))$ $(\phi_0/2\pi\xi_1^2(T))$. The value of $B_z(0)$ is about 77% of its maximum value

$$\partial_{xz}^2 A_x+\partial_{yz}^2 A_y+(\lambda_4^{-2}-\partial_{xx}^2-\partial_{yy}^2)A_z=0,$$

$$a_*\psi_2''-K_1\partial_{xx}^2\psi_2''-K_{123}\partial_{yy}^2\psi_2''-K_4\partial_{zz}^2\psi_2''$$

$$=-\mu\eta_0\delta(\mathbf{r})-\frac{2e\eta_0}{c}(K_2\partial_y A_x+K_3\partial_x A_y). \tag{9}$$

We have introduced here the notation $\lambda_i^{-2}=(32\pi e^2\eta_0^2/c^2)K_i$. These equations can be easily solved in momentum space. We use the Fourier transformation

$$\tilde{\mathbf{A}}(\mathbf{k})=\frac{1}{\sqrt{V}}\int dV\mathbf{A}(\mathbf{r})e^{i\mathbf{kr}},$$

$$\tilde{\psi}_2''(\mathbf{k})=\frac{1}{\sqrt{V}}\int dV\psi_2''(\mathbf{r})e^{i\mathbf{kr}}, \tag{10}$$

which leads to

$$(\lambda_{123}^{-2}+k_y^2+k_z^2)\tilde{A}_x-k_xk_y\tilde{A}_y-k_xk_z\tilde{A}_z-\frac{ic}{2e\eta_0}\lambda_2^{-2}k_y\tilde{\psi}_2''=0,$$

$$-k_xk_y\tilde{A}_x+(\lambda_1^{-2}+k_x^2+k_z^2)\tilde{A}_y-k_yk_z\tilde{A}_z-\frac{ic}{2e\eta_0}\lambda_3^{-2}k_x\tilde{\psi}_2''=0,$$

$$-k_xk_z\tilde{A}_x-k_yk_z\tilde{A}_y+(\lambda_4^{-2}+k_x^2+k_y^2)\tilde{A}_z=0, \tag{11}$$

$$\frac{2ie\eta_0}{ca_*}(K_2k_y\tilde{A}_x+k_3k_x\tilde{A}_y)$$

$$+(1+\xi_1^2k_x^2+\xi_{123}^2k_y^2+\xi_4^2k_z^2)\tilde{w}_2''=-\frac{\mu\eta_0}{a_*}.$$

The solution of this equation is straightforward but gives a rather complicated result (see below Eq. (16) and figure). A good picture of the result can be obtained by solving the last equation for $\psi_2''$ in the absence of the vector potential and then inserting the latter into the other three equations. This approximation may be justified, for example, under the conditions $K_2\sim K_3\ll K_1\sim K_4$. Then the order parameter has the form

$$\tilde{\psi}_2''(\mathbf{k})=-\frac{\mu\eta_0}{a_*(T)}\frac{1}{1+\xi_1^2k_x^2+\xi_{123}^2k_y^2+\xi_4^2k_z^2}, \tag{12}$$

which corresponds to a shape like an anisotropic Yukawa potential in real space. The induced imaginary component $\psi''$ of the order parameter leads to a local $\mathscr{T}$-violation. The length scales over which $\psi_2''$ decays are the anisotropic temperature-dependent coherence lengths $\xi_i^2=K_i/a_*$, which diverge as $T$ approaches $T^*$. Obviously, $\psi''$ is infinite at $r=0$ in real space, because the use of a delta function in Eq. (3) eliminates the lower cutoff-length scale. Within the Ginzburg–Landau theory the natural cutoff length is $\xi_0$. Therefore for the qualitative consideration of quantities at $r=0$ we need a cutoff which is at least of order the zero-temperature coherence length of $\eta$. The order parameter modulation yields supercurrents in the form

$$\tilde{j}_x=4e\eta_0 ik_yK_2\tilde{\psi}_2'', \quad \tilde{j}_y=4e\eta_0 ik_xK_3\tilde{\psi}_2'', \tag{13}$$

and $\tilde{j}_z=0$ if we neglect the screening currents for the moment. We use now these currents as a source and calculate the induced vector potential

$$\tilde{A}_x=\frac{4\pi}{Dc}\,[\tilde{j}_x(R_yR_z-k_y^2k_z^2)+\tilde{j}_yk_xk_y(R_z+k_z^2)],$$

$$\tilde{A}_y=\frac{4\pi}{Dc}\,[\tilde{j}_y(R_xR_z-k_x^2k_z^2)+\tilde{j}_xk_xk_y(R_z+k_z^2)], \tag{14}$$

$$\tilde{A}_z=\frac{4\pi}{Dc}\,[\tilde{j}_xk_xk_z(R_y+k_y^2)+\tilde{j}_yk_yk_z(R_x+k_x^2)],$$

where

$$D=R_xR_yR_z-2k_x^2k_y^2k_z^2-R_xk_y^2k_z^2-R_yk_x^2k_z^2-R_zk_x^2k_y^2,$$
$$R_x=\lambda_{123}^{-2}+k_y^2+k_z^2,\quad R_y=\lambda_1^{-2}+k_x^2+k_z^2, \tag{15}$$
$$R_z=\lambda_4^{-2}+k_x^2+k_y^2.$$

We consider now the magnetic field distribution around the impurity site. Using $\mathbf{B}=\nabla\times\mathbf{A}$ we obtain for the Fourier-transformed magnetic field, $\tilde{\mathbf{B}}=i\mathbf{k}\times\tilde{\mathbf{A}}$,

$$\tilde{B}_x=\frac{16\pi e\,\eta_0\tilde{\psi}_2''k_xk_z}{Dc}\,[K_2k_y^2(\lambda_4^{-2}-\lambda_1^{-2})$$
$$+K_3(\lambda_{123}^{-2}k_x^2+\lambda_4^{-2}R_x)],$$
$$\tilde{B}_y=\frac{16\pi e\,\eta_0\tilde{\psi}_2''k_yk_z}{Dc}\,[-K_2(\lambda_1^{-2}k_y^2+\lambda_4^{-2}R_y)$$
$$+K_3k_x^2(\lambda_{123}^{-2}-\lambda_4^{-2})], \tag{16}$$
$$\tilde{B}_z=\frac{16\pi e\,\eta_0\tilde{\psi}_2''}{Dc}\,[K_2k_y^2(\lambda_4^{-2}k_z^2+R_z\lambda_1^{-2})$$
$$-K_3k_x^2(\lambda_4^{-2}k_z^2+R_z\lambda_{123}^{-2})].$$

The magnetic field distribution has a rather complicated structure, as we show for the $B_z$ component in figure. We do not analyze this structure further, but concentrate on the magnetic field at the site of the impurity. For this purpose we have to perform the Fourier transform from momentum space to real space. At $r=0$ this corresponds simply to the $\mathbf{k}$-integral of $\tilde{\mathbf{B}}(\mathbf{k})$. We see immediately that there are no $x$- and $y$-components, because the angular dependence in $k$-space leads to an exact cancellation. The $z$-component, however, is finite, if we take the lower cutoff length into account properly.

As a consequence the magnetic field would lead to precession of the magnetic moment around the $z$-axis. This precession does not change the $z$-component of the moment so that the coupling term with the superconducting order parameter in Eq. (3) is not changed at all. Therefore the local superconducting state and its field distribution is essentially static despite the precession of the impurity moment. Regarding the muon as an impurity, one could measure the precession in the standard way through the muon decay into positrons. In a very clean material all muons are usually trapped in crystallographically equivalent (very symmetric) points and, consequently, have the same environment. If completely spin-polarized muons are injected, all of them should generate the same local magnetic field distribution and hence have the same precession frequency $\omega$. The frequency $\omega$ depends, however, on the angle $\theta$ between the expectation value of the muon spin and the $z$-axis of the crystal. Because $B_z$ is proportional to $\mu$ the frequency is

$$\omega\propto B_z(r=0)\propto\mu\propto\cos\,\theta. \tag{17}$$

Of course, the precession of the muon spin can only be seen if $0<\theta<\pi/2$.

On the other hand, in a dirty sample the trapping positions of the muons may be scattered so that the magnetic field generated at the muons is spread over many values. Then we would not observe a pure precession, but rather a depolarization for the $x$-$y$-component of the spin. In both cases the effect should become stronger as we approach the transition at $T^*$.

### 3. MICROSCOPIC DERIVATION OF THE IMPURITY TERMS

In the following we discuss briefly the microscopic calculations of the coefficient $\mu$ as well as $\kappa$, $\nu$ and $\gamma$, assuming for simplicity hole–particle symmetry for the energy spectrum. In quasiclassical theory the basic equations for the propagators in the presence of the isolated impurity may be written as follows[6]

$$[i\varepsilon_n\hat{\tau}_3-\hat{\sigma}(\mathbf{k}_F,\mathbf{R}),\hat{g}(\mathbf{k}_F,\mathbf{R};\varepsilon_n)]+i\mathbf{v}_F\nabla_R\hat{g}(\mathbf{k}_F,\mathbf{R};\varepsilon_n)$$
$$=[\hat{t}(\mathbf{k}_F,\mathbf{k}_F;\varepsilon_n),\hat{g}_{\rm int}(\mathbf{k}_F,\mathbf{R}=\mathbf{R}_{\rm imp};\varepsilon_n)]\,\delta(\mathbf{R}-\mathbf{R}_{\rm imp}). \tag{18}$$

Here $\varepsilon_n=(2n+1)\pi T$ is the Matsubara frequency, $\mathbf{k}_F$ is the momentum direction on the Fermi surface, $\mathbf{v}_F(\mathbf{k}_F)$ is the Fermi velocity, and $\hat{\tau}_3$ is the third Pauli matrix in Nambu space.

The normalization condition for the matrix propagator is

$$\hat{g}^2(\mathbf{k}_F,\mathbf{R};\varepsilon_n)=-\pi^2\hat{1}. \tag{19}$$

Equations (18), (19) must be supplemented by the equation for the quasiparticle scattering $t$-matrix of the impurity

$$\hat{t}(\mathbf{k}_F,\mathbf{k}_F';\varepsilon_n)=\hat{\nu}(\mathbf{k}_F,\mathbf{k}_F')+N(0)$$
$$\times\int\frac{d^2\mathbf{k}''}{4\pi}\,\hat{\nu}(\mathbf{k}_F,\mathbf{k}_F'')\hat{g}_{\rm int}(\mathbf{k}_F'',\mathbf{R}$$
$$=\mathbf{R}_{\rm imp};\varepsilon_n)\hat{t}(\mathbf{k}_F'',\mathbf{k}_F';\varepsilon_n). \tag{20}$$

Here $\hat{\nu}(\mathbf{k}_F,\mathbf{k}_F')$ is the matrix of the impurity potential. The auxiliary quantity $\hat{g}_{\rm int}(\mathbf{k}_F,\mathbf{R};\varepsilon_n)$ obeys the normalization condition and Eq. (18) without the $t$-matrix impurity term on the right-hand side.

The impurity potential matrix $\hat{\nu}(\mathbf{k}_F,\mathbf{k}_F')$ may be represented in the form

$$\hat{\nu}(\mathbf{k}_F,\mathbf{k}_F')=w_{kk'}\hat{1}+iv_{kk'}\hat{\tau}_3+u_{kk'}\mathbf{M}\cdot\hat{\mathbf{S}}+i\mathbf{m}_{kk'}\cdot\hat{\mathbf{S}}\hat{\tau}_3. \tag{21}$$

Here terms $w_{kk'}$, $v_{kk'}$, $u_{kk'}$ and $\mathbf{m}_{kk'}$ describe the conventional nonmagnetic potential, the hyperfine interaction, the magnetic spin-spin and spin-orbit coupling respectively. The form of the spin operator $\hat{\mathbf{S}}$ is defined as in Ref. 11.

The Ginzburg–Landau equations are obtained by expanding the self-consistency equation

$$\Delta(\mathbf{k}_F,\mathbf{q})=2TN(0)\sum_{\varepsilon_n}\int\frac{d^2k'}{4\pi}V(\mathbf{k}_F,\mathbf{k}'_F)f(\mathbf{k}'_F,\mathbf{q};\varepsilon_n) \tag{22}$$

in powers of the order parameter and its spatial derivatives. For one-dimensional representations the contributions from an isolated nonmagnetic impurity to the free energy functional were considered in Refs. 12 and 13 for estimation of the vortex pinning potential. Since we are interested in the terms in Eq. (3), we can put $\mathbf{q}=0$ in Eq. (22), omitting gradient terms. This equation is written in the form valid for singlet pairing, $\Delta(\mathbf{k}_F)=i\sigma_y\psi(\mathbf{k}_F)$, and for the particular kind of triplet pairing $(\Delta(\mathbf{k}_F)=i(\mathbf{d}(\mathbf{k}_F)\sigma)\sigma_y$ with $\mathbf{d}\|\mathbf{z}$, where $\mathbf{z}$ is the hexagonal crystalline axis), if one makes use of the notations $\Delta(\mathbf{p})=\psi(\mathbf{p})$ for the former and $\Delta(\mathbf{p})=d_z(\mathbf{p})$ for the latter cases. Mostly, these types of pairing are discussed for the analysis of experimental data of $UPt_3$.[10] We consider the pairing potential of the form

$$V(\mathbf{p},\mathbf{p}')=-\frac{g}{2}[\varphi_1(\mathbf{p})\varphi_1(\mathbf{p}')+\varphi_2(\mathbf{p})\varphi_2(\mathbf{p}')]$$

and assume the basis functions to be real and normalized according to $\int d\Omega\varphi^2_{1,2}(\hat{k})=4\pi$.

The solution of Eq. (20) in the second Born approximation and its substitution into Eq. (18) are straightforward, since one can use the bulk expression for the quantity $\hat{g}_{\rm int}$ in the case $\sigma/\xi_0^2\ll 1$, where $\sigma$ is the quasiparticle cross-section for the impurity potential.[5,6] From the solution of the Eilenberger equations in this approximation we obtain the impurity contribution to the anomalous propagator

$$\begin{aligned}f_{\rm imp}=\left(\frac{\pi}{\varepsilon_n}\right)^2N(0)\int\frac{d\Omega'}{4\pi}[&-(v^2_{kk'}-m^2_{kk'}+M^2u^2_{kk'}\\&-w^2_{kk'})\Delta(\mathbf{k}'_F)-(v^2_{kk'}+m^2_{kk'}+M^2u^2_{kk'}\\&+w^2_{kk'})\Delta(\mathbf{k}_F)+2i(v_{kk'}w_{kk'}\\&-u_{kk'}\mathbf{m}_{kk'}\cdot\mathbf{M})\Delta(\mathbf{k}'_F)].\end{aligned} \tag{23}$$

Only the last term of this expression for the $f$-function, substituted into the self-consistency equation (22), yields a finite value of $\mu$,

$$\begin{aligned}\mu=\frac{\pi^2}{16}M_z\alpha v_F\lambda N^2(0)\int d\Omega\int d\Omega'(&\Phi_{kk'}-bu_{kk'})\\&\times w_{kk'}(\hat{k}_x\hat{k}'_y-\hat{k}_y\hat{k}'_x)\varphi_1(\mathbf{k}_F)\varphi_2(\mathbf{k}'_F).\end{aligned} \tag{24}$$

Here the coupling constant $\lambda=gN(0)$ is expressed in terms of the critical temperature in the conventional way ($T_c\propto\exp(-1/\lambda)$) and the matrix elements for the hyperfine and spin-orbit interactions are represented in the form

$$v_{kk'}=(\mathbf{k}_F\times\mathbf{k}'_F\cdot\mathbf{M})\Phi_{kk'},\quad \mathbf{m}_{kk'}=bw_{kk'}\mathbf{k}_F\times\mathbf{k}'_F.$$

Note, that for the point-like impurity potential, when $w_{kk'}$, $\Phi_{kk'}$ and $u_{kk'}$ do not depend upon the momenta $\mathbf{k}$, $\mathbf{k}'$, the coefficient $\mu$ vanishes for singlet pairing. This is not the case for triplet superconductors due to the different parity properties of the basis functions $\varphi_{1,2}(\mathbf{k}_F)$ for singlet and triplet superconductors. This result may be justified beyond the Born approximation as well. Since the expression Eq. (24) is proportional to $M_z$, the coefficient $\mu$ changes its sign under the time-reversal operation, which ensures the time-reversal symmetry of the whole expression $i\mu(\eta_1\eta_2^*-\eta_1^*\eta_2)$.

## 4. CONCLUSION

We have demonstrated phenomenologically that a magnetic impurity can generate a locally $\mathscr{T}$-violating superconducting phase. This leads to a distribution of supercurrents and magnetic fields which acts on the magnetic moment. For the two representations $E_1$ and $E_2$ considered here, only the $z$-component of the magnetic moment couples to the superconducting order parameter, and the resulting magnetic field has only a finite $z$-component at the impurity site. We have shown that this fact yields the precession of the magnetic moment without changing the locally $\mathscr{T}$-violating order parameter configuration. Thus, for injected muons this may lead to precession of the spin. However, it is difficult to estimate whether the generated magnetic field would be sufficiently large to really give an observable precession. Our discussion may also apply to other systems besides the $UPt_3$ we had in mind here. This is important for the enhancement of effects considered to be in the vicinity of a bulk transition to a superconducting state with broken time reversal symmetry.

We are grateful to T. M. Rice and H. Monien for helpful discussions. M. S. would like to thank the Swiss Nationalfonds for financial support. Yu. S. B. and A. G. G. were supported by Grant No. 96-02-16249 of the RFBR. Two of the authors (Y. B. and M. S.) are grateful to the I. S. I. for hospitality at the Villa Gualino in Torino, where this paper was finished.

[1] M. Sigrist and K. Ueda, Rev. Mod. Phys. **63**, 239 (1991).
[2] R. H. Heffner, J. L. Smith, J. O. Willis *et al.*, Phys. Rev. Lett. **65**, 2816 (1990).
[3] G. M. Luke, A. Keren, L. P. Le *et al.*, Phys. Rev. Lett. **71**, 1466 (1993).
[4] P. D. de Reotier, A. Huxley, A. Yaouanc *et al.*, Phys. Lett. A **205**, 239 (1995).
[5] D. Rainer and M. Vuorio, J. Phys. C: Sol. St. Phys. **10**, 3093 (1977).
[6] E. V. Thuneberg, J. Kurkijärvi, and D. Rainer, J. Phys. C: Sol. St. Phys. **14**, 5615 (1981).
[7] C. H. Choi and P. Muzikar, Phys. Rev. B **39**, 9664 (1989).
[8] C. H. Choi and P. Muzikar, Phys. Rev. B **41**, 1812 (1990).
[9] V. P. Mineev, JETP Lett. **49**, 719 (1989).
[10] J. A. Sauls, Adv. Phys. **43**, 113 (1994).
[11] J. A. X. Alexander, T. P. Orlando, D. Rainer, and P. M. Tedrow, Phys. Rev. B **31**, 5811 (1985).
[12] E. V. Thuneberg, J. Kurkijärvi, and D. Rainer, Phys. Rev. B **29**, 3913 (1984).
[13] M. Friesen and P. Muzikar, Phys. Rev. B **53**, R11953 (1996).

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

# Indirect determination of the ratio $R=\sigma_L/\sigma_T$ at small $x$ from HERA data

A. V. Kotikov[1)]

*Particle Physics Laboratory, Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia*

G. Parente[2)]

*Departamento de Física de Partículas, Universidade de Santiago de Compostela, 15706 Santiago de Compostela, Spain*

(Submitted 20 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 32–36 (July 1997)

The value of the deep inelastic scattering cross-section ratio $R=\sigma_L/\sigma_T$ is found in the range $10^{-4}\leqslant x\leqslant 10^{-2}$ from $F_2$ and $dF_2/d\ln Q^2$ HERA data using very simple relations based on perturbative QCD. © *1997 American Institute of Physics.* [S1063-7761(97)00307-7]

## 1. INTRODUCTION

In recent years the behavior of deep inelastic lepton–hadron scattering at small values of the Bjorken variable $x$ has been studied in depth. One of the many interesting deep inelastic scattering variables is the ratio of cross-sections of the absorption of a longitudinally and transversely polarized photon by a hadron: $R=\sigma_L/\sigma_T$. The ratio $R$, which can be represented as a combination of the longitudinal $F_L(x,Q^2)$ and transverse $F_2(x,Q^2)$ deep inelastic scattering functions,

$$R(x,Q^2)=\frac{F_L(x,Q^2)}{F_2(x,Q^2)-F_L(x,Q^2)}, \tag{1}$$

is a very sensitive QCD characteristic because it vanishes for free quarks. At small values of $x$, $R$ data are not yet available, as they require a rather cumbersome procedure (see Ref. 1, for example) for the extraction from the experiment.

We study the behavior of $R(x,Q^2)$ at small values of $x$, using the H1 data[2,3] and the method[4] of replacement of the Mellin convolution by ordinary products. By analogy with the case of the gluon distribution function (see Refs. 3 and 5–7), it is possible to obtain a relation between $F_L(x,Q^2)$, $F_2(x,Q^2)$, and $dF_2(x,Q^2)/d\ln Q^2$ at small $x$. Thus, the small $x$ behavior of the ratio $R(x,Q^2)$ can be extracted directly from the measured values of $F_2(x,Q^2)$ and its derivative. These extracted values of $R$ may be well considered new small-$x$ ''experimental data''.[3)] Moreover, when accurate experimental data for $R$ at small $x$ become available, the breakdown of this exactly perturbative relation will be indicative of the importance of other effects, such as higher twist contributions or nonperturbative QCD dynamics at small $x$.

We follow the notation of our previous work.[7,8] The singlet quark $s(x,Q_0^2)$ and gluon $g(x,Q_0^2)$ parton distributions[4)] at some $Q_0^2$ are parameterized (see, for example, Ref. 9) by

$$p(x,Q_0^2)=A_p x^{-\delta}(1-x)^{\nu_p}(1+\epsilon_p\sqrt{x}+\gamma_p x) \tag{2}$$

(hereafter $p=s,g$).

Further, we restrict the analysis to the case of large $\delta$ (i.e., $x^{-\delta}\gg 1$) following recent H1 data.[2] A more complete analysis concerning the extraction of the longitudinal structure function $F_L(x,Q^2)$ can be found in Ref. 8, where we also took into account the case $\delta\sim 0$ corresponding to the standard pomeron.

Assuming Regge-like behavior for the gluon distribution and $F_2(x,Q^2)$ at $x^{-\delta}\gg 1$

$$g(x,Q^2)=x^{-\delta}\widetilde{g}(x,Q^2),\quad F_2(x,Q^2)=x^{-\delta}\widetilde{s}(x,Q^2),$$

we obtain the following equation for the $Q^2$ derivative of the structure function $F_2$:[5)]

$$\begin{aligned}\frac{dF_2(x,Q^2)}{d\ln Q^2}=&-\frac{1}{2}x^{-\delta}\sum_{p=s,g}[r_{sp}^{1+\delta}(\alpha)\widetilde{p}(0,Q^2)\\&+r_{sp}^{\delta}(\alpha)x\widetilde{p}\,'(0,Q^2)+O(x^2)],\\F_L(x,Q^2)=&\,x^{-\delta}\sum_{p=s,g}[r_{Lp}^{1+\delta}(\alpha)\widetilde{p}(0,Q^2)\\&+r_{Lp}^{\delta}(\alpha)x\widetilde{p}\,'(0,Q^2)+O(x^2)],\end{aligned} \tag{3}$$

where $r_{sp}^{\eta}(\alpha)$ and $r_{Lp}^{\eta}(\alpha)$ are the combinations of the anomalous dimensions of Wilson operators $\gamma_{sp}^{\eta}=\alpha\gamma_{sp}^{(0),\eta}+\alpha^2\gamma_{sp}^{(1),\eta}+O(\alpha^3)$ and Wilson coefficients[6)] $\alpha B_L^{p,\eta}\times(1+\alpha R_L^{p,\eta})+O(\alpha^3)$ and $\alpha B_2^{p,\eta}+O(\alpha^2)$ of the $\eta$ ''moment'' (i.e., the corresponding variables extended from integer arguments to noninteger ones):

$$\begin{aligned}r_{Ls}^{\eta}(\alpha)=&\,\alpha B_L^{s,\eta}[1+\alpha(R_L^{s,\eta}-B_2^{s,\eta})]+O(\alpha^3),\\r_{Lg}^{\eta}(\alpha)=&\,\frac{e}{f}\,\alpha B_L^{g,\eta}\left[1+\alpha\left(R_L^{g,\eta}-\frac{B_2^{g,\eta}B_L^{s,\eta}}{B_L^{g,\eta}}\right)\right]+O(\alpha^3),\\r_{ss}^{\eta}(\alpha)=&\,\alpha\gamma_{ss}^{(0),\eta}+\alpha^2(\gamma_{ss}^{(1),\eta}+B_2^{g,\eta}\gamma_{gs}^{(0),\eta}+2\beta_0 B_2^{s,\eta})\\&+O(\alpha^3).\\r_{sg}^{\eta}(\alpha)=&\,\frac{e}{f}\{\alpha\gamma_{sg}^{(0),\eta}+\alpha^2[\gamma_{sg}^{(1),\eta}+B_2^{s,\eta}\gamma_{sg}^{(0),\eta}+B_2^{g,\eta}(2\beta_0\\&+\gamma_{gg}^{(0),\eta}-\gamma_{ss}^{(0),\eta}]\}]+O(\alpha^3),\end{aligned} \tag{4}$$

and

$$\widetilde{p}\,'(0,Q^2)\equiv\frac{d}{dx}\widetilde{p}(x,Q^2)\quad\text{at}\quad x=0,$$

where $e=\Sigma_i^f e_i^2$ is the sum of squares of the quark charges.

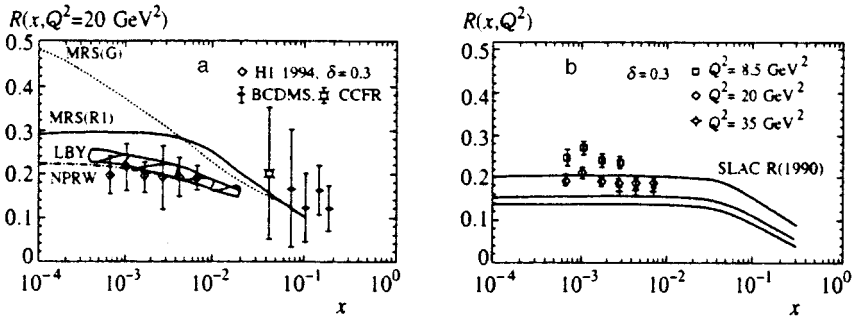



FIG. 1. The ratio $R=\sigma_L/\sigma_T$ at small $x$. The points were extracted from Eqs. (1) and (8) using H1[2,3] data. The dashed-dotted line (NPRW) is the prediction of the Saclay group[19] based on the dipole picture of BFKL dynamics. The band represents the uncertainty from the DGLAP analysis of HERA data in Ref. 18. Also shown are BCDMS data[11] points at high $x$ and the preliminary CCFR data point from Ref. 12. The solid lines (in Fig. 1b) are the SLAC R(1990) parametrization[20] at $Q^2=8.5$, 20 and 35 GeV$^2$ (lower curve corresponds to lower $Q^2$ value).

To accuracy $O(x^{2-\delta})$, we have for Eq. (3)

$$\frac{dF_2(x,Q^2)}{d\ \ln Q^2}=-\frac{1}{2}\left[r_{sg}^{1+\delta}(\xi_{sg})^{-\delta}g\left(\frac{x}{\xi_{sg}},\ Q^2\right)+r_{ss}^{1+\delta}F_2(x,Q^2)+(r_{ss}^{\delta}-r_{ss}^{1+\delta})\times x^{1-\delta}\widetilde{s}\,'(x,Q^2)\right]+O(x^{2-\delta}) \tag{5}$$

$$F_L(x,Q^2)=r_{Lg}^{1+\delta}(\xi_{Lg})^{-\delta}g\left(\frac{x}{\xi_{Lg}},\ Q^2\right)+r_{Ls}^{1+\delta}F_2(x,Q^2)+(r_{Ls}^{\delta}-r_{Ls}^{1+\delta})x^{1-\delta}\widetilde{s}\,'(x,Q^2)+O(x^{2-\delta}) \tag{6}$$

with $\xi_{sg}=r_{sg}^{1+\delta}/r_{sg}^{\delta}$ and $\xi_{Lg}=r_{Lg}^{1+\delta}/r_{Lg}^{\delta}$. From Eqs. (5) and (6) one can obtain $F_L$ as a function of both $F_2$ and the derivative:

$$F_L(x,Q^2)=-\xi^{\delta}\left[2\frac{r_{Lg}^{1+\delta}}{r_{sg}^{1+\delta}}\frac{dF_2(x\xi,Q^2)}{d\ \ln Q^2}+\left(r_{Ls}^{1+\delta}-\frac{r_{Lg}^{1+\delta}}{r_{sg}^{1+\delta}}r_{ss}^{1+\delta}\right)F_2(x\xi,Q^2)+O(x^{2-\delta},\alpha x^{1-\delta})\right], \tag{7}$$

where the result is restricted to $O(x^{2-\delta},\alpha x^{1-\delta})$. To arrive at the above equation we have performed the substitution

$$\xi_{sg}/\xi_{Lg}\to\xi=\gamma_{sg}^{(0),1+\delta}B_L^{g,\delta}/\gamma_{sg}^{(0),\delta}B_L^{g,l+\delta}$$

and neglected the term $\widetilde{s}\,'(x\xi_{sg},Q^2)$.

This replacement is very useful. The anomalous dimensions $\gamma_{sp}^{(1),n}$ in the next-to-leading order approximation (NLO) are singular at both, $n=1$ and $n=0$, and their presence in the arguments of $\widetilde{p}(x,Q^2)$, makes the numerical agreement between this approximate formula and an exact calculation worse (we have checked this point using some MRS sets[9] of parton distributions).

Using the NLO approximation of $r_{sp}^{1+\delta}$ and $r_{Lp}^{1+\delta}$ for the specific value $\delta=0.3$, we obtain (for $f=4$ and $\overline{\text{MS}}$ scheme)

$$F_L(x,Q^2)=\frac{0.84}{1+59.3\alpha}\left[\frac{dF_2(0.48x,Q^2)}{d\ \ln Q^2}+3.59\alpha F_2(0.48x,Q^2)\right]+O(\alpha^2,x^{2-\delta},\alpha x^{1-\delta}). \tag{8}$$

Notice that the $\alpha_s$ correction in the denominator of the factor on the right-hand side of Eq. (8) makes a large contribution. For example, at $Q^2=20$ GeV$^2$ the denominator is $1+0.92$.

With the help of Eqs. (1) and (8), we have extracted the ratio $R(x,Q^2)$ from H1 1994 data,[2] determining the slopes $dF_2/d\ \ln Q^2$ from straight line fits as in Refs. 3 and 5. In the present calculation, only statistical errors from those measurements have been taken into account. To estimate the systematic errors that have been added in quadrature, we have used those from an early analysis performed by H1.[3] In the calculation of the running coupling constant $\alpha_s(Q^2)$ at two loops we have taken $[(4)/\text{MS}]=225$ MeV.

Figure 1a shows the extracted ratio $R$ at $Q^2=20$ GeV$^2$ using the above formula for $\delta=0.3$. This value of $\delta$ is very close to those obtained by various groups from QCD fits to H1 data.[2,10] Fig. 1a also shows BCDMS[11] and preliminary CCFR (see, Ref. 12) data points with much larger errors.

For comparison we have also plotted various predictions for $R$ using QCD formulas at $O(\alpha_s^2)$[13–15 7)] and parton densities extracted from fits to HERA data. The large difference between the result from MRS(G) and the latest set MRS(R1)[17] shows the sensitivity of $R$ to the update of these parton densities to new HERA data. One can also notice that all these predictions remain higher than our extracted points.

On the other hand, recent theoretical calculations of $R$ based on the conventional NLO Dokshitzher–Gribov–Lipatov–Altarelli–Parisi (DGLAP) evolution analysis of HERA data (LBY)[18] and on the dipole picture of the Balitskiĭ–Fadin–Kuraev–Lipatov (BFKL) dynamics (NPRW)[19] are in very good agreement with our points obtained with Eq. (8).

Finally, Fig. 1b shows the extracted $R$ with $\delta=0.3$ at three different $Q^2$ values showing only statistical errors (to avoid the strong overlap between the data points at different $Q^2$ values), in comparison with the SLAC R(1990) parametrization[20] based on larger $x$ data. Relatively good agreement at $x\leqslant 10^{-2}$ is achieved when systematic errors are taken into account. Notice that points at the same $x$ and

different $Q^2$ are correlated by the form in which the derivative term $dF_2/d\ln Q^2$ is determined.

In summary, we have extracted the ratio $R=\sigma_L/\sigma_T$ at small $x$ from the structure function $F_2$ and its $Q^2$ derivative with the help of Eqs. (1), (7), and (8). These equations provide the possibility of an indirect determination of $R$. This is important since the direct extraction of $R$ from experimental data is a cumbersome procedure (see Ref. 1). Moreover, the fulfillment of Eqs. (1), (7), and (8) by deep inelastic scattering experimental data is a cross-check of perturbative QCD at small values of $x$. Our formulas can also be used as a parametrization of $R$ as a function of the most widely used phenomenological $F_2$.

The results depend on the specific value of the slope $\delta$. In the case $\delta=0.3$, which is very close to the values obtained by H1 group[2] in the considered $Q^2$ interval, we found relatively good agreement with the SLAC parametrization[20] and also very good agreement with the studies based on NLO DGLAP and BFKL dynamics (see Refs. 18 and 19, respectively). However, the calculation performed with the latest sets of HERA parton densities using perturbative QCD in second order (see MRS(R1) curve in Fig. 1a) predicts a slightly higher value of $R$.

This work was supported in part by CICYT and by Xunta de Galícia. We are grateful to J. W. Stirling for providing the parton distributions used in this work, and to A. Bodek, M. Klein, and F. J. Yndurain for discussions.

[1)]E-mail: KOTIKOV@SUNSE.JINR.DUBNA.SU
[2)]E-mail: GONZALO@GAES.USC.ES
[3)]Although with the theoretical prejudice contained in the starting QCD relation.
[4)]We use parton distributions multiplied by $x$ and neglect the nonsinglet quark distribution at small $x$.
[5)]Hereafter we use $\alpha(Q^2)=\alpha_s(Q^2)/4\pi$.
[6)]Because we consider here $F_2(x,Q^2)$ but not the singlet quark distribution.
[7)]The quark singlet and nonsinglet kernels in the $\overline{\text{MS}}$ scheme are taken from Ref. 13. It was noted in Ref. 16 that the gluon kernel given in Ref. 13 is erroneous. We use the correct result given in Refs. 14 and 15.

[1] A. M. Cooper-Sarkar, G. Ingelman, K. R. Long, R. G. Roberts, and D. H. Saxon, Z. Phys. C **39**, 281 (1988); M. W. Krasny, W. Placzek, and H. Spiesberger, Z. Phys. C **53**, 687 (1992); L. Favart, M. Gruwe, P. Marage, and Z. Zhang, Z. Phys. C **72**, 425 (1996).
[2] T. Ahmed, V. Andreev, B. Andrieu *et al.*, Nucl. Phys. B **470**, 3 (1996).
[3] T. Aid, V. Andreev, B. Andrieu *et al.*, Phys. Lett. B **354**, 494 (1995).
[4] A. V. Kotikov, Yad. Fiz. **57**, 142 (1994) [ Phys. At. Nucl. **57**, 133 (1994)]; Phys. Rev. D **49**, 5746 (1994).
[5] M. Derrick, D. Krakauer, S. Magill *et al.*, Phys. Lett. B **345**, 576 (1995); Z. Phys. C **65**, 379 (1995).
[6] K. Prytz, Phys. Lett. B **311**, 286 (1993); **332**, 393 (1994); R. K. Ellis, Z. Kunszt, and E. M. Levin, Nucl. Phys. B **420**, 517 (1994).
[7] A. V. Kotikov and G. Parente, Phys. Lett. B **379**, 195 (1996).
[8] A. V. Kotikov and G. Parente, Report No. US-FT-19-96 (E-prints archive hep-ph/9605207), submitted to Mod. Phys. Lett. B.
[9] A. D. Martin, W. S. Stirling, and R. G. Roberts, Phys. Lett. B **306**, 145 (1993); **354**, 155 (1995).
[10] K. Adel, F. Barreiro, and F. J. Yndurain, Report No. FTUAM 96-39 (E-prints archive hep-ph/9610380).
[11] A. C. Benvenuti, D. Bollini, G. Brani *et al.*, Phys. Lett. B **237**, 592 (1990).
[12] H. Abramowitz, *Test of QCD at Low x*, talk given at the EPS Conference, Warsaw (1996).
[13] J. Sánchez-Guillén, J. L. Miramontes, M. Miramontes, G. Parente, and O. A. Sampayo, Nucl. Phys. B **353**, 337 (1991).
[14] E. B. Zijlstra and W. L. van Neerven, Nucl. Phys. B **383**, 525 (1992).
[15] D. I. Kazakov and A. V. Kotikov, Phys. Lett. B **291**, 171 (1992).
[16] S. Larin and J. Vermaseren, Z. Phys. C **57**, 93 (1993).
[17] A. D. Martin, R. G. Roberts, and W. S. Stirling, Phys. Lett. B **387**, 419 (1996).
[18] C. Lopez, F. Barreiro, and F. J. Yndurain, Z. Phys. C **72**, 561 (1996).
[19] H. Navelet, R. Peschanski, Ch. Royon, and S. Wallon, Phys. Lett. B **385**, 357 (1996).
[20] L. W. Whitlow, S. Rock, A. Bodek, S. Dasu, and E. M. Riordan, Phys. Lett. B **250**, 193 (1990).

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

# Calculation of the effective permeability of a randomly inhomogeneous porous medium

É. V. Teodorovich[1)]

*Institute of Problems in Mechanics, Russian Academy of Sciences, 117526 Moscow, Russia*
(Submitted 11 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 313–324 (July 1997)

The effective permeability of a porous medium is calculated nonperturbatively. An exact expression for the permeability in terms of a double path integral is derived on the assumption that the permeability obeys a log-normal distribution function. Path integration is carried out in general form in the large-scale limit. The result confirms the well-known conjecture that the effective permeability is independent of the structure of the correlation function, but it disagrees with the hypothesis that the effective permeability depends exponentially on the variance.
 [S1063-7761(97)02707-8]

## 1. INTRODUCTION

The description of filtration flows in porous media poses a very timely problem in view of its applications to various problems in hydrogeology, petroleum recovery, chemical technology, etc. It is natural to regard a porous medium as randomly inhomogeneous and to specify its properties statistically by means of distribution functions or statistical moments and correlation functions of various orders. When the spatial scales of the random inhomogeneities (correlation lengths) are small in comparison with the characteristic filtration flow scales, for their description it is sufficient to know the effective permeability, which determines the relation between the average filtration flow rate and the average pressure gradient in accordance with Darcy's law.

The problem of calculating the permeability, which essentially entails investigating and finding statistical solutions to stochastic differential equations, has a long history. The simplest and most natural approach to its solution is by perturbation theory. The solution is represented by a power series in fluctuations of the permeability, and then each term of the series is averaged on the basis of certain assumptions regarding the nature of the statistics for the permeability fluctuations (a normal or log-normal distribution is customarily assumed).[1–3] This approach is similar to the method of description proposed by Wyld in the theory of turbulence, where the solution is represented by a functional power series in the density of random Langevin force sources modeling the onset of stochasticity due to the development of instability.[4] The authors of most papers have confined their calculations of the effective permeability to the first nonvanishing approximation of perturbation theory. The adopted boundary conditions stipulate that the pressure gradient is a given constant, and the problem is reduced to calculating the average filtration flow rate. Corresponding calculations have been carried out to estimate the role of higher approximations,[5,6] showing that the results of the second perturbation approximation do not contradict the hypothesis stated in Ref. 2: that the result of the first approximation is the first term of the Taylor series expansion of the exponential function describing the exact solution of the problem (at least in the case of a log-normal distribution of the permeability of the medium). It follows from the results obtained in the lowest perturbation approximation that the effective permeability does not depend on the form of the correlation function and is determined entirely by the variance of the fluctuating variable at the given point.[5]

The question arises, to what extent are these conclusions that the effective permeability depends exponentially on the variance and not on the form of the correlation functions bound to the application of perturbation theory? In particular, De Wit[7] has recently shown on the basis of calculations up to and including the third perturbation approximation that the hypothesis of an exponential dependence of the effective permeability on the variance is not justified and that in the case of arbitrary statistics the permeability in higher approximations is found to depend on the functional form of the pair correlation function of the permeability fluctuations and not merely on the variance. The description of the structure of any term of the permeability series and the refinement of permeability theory by the partial summation of certain infinite subsequences of the complete series can be implemented on the basis of methods analogous to Wyld's approach,[4] which are borrowed from quantum field theory and utilize Feynman diagrams, the Dyson equation, and a renormalization procedure. Working within the framework of the field approach and assuming that a log-normal distribution for the permeability is equivalent to a Gaussian distribution for small fluctuations, King[8] has reproduced the results obtained in the lowest permeability approximations, but at the same has succeeded in justifying the hypothesis that the effective permeability depends exponentially on the variance.

We look at the problem from the standpoint of finding the effective permeability outside perturbation theory, because in reality even the application of methods of quantum field theory is based on perturbation theory and the refinement thereof by partial summation of diagrams of a certain kind. A nonperturbative approach can be based on the representation of the solution of the stochastic differential equation in the form of a Feynman path integral previously constructed by the present author in the problem of the diffusion of a passive impurity in a field of random velocities[9] without relying on the assumption of smallness of the relative permeability fluctuations.

## 2. STATEMENT OF THE PROBLEM

The relation between the filtration flow rate $\mathbf{v}$ and the pressure gradient $p$ is given by Darcy's law, which is written as follows for an isotropic system:

$$\mathbf{v}(\mathbf{r}) = -k(\mathbf{r})\nabla p(\mathbf{r}), \tag{2.1}$$

where the susceptibility (permeability) coefficient $k(\mathbf{r})$ is a certain random function of the coordinates (the generalization to the case of an anisotropic medium requires special investigation). The condition of continuity for an incompressible fluid is given by the equation

$$\nabla \cdot \mathbf{v}(\mathbf{r}) = \rho(\mathbf{r}), \tag{2.2}$$

where $\rho(\mathbf{r})$ is the number density of real deterministic sources or equivalent sources governing the boundary conditions of the problem. It follows from Eqs. (2.1) and (2.2) that the pressure is the solution of the equation with random coefficients

$$\nabla \cdot [k(\mathbf{r})\nabla] p(\mathbf{r}) = -\rho(\mathbf{r}). \tag{2.3}$$

We define the effective permeability operator of the medium by the relation

$$\langle \mathbf{v}(\mathbf{r}) \rangle = -\hat{k}_{\text{eff}} \langle \nabla p(\mathbf{r}) \rangle . \tag{2.4}$$

In the general case $\hat{k}_{\text{eff}}$ is a certain integral operator of the form

$$\langle \mathbf{v}(\mathbf{r}) \rangle = -\int d\mathbf{r}_1 k_{\text{eff}}(\mathbf{r},\mathbf{r}_1) \langle \nabla p(\mathbf{r}_1) \rangle. \tag{2.5}$$

For a statistically homogeneous medium the kernel $k_{\text{eff}}(\mathbf{r},\mathbf{r}_1)$ depends only on the difference in the coordinates, and the transformation to Fourier transform space takes Eq. (2.5) into the form

$$\langle \mathbf{v}(\mathbf{q}) \rangle = k_{\text{eff}}(\mathbf{q}) i\mathbf{q} \langle p(\mathbf{q}) \rangle.$$

The objective of the present study is to calculate $k_{\text{eff}}(\mathbf{q}()$ and its Fourier transform $k_{\text{eff}}(\mathbf{r}-\mathbf{r}_1)$.

Averaging (2.2), we then obtain from this equation

$$\langle \mathbf{v}(\mathbf{r}) \rangle = \nabla^{-1} \rho(\mathbf{r}) + \text{curl } \mathbf{A}, \tag{2.6}$$

where the operator $\nabla^{-1} = \nabla(\nabla^2)^{-1}$ is well-defined, since $(\nabla^2)^{-1}$ represents the Green's function for the Laplace equation. With the exception of certain exotic conditions for an anisotropic medium we can set $\mathbf{A} = 0$ in Eq. (2.6)).

As a result, we obtain the expression for the Fourier transform for the inverse effective permeability operator

$$k_{\text{eff}}^{-1}(\mathbf{q}) = \frac{q^2 \langle p(\mathbf{q}) \rangle}{\rho(\mathbf{q})}, \tag{2.7}$$

and the problem reduces to the calculation of the Fourier transform of the average pressure for a given density of sources.

To facilitate the ensuing discussion, we introduce the notation $k(\mathbf{r}) = k_0 e^{u(\mathbf{r})}$. In this expression $k_0$ and $u(\mathbf{r}) = \ln[k(\mathbf{r})/k_0]$ are not uniquely defined, and the arbitrariness in their definition will be exploited in writing an expression for the characteristic functional of the random function $u(\mathbf{r})$. As a result, the pressure equation assumes the form

$$[\nabla^2 + \nabla u \cdot \nabla] p = -k_0^{-1} e^{-u} \rho. \tag{2.8}$$

The formal solution of this equation can be represented by the relation

$$p = k_0^{-1} (-\nabla^2 - \nabla u \cdot \nabla)^{-1} e^{-u} \rho. \tag{2.9}$$

Averaging Eq. (2.9), we obtain an explicit representation for $\langle p(\mathbf{r}) \rangle$, which can be used to find the effective permeability according to (2.4) or (2.7).

## 3. CONSTRUCTION OF THE INVERSE OPERATOR

To construct the inverse operator $(-\nabla^2 - \nabla u \cdot \nabla)^{-1}$ (the Green's function), we invoke the Feynman operator formalism, which essentially entails the following. In the case of an equation with constant coefficients, $\nabla u = \text{const}$, the inverse operator can be written in the operator exponential form proposed by Fock:

$$(-\nabla^2 - \nabla u \cdot \nabla)^{-1} = \int_0^\infty d\tau \, \exp\{-(-\nabla^2 - \nabla u \cdot \nabla)\tau\}. \tag{3.1}$$

which must be interpreted as an expansion in an infinite Taylor series in powers of the operators. In our case, however, where the coefficients of the operator are not constant, the operators $\nabla u(\mathbf{r})$ and $\nabla$ do not commute, $(\partial_i) \cdot (\partial_j u) - (\partial_j u) \cdot (\partial_i) = (\partial_i \partial_j u)$, and in the series expansion of the exponential function is it necessary to specify the order of operation of the operators of differentiation $\partial_i$ and multiplication by the function of the coordinates $\partial_j u$. Following Feynman,[10] we introduce the ordering parameter $s$ (''proper time''), assuming that the operators act in whatever sequence corresponds to increasing value of the parameter $s$.

Once the order of the noncommuting operators has been defined, they can be treated as numbers, whereupon the inverse operator can be written in the form

$$(-\nabla^2 - \nabla u \cdot \nabla)^{-1} = \int_0^\infty d\tau \, \exp\left\{ -\int_0^\tau ds[-\nabla^2(s) - \nabla u(\mathbf{r},s)\nabla(s)] \right\}. \tag{3.2}$$

To ''disentangle'' the operator exponential, it is necessary to dispose of the square of the operator $\nabla(s)$ in the exponential, whereupon the remaining expression can be interpreted as an operator of translation of the argument of the function according to the relation

$$\exp\{\mathbf{a}\nabla\} f(\mathbf{r}) f(\mathbf{r}+\mathbf{a}). \tag{3.3}$$

For this purpose we use the Stratonovich transform,[11] which is a functional analog of the well-known Weierstrass transform in the theory of integral transforms[12] In application to our situation (see Appendix A) this transform has the form

$$\exp\left\{ \int_0^\tau ds[\nabla^2(s) + \nabla u(\mathbf{r},s)\nabla(s)] \right\}$$

$$= A^{-1}\left(\frac{1}{4},\tau\right) \int d[\mathbf{X}(s)] \exp\left\{ -\frac{1}{4}\int_0^\tau ds[\mathbf{X}(s)\right.$$

$$+\nabla u(\mathbf{r},s)]^2\Big\}\exp\Big\{-\int_0^\tau ds\mathbf{X}(s)\nabla(s)\Big\}, \tag{3.4}$$

$$A(\alpha,\tau)=\int d[\mathbf{X}(s)]\exp\Big\{-\alpha\int_0^\tau ds\mathbf{X}^2(s)\Big\}.$$

Equation (3.4) involves an integral over all possible $d$-components of the vector functions $\mathbf{X}(s)$ defined on the interval $0\leqslant s\leqslant\tau$. The ambiguity in the definition of the measure $d[\mathbf{X}(s)]$, which specifies the weight of the paths, is unimportant, because expression (3.4) contains an integral with respect to the measure in the numerator and the denominator [in terms of $A(\alpha,\tau)$].

Equations (2.9), (3.2), (3.4) and the rules governing the action of the translation operators reduces the argument (3.3) to the following representation of the solution for $p(\mathbf{r})$:

$$p(\mathbf{r})=k_0{}^{-1}\int_0^\infty d\tau\, A^{-1}\Big(\frac{1}{4},\tau\Big)\int d[\mathbf{X}(s)]\rho\Big(\mathbf{r}-\int_0^\tau ds\mathbf{X}(s)\Big)$$
$$\times\exp\Big\{-\frac{1}{4}\int_0^\tau ds\Big[\mathbf{X}(s)+\nabla u\Big(\mathbf{r}-\int_0^\tau ds'\mathbf{X}(s')\Big)\Big]^2$$
$$-u\Big(\mathbf{r}-\int_0^\tau ds'\mathbf{X}(s')\Big)\Big\} \tag{3.5}$$

[the explicit dependence of $u(\mathbf{r},s)$ on the proper times $s$ is omitted from (3.5), because no noncommuting operators are left in the exponential].

The evaluation of $\langle p(\mathbf{r})\rangle$ reduces to averaging (3.5) over the ensemble of realizations of the permeability. In the averaging operation it is convenient to avoid the square of the expression containing $\nabla u(\mathbf{r}-\int_s^\tau ds'\mathbf{X}(s'))$ in the exponent. To do so, once again we invoke the functional Weierstrass transform

$$\exp\Big\{-\frac{1}{4}\int_0^\tau ds\Big[\mathbf{X}(s)+\nabla u\Big(\mathbf{r}-\int_0^\tau ds'\mathbf{X}(s')\Big)\Big]^2\Big\}$$
$$=A^{-1}(1,\tau)\int d[\mathbf{Y}(s)]\exp\Big\{-\int_0^\tau ds\Big[\mathbf{Y}^2(s)$$
$$+i\mathbf{Y}(s)\cdot\mathbf{X}(s)+i\mathbf{Y}(s)\nabla u\Big(\mathbf{r}-\int_s^\tau ds'\mathbf{X}(s')\Big)\Big]\Big\}. \tag{3.6}$$

As a result, we obtain

$$\langle p(\mathbf{r})\rangle=k_0^{-1}\int_0^\infty d\tau\, A^{-1}\Big(\frac{1}{4},\tau\Big)A^{-1}(1,\tau)$$
$$\times\int d[\mathbf{X}(s)]d[\mathbf{Y}(s)]\exp\Big\{-\int_0^\tau ds[\mathbf{Y}^2(s)$$
$$+i\mathbf{Y}(s)\cdot\mathbf{X}(s)]\Big\}\rho\Big(\mathbf{r}-\int_0^\tau ds\mathbf{X}(s)\Big)$$
$$\times\Big\langle\exp\Big\{i\int_0^\tau ds[-\mathbf{Y}(s)\nabla+2i\delta(s)]u\Big(\mathbf{r}$$
$$-\int_0^\tau ds'\mathbf{X}(s')\Big)\Big\}\Big\rangle. \tag{3.7}$$

It is readily verified that the ensemble average of the random permeability field in (3.7) admits the representation

$$\Psi[\theta(\mathbf{x})]=\Big\langle\exp\Big\{i\int d\mathbf{x}\theta(\mathbf{x})u(\mathbf{x})\Big\}\Big\rangle, \tag{3.8}$$
$$\theta(\mathbf{x})=\int_0^\tau ds[\mathbf{Y}(s)\nabla+2i\delta(s)]\delta\Big(\mathbf{x}-\mathbf{r}+\int_s^\tau ds'\mathbf{X}(s')\Big)$$

and can be interpreted as a characteristic (generating) functional of the random field $u(\mathbf{x})$ and, as such, is the functional Fourier transform of the distribution function for $u$.

Further progress can be made under certain assumptions about the nature of the statistics for the permeability field $k(\mathbf{x})$. If we assume that $k$ has a log-normal distribution for $u=\ln[k(\mathbf{x})/k_0]$ (Refs. 3, 6, and 8, then the characteristic functional is easily calculated, and in the case $\langle u\rangle=0$ it has the form

$$\Psi[\theta(\mathbf{x})]=\exp\Big\{-\frac{1}{2}\int d\mathbf{x}\,d\mathbf{x}'\theta(\mathbf{x})\theta(\mathbf{x}')B(\mathbf{x},\mathbf{x}')\Big\}, \tag{3.9}$$

where $B(\mathbf{x},\mathbf{x}')$ is the pair correlation function of the random field $u(\mathbf{x})$, and

$$B(\mathbf{x},\mathbf{x}')=\langle u(\mathbf{x})u(\mathbf{x}')\rangle, \tag{3.10}$$

which in our case of a statistically homogeneous system depends only on the difference in its arguments. It is always possible to satisfy the condition $\langle u\rangle=0$ by virtue of the latitude in the choice of $k_0$.

Calculating the Fourier transform for $\langle p(\mathbf{r})\rangle$ with allowance for the fact that the functional $\Psi[\theta]$ in (3.7) is independent of $\mathbf{r}$, substituting $\langle p(\mathbf{q})\rangle$ into (2.7), and using Eqs. (3.8) and (3.9), we obtain an expression for the Fourier transform of the reciprocal effective permeability:

$$k_{\text{eff}}^{-1}(\mathbf{q})=k_0^{-1}q^2\int_0^\infty d\tau\, A^{-1}\Big(\frac{1}{4},\tau\Big)A^{-1}(1,\tau)$$
$$\times\int d[\mathbf{X}(s)]d[\mathbf{Y}(s)]\exp\Big\{-\int_0^\tau ds[\mathbf{Y}^2(s)$$
$$+i\mathbf{Y}(s)\cdot\mathbf{X}(s)+i\mathbf{q}\cdot\mathbf{X}(s)]\Big\}\exp\Big\{-\frac{1}{2}$$
$$\times\int_0^\tau ds[-\mathbf{Y}(s)\nabla+2i\delta(s)]$$
$$\times\int_0^\tau ds'[\mathbf{Y}(s')\nabla+2i\delta(s')]B\Big(\int_{s'}^s ds''\mathbf{X}(s'')\Big)\Big\}. \tag{3.11}$$

We note that Eq. (3.11) is exact under the sole condition that the permeability field is assumed to have log-normal statistics; no assumptions have been made about smallness of the permeability fluctuations. The independence of the effective permeability from the density of sources $\rho$ indicates that this quantity characterizes the medium and has no connection with the properties of the filtration flows.

#### 4. EFFECTIVE PERMEABILITY IN THE LARGE-SCALE LIMIT

Normally the case of interest is the one in which the characteristic scales of the filtration flow are large in com-

parison with the scales of the permeability inhomogeneities (large-scale limit). In the large-scale limit we can set $\langle p(\mathbf{r}')\rangle \simeq \langle p(\mathbf{r})\rangle$ in the integrand of Eq. (2.5) and take this expression outside the integral sign. The latter approximation corresponds to the substitution

$$k_{\text{eff}}(\mathbf{r}-\mathbf{r}') \to \delta(\mathbf{r}-\mathbf{r}') \int d\mathbf{r}' k_{\text{eff}}(\mathbf{r}-\mathbf{r}') \equiv \delta(\mathbf{r}-\mathbf{r}') k_{\text{eff}}(\mathbf{q})|_{\mathbf{q}=0}. \tag{4.1}$$

Consequently, the large-scale limit for $k_{\text{eff}}$ is obtained by passing to the limit $q \to 0$ in Eq. (3.11), a nontrivial operation in view of the nonanalyticity of the integral on the right–hand side of (3.11) with respect to $q$ at the point $q = 0$. When such a transition is made in (3.11), one can make the substitutions $\tau \to \tau/q^2$, $s \to s/q^2$ and then $\mathbf{X}(s/q^2) \to q\mathbf{X}(s)$, $\mathbf{Y}(s/q^2) \to q\mathbf{Y}(s)$. The results of these substitutions are the disappearance of the factor $q^2$ on the right-hand side of (3.11) and the substitution

$$\mathbf{q} \to \mathbf{q}_0 = \frac{\mathbf{q}}{|\mathbf{q}|}, \quad B\left(\int_{s'}^{s} ds''\mathbf{X}(s'')\right) \to B\left(\frac{1}{q}\int_{s'}^{s} ds''\mathbf{X}(s'')\right).$$

The correlation function $B(\mathbf{x})$ is nonvanishing in an interval on the order of the correlation length. The substitution $B(\mathbf{x}) \to B(\mathbf{x}/q)$ in the limit $q \to 0$ corresponds to the case when the correlation length tends to zero. As a result, only minimum-length paths $x = \int_{s'}^{s} ds''|\mathbf{X}(s'')|$ are given in $B(\mathbf{x})$. Inasmuch as minimum-length paths correspond to motion with a constant velocity $\mathbf{X}(s) = \text{const}$, we can set $\mathbf{X}(s) \approx \mathbf{X}_0$ in the argument of the correlation function in (3.11) in the large-scale limit.

Functional integration can be carried out completely with the aid of Fourier series in this approximation.[13] Accordingly, we represent the vector functions $\mathbf{X}(s)$ and $\mathbf{Y}(s)$ in the interval $(0,\ \tau)$ by the expansions

$$\mathbf{X}(s) = \mathbf{X}_0 + \sum_{n=1}^{\infty} \mathbf{X}_n \cos\frac{\pi n s}{\tau}, \quad \mathbf{Y}(s) = \mathbf{Y}_0 + \sum_{n=1}^{\infty} \mathbf{Y}_n \cos\frac{\pi n s}{\tau}, \tag{4.2}$$

so that path integration is replaced by integration of infinite multiplicity with respect to the coefficients $\mathbf{X}_0$, $\mathbf{Y}_0$, $\mathbf{X}_n$, and $\mathbf{Y}_n$.

The integrations with respect to $\mathbf{X}_n$ are carried out by means of the readily verifiable relation

$$A_n^{-1}\left(\frac{1}{4},\tau\right) A_n^{-1}(1,\tau) \int d\mathbf{X}_n \exp(i\mathbf{Y}_n\cdot\mathbf{X}_n\tau) = \delta(\mathbf{Y}_n), \tag{4.3}$$

which facilitates the subsequent integrations with respect to $\mathbf{Y}_n$. Making use of the identity

$$\nabla B[(s-s')\mathbf{X}_0] = \frac{\mathbf{X}_0}{\mathbf{X}_0^2}\frac{\partial}{\partial s} B[(s-s')\mathbf{X}_0] = -\frac{\mathbf{X}_0}{\mathbf{X}_0^2}\frac{\partial}{\partial s'} B[(s-s')\mathbf{X}_0] \tag{4.4}$$

and the rules of integration using delta functions, we then perform integration with respect to $s$ and $s'$ to obtain

$$k_{\text{eff}}^{-1}(\mathbf{q}) = k_0^{-1} e^{B(0)/2} \int_0^{\infty} d\tau\, A^{-1}\left(\frac{1}{4},\tau\right) A^{-1}(1,\tau) \times \int d\mathbf{X} d\mathbf{Y} \exp\{-(\mathbf{Y}^2 + i\mathbf{Y}\cdot\mathbf{X} + i\mathbf{q}\cdot\mathbf{X})\tau\} \times \exp\left\{-\frac{\mathbf{Y}\cdot\mathbf{X}}{\mathbf{X}^2}\left(\frac{\mathbf{Y}\cdot\mathbf{X}}{\mathbf{X}^2} + i\right)[B(0) - B(\mathbf{X}\tau)]\right\} \tag{4.5}$$

[in Eq. (4.5) we have dropped the subscript 0 from $\mathbf{X}_0$, $\mathbf{Y}_0$, and $A_0(\alpha,\tau) = \int d\mathbf{X}_0 \exp(-\alpha\mathbf{X}_0^2\tau)$].

Introducing the substitution $\mathbf{Y} \to \mathbf{Y} - i\mathbf{X}/2$, we carry out the integration with respect to the $d$-dimensional vector $\mathbf{Y}$ by means of the relation

$$\int d\mathbf{Y} \exp\left\{-\mathbf{Y}^2\tau - \frac{(\mathbf{Y}\cdot\mathbf{X})^2}{\mathbf{X}^4}[B(0) - B(\mathbf{X}\tau)]\right\} = \left(\frac{\pi}{\tau}\right)^{d/2} \frac{1}{\sqrt{1 + [B(0) - B(\mathbf{X}\tau)]/\mathbf{X}^2\tau}}. \tag{4.6}$$

The rest of the integration with respect to $\mathbf{X}$ and the passage to the limit $q \to 0$ are conveniently executed by means of Fourier inversion for $k_{\text{eff}}^{-1}(q)$ with subsequent integration over the entire volume according to Eq. (4.1)).

The inverse Fourier transform produces $\delta(\mathbf{r} - \mathbf{X}\tau)$ in the integrand of Eq. (4.5), making it possible to integrate with respect to $\mathbf{X}$. As a result, we have

$$k_{\text{eff}}^{-1}(\mathbf{r}) = -\frac{k_0^{-1} e^{B(0)/2}}{(4\pi)^{d/2}} \nabla^2 \frac{1}{r^{d-2}} \varphi_d(\mathbf{r}), \quad \varphi_d(\mathbf{r}) = e^{-[B(0)-B(\mathbf{r})]/4} \int_0^{\infty} dt\ t^{(d-3)/2} [t + B(0) - B(\mathbf{r})]^{-1/2} e^{-t/4}, \tag{4.7}$$

where $t = r^2/\tau$. Integrating $k_{\text{eff}}^{-1}(\mathbf{r})$ over infinite volume and transforming the volume integral into a surface integral, we obtain

$$k_{\text{eff}}^{-1} \equiv k_{\text{eff}}^{-1}(q)|_{q=0} = \frac{k_0^{-1} e^{B(0)/2}}{(4\pi)^{d/2}} S_d \left[r^{d-1} \frac{\partial}{\partial r} \frac{\varphi(r)}{r^{d-1}}\right]_{r\to\infty}, \tag{4.8}$$

where $S_d = 2\pi^{d/2}/\Gamma(d/2)$ is the surface area of the $d$-dimensional sphere of unit radius.

In calculating the derivative in (4.8), it must be borne in mind that the derivative satisfies $\partial\varphi(r)/\partial r \sim \partial B(r)/\partial r$, which is assumed to decay faster than $r^{-1}$ as $r \to \infty$. As a result, taking into account the relation $\langle k^{-1}\rangle = k_0^{-1}\exp\{B(0)/2\}$ (see Appendix B), we find

$$k_{\text{eff}}^{-1} = \langle k^{-1}\rangle \frac{(d-2)S_d}{(4\pi)^{d/2}} \varphi_d(\infty). \tag{4.9}$$

It is evident from Eq. (4.9) that, since $B(\infty) = 0$, in the large-scale limit the effective permeability actually depends only on the quantity $B(0)$, which is related to the relative variance of the permeability $D$ by the equation $B(0)$

$= \ln(1+D)$ (see Appendix B). We have thus confirmed the conclusion on which perturbation theory is based, that the effective permeability of an isotropic system is independent of the structure of the correlation function.

In the general case of arbitrary dimensionality $d$ the integral $\varphi(\infty)$ defined by Eq. (4.7) is expressed in terms of Whittaker functions, which have singularities of the type $\Gamma[\pm(d-2)/2]$ for $d-2=0$, so that the coefficient $k_{\text{eff}}^{-1}$ is finite for $d=2$.

Of immediate interest here is the three-dimensional case ($d=3$), in which $\varphi_3(\infty)$ is expressed in terms of the probability integral. The resulting expression has the form

$$k_{\text{eff}}=\frac{\langle k\rangle}{(1+D)[1-\Phi(\sqrt{B(0)}/2)]}. \tag{4.10}$$

For small fluctuations we can set $B(0)\approx D$ and expand $k_{\text{eff}}$ in powers of $D$. Invoking the expansion of the probability integral $\Phi(x)\approx 2x/\sqrt{\pi}$, we see at once that the expansion contains half-integer powers of $D$, and the results of perturbation theory are not reproduced. This fact indicates that the perturbation expansion must be regarded as an asymptotic series, and the inclusion of higher approximations can significantly alter the picture of the kinds of laws involved when based on an analysis of results obtained in the lowest approximations. On the other hand, we should not rule out the possibility that the inferred asymptotic character of the perturbation series is associated with the use of a log-normal distribution for the permeability and that the application of perturbation theory might be justified in the presence of alternative statistics. All the same, we still point out the shortcomings of the traditional perturbation scheme whereby the boundary conditions correspond to the stipulation of a constant pressure gradient. At the same time, the renormalization of the pressure gradient has been disregarded. In our proposed approach the renormalization of the pressure is taken into account by the substitution of $\rho\exp(-u)$ for $\rho$ on the right-hand side of Eq. (2.8). Another reason for the nonreproducibility of the results of perturbation theory is that usually only the statistical moments of even orders are included in the term-by-term averaging of the functional series, even though the statistical moments of odd orders have nonzero values in the case of a log-normal distribution (see Appendix B).

Another conclusion that can be drawn from Eq. (4.9), which is exact in the large-scale limit and under the assumption of a log-normal distribution of the permeability, is the incorrectness of the assumption that the lowest perturbation approximations correspond to an expansion of the exponential function representing an exact solution of the problem of the effective permeability of a randomly inhomogeneous medium.

We close with the observation that the final result (4.10) can be applied without modification of any kind to the problem of the effective electrical conductivity of mixtures and alloys[14,15] and to a number of other problems.

## APPENDIX A: FUNCTIONAL WEIERSTRASS TRANSFORM

The Weierstrass transform is defined by the equation[12]

$$\exp\{a\mathbf{F}^2(s)\}=A^{-1}(a)\int d\mathbf{X}\,\exp\{-a[\mathbf{X}^2+2\mathbf{X}\cdot\mathbf{F}(s)]\}, \tag{A1}$$

$$A(a)=\int d\mathbf{X}\,\exp\{-a\mathbf{X}^2\}.$$

This transform can be used to express an exponential of the square of a certain function (vector or scalar) in terms of an exponential of the first power of this function.

If the function $\mathbf{F}(s)$ is specified in the interval $(0,\ \tau)$, we can partition the interval into $N$ subintervals of length $\Delta s_i$ and approximate $\mathbf{F}(s)$ by piecewise-constant functions $\mathbf{F}_i$. In the continuous limit $N\to\infty$, $\Delta s_i\to 0$ such an approximation is equivalent to specification of the function $\mathbf{F}(s)$.

Let us consider the identity

$$\exp\{\alpha\mathbf{F}_i^2\Delta s_i\}=A_i^{-1}(\alpha)\int d\mathbf{X}_i\,\exp\{-\alpha(\mathbf{X}^2+2\mathbf{X}_i\cdot\mathbf{F}_i)\Delta s_i\}, \tag{A2}$$

$$A_i(\alpha)=\int d\mathbf{X}\,\exp\{-\alpha\mathbf{X}^2\Delta s_i\}.$$

We form the product of the left-hand sides of Eqs. (A2) and pass to the continuous limit:

$$\lim_{N\to\infty}\prod_{i=1}^{N}\exp\{\alpha\mathbf{F}_i^2\Delta s_i\}=\lim_{N\to\infty}\exp\left\{\alpha\sum_{i=1}^{N}\mathbf{F}_i^2\Delta s_i\right\}$$

$$\equiv\exp\left\{\alpha\int_0^\tau ds\mathbf{F}^2(s)\right\}. \tag{A3}$$

Analogously, we write the product of the right-hand sides of Eqs. (A2) in the form

$$\lim_{N\to\infty}\prod_{i=1}^{N}\left\{\frac{d\mathbf{X}_i}{A_i(\alpha)}\exp\{-\alpha(\mathbf{X}_i^2+2\mathbf{X}_i\cdot\mathbf{F}_i)\Delta s_i\}\right\}$$

$$\equiv\int d[\mathbf{X}(s)]\exp\left\{-\alpha\int_0^\tau ds[\mathbf{X}^2(s)+2\mathbf{X}(s)\cdot\mathbf{F}(s)]\right\}, \tag{A4}$$

where the integral measure

$$d[\mathbf{X}(s)]=\lim_{N\to\infty}\prod_{i=1}^{N}\left\{\frac{d\mathbf{x}_i}{A_i(\alpha)}\right\}$$

defines the contribution of the ''path'' $\mathbf{X}(s)$ running through the sequence of points $\mathbf{X}_i$ the interval $d\mathbf{X}_i$, evaluated at the ''time'' $s_i$. In the continuous limit the measure integral must be regarded as a path integral. The concept of the path integral was first introduced by Einstein and Smoluchowski in the theory of Brownian motion and was later substantiated mathematically in the work of Wiener.

The resulting equation

$$\exp\left\{\alpha\int_0^\tau ds\mathbf{F}^2(s)\right\}$$

$$= A^{-1}(\alpha,\tau)\int d[\mathbf{X}(s)]\exp\int d[\mathbf{X}(s)]$$
$$\times\exp\left\{-\alpha\int_0^\tau ds[\mathbf{X}^2(s)+2\mathbf{X}(s)\cdot\mathbf{F}(s)]\right\},$$
$$A(\alpha,\tau)=\int d[\mathbf{X}(s)]\exp\left\{-\alpha\int_0^\tau ds\mathbf{X}^2(s)\right\} \tag{A5}$$

is a functional analog of the Weierstrass transform (A1). It was proposed by Stratonovich[11] and can be used to express the integral of the square of a certain function in the argument of an exponential in terms of an exponential containing an integral of the first power of this function.

## APPENDIX B: EQUATIONS FOR THE LOG-NORMAL DISTRIBUTION

The case of a log-normal distribution for $k(\mathbf{x})$ corresponds to a normal distribution for the quantity $u(\mathbf{x}) = \ln[k(\mathbf{x})/k_0]$. It is convenient to replace the distribution function for $u(\mathbf{x})$ by its functional Fourier transform or so-called characteristic functional:

$$\Psi[\theta(\mathbf{x})]=\left\langle\exp\left\{i\int d\mathbf{x}\ \theta(\mathbf{x})u(\mathbf{x})\right\}\right\rangle. \tag{B1}$$

Knowing the characteristic functional, we can find the statistical moments of $u(\mathbf{x})$ by means of the variational differentiation operations

$$\langle u(\mathbf{x}_1)\rangle=\left.\frac{\delta\Psi[\theta(\mathbf{x})]}{i\,\delta\theta(\mathbf{x}_1)}\right|_{\theta=0},$$
$$\langle u(\mathbf{x}_1)u(\mathbf{x}_2)\rangle=-\left.\frac{\delta^2\Psi[\theta(\mathbf{x})]}{\delta\theta(\mathbf{x}_1)\,\delta\theta(\mathbf{x}_2)}\right|_{\theta=0}. \tag{B2}$$

The statistical moments for $k(\mathbf{x})$ are calculated from the relations

$$\langle k(\mathbf{x}_1)\rangle=k_0\Psi[-i\,\delta(\mathbf{x}-\mathbf{x}_1)],$$
$$\langle k(\mathbf{x}_1)k(\mathbf{x}_2)\rangle=k_0^2\Psi[-i\,\delta(\mathbf{x}-\mathbf{x}_1)-i\,\delta(\mathbf{x}-\mathbf{x}_2)], \tag{B3}$$
$$\langle k^{-1}(\mathbf{x}_1)\rangle=k_0^{-1}\Psi[i\,\delta(\mathbf{x}-\mathbf{x}_1)],$$

etc. The characteristic functional for the log-normal distribution is specified by Eq. (3.9). In this case Eqs. (B3) give

$$\langle k(\mathbf{x})\rangle=k_0e^{B(0)/2},\langle k^{-1}(\mathbf{x})\rangle=k_0^{-1}e^{B(0)/2},$$
$$\langle k(\mathbf{x}_1)k(\mathbf{x}_2)\rangle=\langle k\rangle^2e^{B(\mathbf{x}_1-\mathbf{x}_2)}, \tag{B4}$$
$$\langle k(\mathbf{x}_1)k(\mathbf{x}_2)k(\mathbf{x}_3)\rangle=\frac{\langle k(\mathbf{x}_1)k(\mathbf{x}_2)\rangle\langle k(\mathbf{x}_1)k(\mathbf{x}_3)\rangle\langle k(\mathbf{x}_2)k(\mathbf{x}_3)}{\langle k\rangle^3}.$$

From (B4), in particular, we deduce the expression for the relative variance

$$D=\frac{\langle k^2\rangle}{\langle k\rangle^2}-1=e^{B(0)}-1,\quad B(0)=\ln(1+D). \tag{B5}$$

[1)]E-mail: teodor@ipm.net.ru

[1] A. L. Gutjahr, L. W. Gelhar, A. A. Bakr, and J. R. MacMillan, Water Resour. Res. **14**, 953 (1978).
[2] L. W. Gelhar and C. L. Axness, Water Resour. Res. **19**, 161 (1983).
[3] M. I. Shvidler, *Stochastic Hydrodynamics of Porous Media* [in Russian], Nedra, Moscow (1985).
[4] H. W. Wyld, Ann. Phys. (N.Y.) **14**, 143 (1961).
[5] P. Indelman and B. Abramovich, Water Resour. Res. **30**, 1857 (1994).
[6] G. Dagan, Transp. Porous Media **12**, 279 (1993).
[7] A. De Wit, Phys. Fluids **7**, 2553 (1995).
[8] P. R. King, J. Phys. A **20**, 4935 (1987).
[9] É. V. Teodorovich, Prikl. Mat. Mekh. **55**, 275 (1992).
[10] R. P. Feynman, Phys. Rev. **84**, 118 (1951).
[11] R. L. Stratonovich, Dokl. Akad. Nauk SSSR **115**, 1097 (1957) [Sov. Phys. Dokl. **2**, 416 (1957)].
[12] I. I. Hirschman and D. V. Widder, *The Convolution Transform*, Princeton Univ. Press, Princeton, N.J. (1955) [Russ. transl., IIL, Moscow (1958)].
[13] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York (1965) [Russ. transl., Mir, Moscow (1968)].
[14] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed. (rev. and enl., with L. P. Pitaevskii), Pergamon, Press, Oxford, New York (1984) [Russ. original, Nauka, Moscow (1982)].
[15] B. Abramovich and P. Indelman, J. Phys. A **28**, 693 (1995).

Translated by James S. Wood

# Role of decay processes in the propagation of slightly nonequilibrium thermal-pulse phonons in $YAlO_3$:Lu crystals

B. A. Danil'chenko and D. V. Poplavskiĭ

*Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia*

S. N. Ivanov, A. V. Taranov, and E. N. Khazanov

*Institute of Physics, National Academy of Sciences of Ukraine, 252142 Kiev, Ukraine*
(Submitted 15 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 325–331 (July 1997)

Data on the propagation of slightly nonequilibrium phonons, $\Delta T \ll T_b$ ($T_b$ is the cryostat bath temperature) in $Y_{0.9}Lu_{0.1}AlO_3$ crystals with effective elastic scattering by rare-earth metal atoms are analyzed. The experimental results are interpreted on the basis of the ''quasidiffusion'' model, where three-phonon inelastic decay processes play a decisive role in the phonon kinetics. Monte Carlo calculations with these processes taken into account yield good quantitative agreement with the experimental data. © *1997 American Institute of Physics.* [S1063-7761(97)02807-2]

## 1. INTRODUCTION

The kinetics of nonequilibrium phonons injected into a semiconductor or insulator crystal is determined by two main scattering processes: 1) elastic scattering by structural defects (including isotopes); 2) three-phonon decay processes, in which a high-frequency phonon decays into two lower-energy phonons.

Whereas elastic scattering processes are fairly efficient, nonequilibrium phonons injected into a crystal propagate diffusely, and in work with a thermal-pulse scheme[1] the bolometer signal has a characteristic diffusion-induced, bell-shaped profile. The analysis of this curve can yield information about the efficiency and nature of the scattering centers.[2] The diffusive motion of nonequilibrium phonons in the presence of elastic scattering can be influenced by their decay processes, which cause phonons to travel more rapidly through the crystal. This phenomenon is known as ''quasidiffusion'' and was first treated analytically in Refs. 3–5. In these and later papers,[6,7] including numerical modeling studies of the decay processes, the injected phonons are assumed to be highly nonequilibrium particles with respect to frequency, which is close to the Debye frequency. This hypothesis follows from a quantitative analysis of the efficiency of the elastic and decay mechanisms of phonon scattering for several well-known materials used in experiments using thermal pulses.

It is necessary to distinguish between the strong nonequilibrium of phonons with respect to their frequency $\omega$ and with respect to their occupation number $N(\omega)$. The foregoing discussion applies to the case $N(\omega) \ll 1$. For $N \gg 1$ the coalescence of phonons becomes significant, altering their kinetics. A qualitatively new phenomenon emerges here, the formation of a ''hot spot,'' which was first investigated theoretically in Ref. 8 and observed experimentally in Refs. 9 and 10.

Table I shows data on the efficiency of elastic scattering and decays for materials most commonly used in experiments on the propagation of nonequilibrium phonons.

In compiling the table, we have assumed that elastic scattering in the low-frequency range can be written in the form

$$1/\tau_B = BT^4, \tag{1}$$

where $T$ is the phonon energy in kelvins. For the injection of phonons into the investigated material with a Planck distribution we interpret $T$ as the energy of the dominant group of thermal-pulse phonons.[2]

The decay rate for anharmonic processes, averaged over the phonon polarizations and directions in $q$ space, can be written in the form

$$1/\tau_A = AT^5. \tag{2}$$

Maris[11] has introduced the dimensionless ratio of the efficiencies of elastic and anharmonic decays

$$B^* = B/(A^4 V_D / L)^{1/5}, \tag{3}$$

where $V_D$ is the averaged phonon velocity in the material of the sample, which has length $L$. For the data in Table I the ballistic transit time $t_{\rm bal} = L/V_D$ is taken equal to $10^{-6}$ s for all the materials, which is consistent in order of magnitude with typical experimental conditions ($L \simeq 1$ cm and $V_D = 10^5 - 10^6$ cm/s). It has been shown[11] that analytical calculations for the description of quasidiffusion processes are possible only under the condition

$$B^* \geqslant 3 \times 10^3. \tag{4}$$

It is evident from the table that this condition is not satisfied for conventional semiconductor-class materials, and numerical simulation is required to take account of the contribution of decay processes.

A significantly higher level of elastic scattering can be encountered in substitutional solid solutions such as (e.g.) yttrium rare-earth aluminum garnets and yttrium rare-earth aluminates. The closeness of the ionic radii of yttrium atoms and the rare-earth metal ion (e.g., Lu) ensures total mutual substitution, Y↔Lu, without any appreciable lattice distor-

TABLE I.

| Compound: | $CaF_2$ [Ref. 13] | Si [Ref. 10] | Ge [Ref. 10] | GaAs [Ref. 15] | $Y_3Al_5O_{12}$:$Lu_{0.3}$ [Ref. 1, 2] | $Y_{0.9}Lu_{0.1}AlO_3$ |
|---|---|---|---|---|---|---|
| $A$, $s^{-1}\times K^{-5}$ | $8\times10^{-4}$ | $1.6\times10^{-4}$ | $7.4\times10^{-4}$ | $39\times10^{-4}$ | $2.07\times10^{-4}$ | $48\times10^{-4}$ |
| $B$, $s^{-1}\times K^{-4}$ | 0.5 | 0.46 | 7.0 | 1.4 | 195 | $7\times10^3$ |
| $B^*$ | $\approx9.5$ | 31.5 | 141 | 7.5 | $\approx10^4$ | $3.16\times10^4$ |
| $T=T_{\min}$ K | $T\gg38$ K | $T\gg38$ K | $T\gg19$ K | $T\gg29$ K | $T\gg10$ K | $T\gg3.5$ K |

*Note:* The literature sources are given in brackets after each compound in the column headings; the data in the last column are from the present article.

tions, whereas the large difference in their masses creates a high level of elastic scattering and makes $B^*$ large. We have calculated the coefficients $A$ and $B$ for yttrium garnets and aluminates in accordance with a procedure proposed earlier[12,13] on the assumption that the decay rate is devoid of anisotropy in $q$ space.

The analytical theory[3–5] rests on the assumption that the relation between the times $\tau_B$, $\tau_A$, and $t_{\rm bal}$ satisfies the condition

$$\tau_B \ll \tau_A, t_{\rm bal}, \tag{5}$$

which can be used to estimate the range of temperatures (phonon energies) at which quasidiffusion is possible. The relation $\tau_B \ll \tau_A$ establishes an upper limit roughly equal to $T\approx10^3$ K for the materials listed in the table. A more significant factor is the lower limit

$$T_{\min}=(V_D/LB)^{1/4}, \tag{6}$$

which is given in the fourth row of the table. The quantity $V_D/L$ is again set equal to $10^6$ $s^{-1}$.

For the quasidiffusion effect to be observable for the majority of the materials investigated, it is necessary that $T\gg20$ K. It is evident from Fig. 1 (curve *1*) that this condition requires the injection of highly nonequilibrium phonons with a temperature close to the Debye temperature. In working with slightly nonequilibrium phonons injected while the phonon source is subjected to moderate heating ($\Delta T<0.5$ K), strong elastic scattering makes the quasidiffusion effect observable only for $YAlO_3$:Lu crystals, which are the object of the present investigation. It is evident from Fig. 1 (curve *2*) that the energy of the dominant group of phonons satisfies the condition from the table within certain error limits.

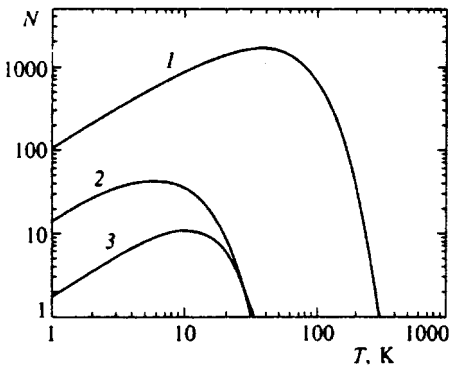



FIG. 1. Spectral density of the number of phonons injected into a crystal for $T_b=3.81$ K. *1)* $\Delta T\gg T$; *2)* $\Delta T\ll T$; *3)* equilibrium distribution of phonons in the hater at $T=3.81$ K.

## 2. EXPERIMENTAL METHODS

Here we investigate the propagation of a thermal pulse in a $Y_{0.9}Lu_{0.1}AlO_3$ crystal. The yttrium aluminate crystal have a distorted perovskite (orthorhombic) structure with space group $D_{2h}^{16}$. As a rule, the investigated samples were cut perpendicular to the $\langle001\rangle$ direction with an area of 1 $cm^2$ and a length of 0.5 cm in the direction of propagation of the injected phonons. The planes of the crystal were polished to optical standards; a gold film of thickness 800–1000 Å was deposited on one of the faces by thermal spraying to create a phonon injector. A bolometer in the form of a thin indium film was deposited on the opposite face in a meander pattern with an area of $0.3\times0.25$ $mm^2$ and an impedance of approximately 50 Ω.

The sample was placed in a helium cryostat, whose temperature was maintained by evacuation and vapor-pressure monitoring. This technique enabled us to work in the interval 1.7–3.8 K with a temperature instability less than or equal to $10^{-3}$ K.

Nonequilibrium phonons were generated by a short 100-ns current pulse, which heated the gold film. The meander-resistor bolometer was in a state close to the superconducting transition point and was used to detect phonons transmitted through the sample. The superconducting transition temperature of the bolometer was varied in accordance with the bath (cryostat) temperature by the application of a weak magnetic field.

In our experiments the power dissipated in the phonon injector ranged from $5\times10^{-3}$ $W/mm^2$ to $10^{-1}$ $W/mm^2$. The heater injected nonequilibrium Planck phonons with a heater temperature $T_h$, which is close to the bath temperature at such a low dissipated power level, i.e., $\Delta T=(T_h-T_b)\ll T_b$.

This ''quasiequilibrium'' phonon injection regime has definite advantages: From an analysis of the behavior of a slightly nonequilibrium thermal pulse, in garnets for example, we know the temperature of the dominant group of injected phonons forming the bolometer response signal, $(3.2–3.8)T_b$. Simple estimates show that the condition $T\gg3.5$ K holds over the entire temperature range, and the quasidiffusion effect should be appreciable.

Figure 2 shows typical plots of the bolometer response vs time at several thermostat temperatures. The graphs exhibit a typical diffusion profile. We see that the maximum of the bolometer signal and the temperature both decrease with time. Figure 3 shows the temperature dependence of the position of the diffusion maximum $T_{\max}$. The dependence of $t_{\max}$ on $T_b$ can be approximated by the relation $t_{\max}=CT^{1.8}$ (curve *1*) within the error limits of measurement of this

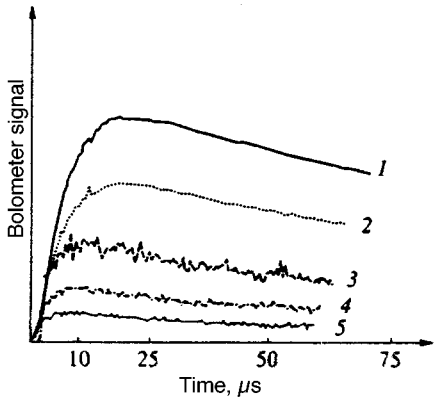



FIG. 2. Experimental time plots of the bolometer response for a $Y_{0.9}Lu_{0.1}AlO_3$ sample of length $L=0.6$ cm at various temperatures. *1*) 3.81 K; *2*) 3.38 K; *3*) 2.40 K; *4*) 2.31 K; *5*) 2.21 K.

quantity. Simple estimates in the limit $\Delta T \ll T_b$ indicate that a dependence of the form $t_{\max}=CT^4$ should be expected, as shown in Ref. 2. Since the authors of Ref. 2 proceed from a pure diffusion model, the resulting disparity can be attributed to the involvement of the decay process in the given material. Also shown in Fig. 3 are the results of a model experiment (curves *2* and *3*), which will be discussed below.

### 3. MONTE CARLO SIMULATION

In our simulation we solve the problem of the motion of phonons of a particular frequency from the center to the surface of a sphere. To determine the partial contribution of this group of phonons to the bolometer signal, we multiply the number of phonons reaching the surface by the sphere by their frequency. The number of phonons injected into the crystal at a given frequency is given by the equation

$$N(\omega)=\omega^2[(e^{\hbar\omega/kT_h}-1)^{-1}-(e^{\hbar\omega/kT_b}-1)^{-1}], \qquad (7)$$

where $T_b$ is the bath (crystal) temperature, and $T_h=T_b+\Delta T$ is the temperature of the heater. The value of $\Delta T$ is assumed to be close to the experimental values of 0.2 K for $T_b=2.21$ K and $\Delta T=0.4$ K at all other temperatures. To simulate the total bolometer signal, we sum the resulting partial contributions with respect to the initial frequency of the distribution (7). The experimental results shown in Fig. 2 represent an intermediate case between planar and point sources. The simulation results for a point source are reduced to the experimental geometry by multiplying the time scales by 1.9 according to the procedure set forth in Ref. 2.

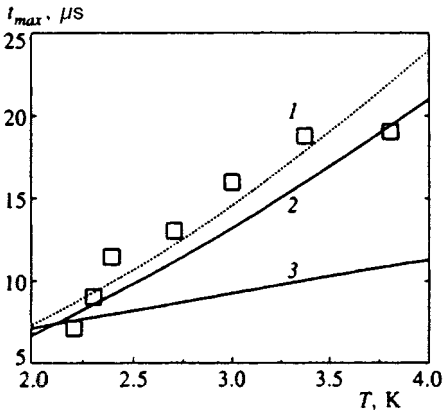



FIG. 3. Temperature dependence of the delay time of the maximum of the diffusion signal. *1*) Experimental result; *2*) simulation with decays; *3*) simulation without decays.

Simulation is carried out with only elastic scattering taken into account and again with both scattering mechanisms (elastic scattering and three-phonon decay processes) taken into account.

It is important to note that simulation with the spectral density of the number of phonons in the form (7) gives a sharp peak at small times. This peak corresponds to the partial contribution of low-frequency phonons, which do not decay ballistically and reach the detector. The absence of the ballistic peak in experiment can be attributed to two causes: 1) the escape of acoustic phonons into the liquid helium;[14] 2) ''cutoff'' of the spectrum of injected phonons in the low-frequency range, an effect observed for thin ($<2000$ Å) heaters.[15] To incorporate this phenomenon into the simulation, the spectrum of injected phonons is depleted at frequencies $\nu \leqslant 0.15$ THz.

Simulation is carried out for initial *LA* and *TA* phonons. The results do not depend on the polarization of the initial phonons, because rapid mode conversion takes place in the elastic scattering of phonons in the investigated crystal.

The Monte Carlo simulation process is a two-parameter problem, where the parameters are the elastic scattering and anharmonicity constants. The best agreement with experiment is obtained when these constants have the values

$$A=800\times10^{-4}\text{s}^{-1}\text{K}^{-5}, \quad B=2.5\times10^{3}\ \text{s}^{-1}\text{K}^{-4}.$$

Clearly, the elastic scattering constant is close to the theoretically calculated value (see Table I) in order of magnitude, whereas the anharmonicity constant, as the factor responsible for decay, is an order of magnitude higher than the theoretically calculated value. This disparity indicates the more significant role of anharmonicity in the investigated crystals. It is evident that $B^*=1.2\times10^3$ holds for the parameters obtained above. Hence, according to the criterion (4), it is impossible to solve the phonon ''quasidiffusion'' problem analytically, and analytical estimates could be incorrect, thereby justifying recourse to the Monte Carlo approach.

The results of simulation of the bolometer signal with all the above-stated conditions taken into account are shown in Fig. 4. A comparison of the simulation results shown in this figure reveals a substantial difference in the phonon fluxes obtained with (curves *1* and *2*) and without (curves *1′* and *2′*) regard for their decay. This difference is manifested in the values of the phonon fluxes and in the position of their maxima ($A_{\max}$) on the time scale ($t_{\max}$). If only elastic scattering processes are taken into account, the quantities $A_{\max}$ and $t_{\max}$ are always smaller than simulation results obtained with allowance for phonon decay processes, even though the number of injected phonons is the same in both cases.

At first glance the difference in the values of $t_{\max}$ appears to contradict the quasidiffusion model. According to this model, the decay of phonons should form a maximum of the phonon flux at shorter times than for the flux of nonde-

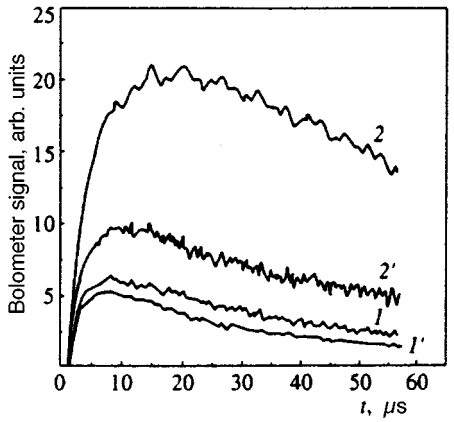



FIG. 4. Results of simulation of the bolometer signal at two crystal temperatures. *1, 2*) With decays at $T=2.21$ K and $T=3.81$ K, respectively; *1′, 2′*) without decays at $T=2.21$ K and $T=3.81$ K, respectively.

caying phonons. This conclusion is obvious in an analysis of the kinetics of phonons with a fixed initial frequency $\omega_0$ and has been demonstrated in the example of GaAs.[7] In our case the apparent contradiction is attributable to simulation of the total spectrum of injected phonons as opposed to simulation of the kinetics of monochromatic phonons as in Ref. 7.

As a matter of fact, in the case of simulation with a set of fixed phonon frequencies each frequency produces a result consistent with quasidiffusion theory. The summation of the partial contributions to the signal of phonons with the distribution function (7) leads to the formation of a signal maximum that is delayed relative to the pure diffusion signal. The presence of decay processes leads to redistribution of the long-time pure-diffusion phonon tail in the range of small arrival times. These considerations also account for the evolution of such an unusual profile of the time distribution of the phonon flux when phonon decay is taken into account.

Figure 3 shows how the position of the diffusion maximum on the time scale depends on the temperature of the crystal for the two cases under consideration; the experimental points are also plotted. Clearly, the experiment is described quite well by the curve obtained with allowance for decays. At lower values of $T_b$ the energy of the dominant phonon group is lower, so that the contribution of decay processes is smaller as well. This also accounts for the convergence of curves *1*, *2*, and *3* at low crystal temperatures, where decay processes no longer play a significant role in the phonon kinetics.

## 4. CONCLUSIONS

We have investigated the kinetics of slightly nonequilibrium phonons in $YAlO_3$:Lu crystals. We have shown that it cannot be described within the framework of phonon diffusion without regard for three-phonon decay processes. We have carried out a Monte Carlo simulation of the diffusion of nonequilibrium phonons with a continuous spectral distribution. The simulation results well describe the experimentally observed behavior for elastic scattering constants close to the calculated values, whereas the anharmonicity constant must exceed the value calculated from the known parameters of the crystal by more than an order of magnitude. For this reason the Grüneisen constant can be expected to exhibit anomalous behavior at low temperatures.

The authors are grateful to T. Paszkiewicz and I. Kaganova for valuable discussions and to I. Obukhov for assistance rendered in the Monte Carlo simulation.

This work has received support from the Russian Fund for Fundamental Research (Project No. 970216810).

[1] S. N. Ivanov, A. V. Taranov, and E. N. Khazanov, Zh. Éksp. Teor. Fiz. **99**, 1311 (1990) [Sov. Phys. JETP **72**, 731 (1990)].
[2] S. N. Ivanov, E. N. Khazanov, P. Paszkiewicz, A. V. Taranov, and M. Wilczynski, Z. Phys. B **99**, 535 (1996).
[3] D. V. Kazakovtsev and Y. B. Levinson, JETP Lett. **27**, 169 (1978).
[4] D. V. Kazakovtsev and Y. B. Levinson, Phys. Status Solidi B **96**, 117 (1979).
[5] Y. B. Levinson, in *Nonequilibrium Phonons in Nonmetallic Crystals*, W. Eisenmenger and A. A. Kaplyanskii (eds.), North-Holland, Amsterdam-New York (1986).
[6] P. Wolfe and G. A. Northrop, in *Phonon Scattering in Condensed Matter*, W. Eisenmenger, K. Lassmann, and S. Dottinger (eds.), Springer-Verlag, Berlin-New York (1984), p. 100.
[7] B. A. Danil'chenko, D. V. Kazakovtsev, and I. A. Obukhov, Zh. Éksp. Teor. Fiz. **106**, 1439 (1994) [JETP **79**, 777 (1994)].
[8] D. V. Kazakovtsev and I. B. Levinson, Zh. Éksp. Teor. Fiz. **88**, 2228 (1985) [Sov. Phys. JETP **61**, 1318 (1985)].
[9] B. A. Danilchenko, M. I. Slutskii, and D. V. Kazakovtsev, Phys. Lett. A **1–2**, 77 (1989).
[10] A. V. Akhimov, A. A. Kaplyanskiĭ, M. A. Pogarskiĭ, and V. K. Tikhomirov, JETP Lett. **43**, 333 (1986).
[11] H. J. Maris, Phys. Rev. B **41**, 9736 (1990).
[12] H. J. Maris and S. Tamura, Phys. Rev. B **47**, 727 (1993).
[13] H. S. Tamura, Phys. Rev. B **31**, 2574 (1985).
[14] B. A. Danilchenko and V. N. Poroshin, Cryogenics **23**, 546 (1983).
[15] B. A. Danil'chenko, I. A. Obukhov, and S. Kh. Rozhko, Zh. Éksp. Teor. Fiz. **98**, 256 (1990) [Sov. Phys. JETP **71**, 142 (1990)].

Translated by James S. Wood

# Microphase separation in multiblock copolymers

S. V. Panyukov

*P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia*

I. I. Potemkin

*M. V. Lomonosov Moscow State University, 119899 Moscow, Russia*

(Submitted 22 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 332–341 (July 1997)

The phase diagram of a multiblock copolymer containing domains of coexistence of phases having different superlattice symmetries with a density of monomers of a given type that varies periodically in space. The parameters of such superlattices are calculated in the mean-field approximation, and it is shown that their wave vector varies continuously with the temperature.
 [S1063-7761(97)02907-7]

## 1. INTRODUCTION

Heteropolymer systems consisting of two or more types of monomers are intriguing by virtue of their capability of forming periodic superlattices. Such structures were first predicted and subsequently studied in detail in the case of a polymer having molecules of the same species consisting of two blocks whose monomers ''do not like each other'' (Ref. 1). The possibility of the formation of superlattices by multiblock copolymers is still an open question. It has been hypothesized that such a structure is formed by a third-order transition.[2,3] It was later remarked that this continuous transition must be replaced by a first-order transition when the polymer composition is highly asymmetric.[4,5] Angerman *et al.*[5] have proposed a phase diagram of a multiblock copolymer consisting of two types of chains.

Multiblock copolymers differ from previously studied monodisperse polymer systems mainly in the possibility of local variations of their molecular-structural distribution, which is described by so-called nonlocal terms in the free-energy functional.[2,3,6] We intend to show that this phenomenon significantly alters the phase diagram of the system in comparison with the results of Ref. 5. In the mean-field approximation we construct the phase diagram of a multiblock copolymer consisting of two types of monomers with a Markov distribution of block lengths, and we show that the system goes through a series of first-order transitions into phases with different superlattice symmetries as the temperature of the system is varied. A distinctive feature of multiblock heteropolymers is the existence of temperature intervals in which different phases coexist, and their relative volume changes with the temperature. Both the amplitudes and the wave vectors of such superlattices are strongly dependent on the temperature.

## 2. FREE ENERGY

We consider a melt of heteropolymer chains with Markov statistics of units $A$ and $B$, for which the probability (transition matrix) $\nu_{ij}$ that a unit of type $j=A,B$ will follow a unit of type $i$ along the chain does not depend on units farther removed along the chain. The average density $\rho_i$ of units of a given type $i$ and the average block length $N$ for such units are expressed in terms of elements of the transition matrix $\nu$ (Ref. 3):

$$\overline{\rho_A}=\frac{\nu_{BA}}{\nu_{AB}+\nu_{BA}}\rho, \quad \overline{\rho_B}=\frac{\nu_{AB}}{\nu_{AB}+\nu_{BA}}\rho, \quad N=\frac{1}{\nu_{AB}+\nu_{BA}}\ , \tag{1}$$

We assume that the block length $N$ is small in comparison with the average length of the chain, i.e., each chain contains a large number of blocks. By the incompressibility condition we have $\rho_A(\mathbf{x})+\rho_B(\mathbf{x})=\rho$, where $\rho$ is the total density of monomers. The deviations $\Delta\rho_A(\mathbf{x})=-\Delta\rho_B(\mathbf{x})$ of the density of monomers of a given type from their average are conveniently described by the dimensionless order parameter

$$\psi(\mathbf{x})=\Delta\rho_A(\mathbf{x})/\rho, \quad \rho=\overline{\rho_A}+\overline{\rho_B}. \tag{2}$$

We show below that a critical point corresponding to a symmetrical polymer composition, $\overline{\rho_A}=\overline{\rho_B}$, exists in the investigated system. Near this point the transition to a spatially inhomogeneous phase takes place as a first-order, almost second-order transition. Consequently, the free energy of decay of multiblock chains can be expanded in powers of the order parameter (2) in the vicinity of this point. Since the expansion is made relative to the spatially homogeneous state, it is useful to consider the free energy of a system of given volume $V$ at a temperature $T$:

$$\begin{aligned}\frac{\mathcal{F}[\psi]}{\rho T}=&\frac{1}{2}\int\frac{d^3\mathbf{q}}{(2\pi)^3}(\chi_c-\chi+a^2q^2)\psi_{\mathbf{q}}\Psi_{-\mathbf{q}}\\&-\frac{\lambda}{6}\int\delta\Big(\sum_{i=1}^{3}\mathbf{q}_i\Big)\prod_{i=1}^{3}\psi_{\mathbf{q}_i}\frac{d^3\mathbf{q}_i}{(2\pi)^3}\\&+\frac{g}{24}\int\delta\Big(\sum_{i=1}^{4}\mathbf{q}_i\Big)\prod_{i=1}^{4}\psi_{\mathbf{q}_i}\frac{d^3\mathbf{q}_i}{(2\pi)^3}\\&+\frac{k}{4V}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\int\frac{d^3\mathbf{q}'}{(2\pi)^3}\frac{\psi_{\mathbf{q}}\psi_{-\mathbf{q}}\psi_{\mathbf{q}'}\psi_{-\mathbf{q}'}}{a^2[q^2+(q')^2]}\end{aligned} \tag{3}$$

as a functional of the Fourier components of the dimensionless order parameter

$$\psi_{\mathbf{q}}=\int d\mathbf{x}\psi(\mathbf{x})\exp(i\mathbf{q}\cdot\mathbf{x}). \tag{4}$$

The temperature dependence enters into the free energy through the dimensionless monomer interaction parameter $\chi$, which has a critical value $\chi_c$ corresponding to a spinode (stability threshold) of the spatially homogeneous state. We note that destabilization [corresponding to loss of the quadratic-$\psi$ term in the free energy (3)] takes place at $\chi=\chi_c$ for zero wave vector. It has been noted[2,3] that in the case of polydisperse polymer systems this fact does not imply the emergence of a new spatially homogeneous phase (macrophase separation of the polymer melt) as in the case of monodisperse systems.

The first three terms in Eq. (3) have the form of the usual Landau free-energy expansion, where the coefficient $\lambda$ of the cubic term vanishes at the critical point $\overline{\rho_A}=\overline{\rho_B}$. The last, ''nonlocal'' term ($\sim k$), which is not included in Ref. 7, describes effects associated with the presence of polydispersity due to the finite width of the block-length distribution. This term gives the energy cost due to local variation of the block-length distribution to create the density inhomogeneity $\Delta\rho_A(\mathbf{x})$ with a characteristic spatial scale of order the reciprocal wave vector of the inhomogeneity $q^{-1}$. Since the formation of such an inhomogeneity involves only blocks with a characteristic length $q^{-1}$, in effect it ''sucks in'' chains with such blocks from the surrounding space. In the case of polydisperse systems this phenomenon depletes the block-length distribution throughout the entire volume of the inhomogeneity and, as a consequence, enhances interaction of two (or more) inhomogeneities arbitrarily far apart. We shall not discuss the processes of relaxation to thermodynamic equilibrium on the scale $q^{-1}$, which can last quite a long time for small values of $q$.

The parameters $\lambda$, $g$, and $k$ of the Landau free-energy expansion (3) are determined from microscopic theory:[3]

$$\lambda=\frac{3}{4N}\frac{1-2f}{f^2(1-f)^2},\quad g=\frac{3}{8N}\frac{5-16f(1-f)}{f^3(1-f)^3},$$

$$k=\frac{1}{4N^2}\frac{1}{f^3(1-f)^3}, \tag{5}$$

where $N$ is the average number of monomers of per block, and $f=\overline{\rho_A}/\rho$ is the fraction of monomers of a given type $A$. In the case $f=1/2$ (i.e., $\overline{\rho_A}=\overline{\rho_B}$) the coefficient $\lambda$ of the cubic term vanishes by virtue of the symmetry of the system under a change of monomer type. The parameter $a=b(\chi_c N)^{1/2}$ can be expressed in terms of the monomer length $b$ and the critical value of the interaction parameter $\chi_c=1/2Nf(1-f)$.

### 3. PHASE DIAGRAM

In the presence of microphase separation the equilibrium value of the order parameter $\psi(\mathbf{x})$ is determined from the condition of minimization of the free energy (3). In the minimization operation we allow for the fact that the system can exist in either the single-phase or the two-phase state. We first consider the single-phase state of the polymer. Near the phase transition line the order parameter can be represented by a superposition of $n$ plane waves with wave vectors $\mathbf{q}_k$, $|\mathbf{q}_k|=q$ ($n=0$, 1, 3, 6 for the isotropic, lamellar, hexagonal, and body-centered cubic phases, respectively):[1]

$$\psi_{\mathbf{q}}|_n=\frac{A_n}{\sqrt{n}}\sum_{k=1}^{n}[\delta(\mathbf{q}-\mathbf{q}_k)+\delta(\mathbf{q}+\mathbf{q}_k)]. \tag{6}$$

Substituting this expression into Eq. (3), we represent the free energy of the phase with a given number $n$ of plane waves

$$\frac{\mathscr{F}_n}{T\rho V}=F_n[q,A_n]+\frac{k}{2}\frac{A_n^4}{a^2q^2} \tag{7}$$

by the sum of the local contribution

$$F_n[q,A_n]\equiv(\chi_c-\chi+a^2q^2)A_n^2-\frac{\lambda_n}{6}A_n^3+\frac{g_n}{24}A_n^4 \tag{8}$$

and the nonlocal contribution. Here $\lambda_1=0$, $g_1=6g$ for the lamellar phase, $\lambda_6=4\lambda\sqrt{6}/3$, $g_6=15g$ for the bcc phase, and $\lambda_3=4\lambda\sqrt{3}/3$, $g_3=10g$ for the hexagonal phase. Minimizing the resulting free energy with respect to the amplitude $A_n$ and the wave vector $q$, we find

$$A_n=\frac{3}{g_n}\left[\frac{\lambda_n}{2}-3\sqrt{2k}+\sqrt{\left(\frac{\lambda_n}{2}-3\sqrt{2k}\right)^2+\frac{4g_n}{3}(\chi-\chi_c)}\right], \tag{9}$$

$$aq=[kA_n/2]^{1/4}.$$

Consequently, in the single-phase domain both the amplitude and the wave vector of the structure increase monotonically with the interaction parameter $\chi$. It is important to note that for sufficiently large values of $\chi$ the period of the structure becomes of the order of the block length, and domain structures with very distinct boundaries appear in the investigated system. The harmonic approximation (6) is inadequate for the description of such structures in this case, and higher harmonics must also be included in the expansion of the order parameter $\psi(\mathbf{x})$.

To describe the transition between two phases $n$ and $m$, it is necessary to write the free energy of a system in which there are two coexisting phases with wave vectors $q_n$ and $q_m$. In each of these phases the Fourier component of the order parameter is given by Eqs. (6) with $n$ plane waves and with $m$ plane waves, respectively. Substituting this order parameter into Eq. (3), we find the free energy of the two-phase system:

$$\frac{\mathscr{F}_{nm}}{TV}=\phi F_n[q_n,A_n]+(1-\phi)F_m[q_m,A_m]+k\left(\frac{\phi^2A_n^4}{2a^2q_n^2}+\frac{2\phi(1-\phi)A_n^2A_m^2}{a^2(q_n^2+q_m^2)}+\frac{(1-\phi)^2A_m^4}{2a^2q_m^2}\right). \tag{10}$$

Here $\phi$ and $1-\phi$ are the volume fractions occupied by phases $n$ and $m$, respectively, and the dimensionless free energy $F_n$ is defined in Eq. (8). The first two terms of Eq.

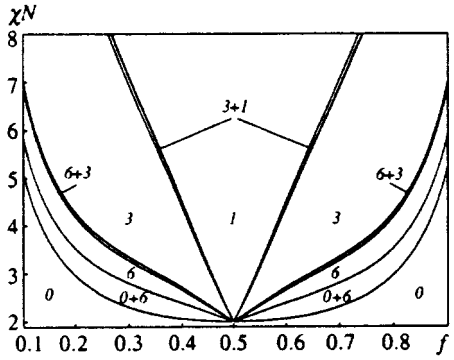



FIG. 1. Phase diagram of a Markovian multiblock copolymer melt in the variables $(\chi N, f)$, where $\chi$ is the interaction parameter, $N$ is the average number of monomers per block, and $f$ is the constituency of the polymer. The curves are numbered according to the domain of existence of different phases: *0*) isotropic; *1*) lamellar; *3*) hexagonal; *6*) body-centered cubic.

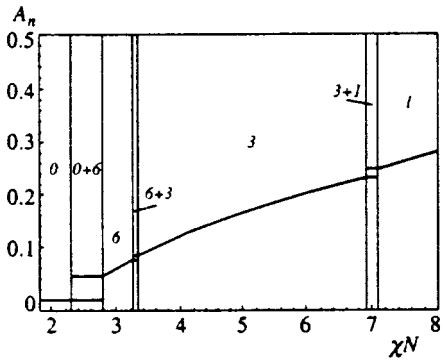



FIG. 3. Amplitude $A_n$ of superlattices vs temperature for $f=0.3$. The numbers have the same significance as in Fig. 1.

(10) give the local contributions of each phase to the free energy and are proportional to the volumes $V_n=\phi V$ and $V_m=(1-\phi)V$ of these phases. The last term describes effects of redistribution of chains with different block lengths between the two coexisting phases and includes a free-energy contribution proportional to the product of the volumes of these phases. The free energy (10) is minimized with respect to the parameters of such phases in Appendix A, where, in particular, it is shown that the wave vectors in the coexisting phases must have the same magnitude: $q_n=q_m=q$. The minimization of the functional (10) with respect to $\phi$ and $q$ gives the dependence of the wave vector on the interaction (temperature) parameter $\chi$:

$$3a^2q^2=\chi-\chi_c+\frac{\lambda_m}{4}A_m-\frac{g_m}{12}A_m^2, \tag{11}$$

along with the volume fraction of phase $n$:

$$\phi=\frac{2a^4q^4/k-A_m^2}{A_n^2-A_m^2}. \tag{12}$$

When two phases coexist their amplitudes $A_n$ and $A_m$ are determined from the condition for the minimumizations of the free energy (10). It is shown in Appendix A that they do not depend on $\chi$ or the fraction $\phi$. According to Eq. (12), the volume fraction of phase $n$ varies from 0 to 1 as the interaction parameter $\chi$ varies.

The phase diagram of the system in variables $(\chi N, f)$ is shown in Fig. 1. The numbers alongside the curves indicate different types of phases: *0*) isotropic; *1*) lamellar; *3*) hexagonal; *6*) bcc. This figure differs from the phase diagram of a melt of monodisperse diblock copolymers[1] primarily in having bands of coexistence of two phases. The existence of such domains in a system of polydisperse heteropolymers was first predicted in Ref. 8. Most notable is the broad domain of coexistence of the isotropic and bcc phases in conjunction with the relative narrowness of the domains of coexistence of all other phases. More detailed information on the properties of the coexisting phases is obtained from an investigation of the temperature dependence of their parameters — the wave vector $q$, whose magnitude is identical in both coexisting phases, and the amplitudes of the superlattices in each of these phases. The dependence of the wave vector of the superlattices on the parameter $\chi$ in the case $f=0.3$ is shown in Fig. 2 in variables $(q^2b^2N, \chi N)$. The vertical lines indicate the values of $\chi$ for the beginning and end of the corresponding transitions between different phases ($\phi=0$ and $\phi=1$, respectively). We note that $q$ increases continuously as the parameter $\chi$ increases, becoming of the order of the block length for large $\chi$. Figure 3 shows the $\chi$ dependence of the amplitudes $A_n$ for all phases in the case $f=0.3$. In the single-phase domains the amplitude of the superlattice increases monotonically with $\chi$. In the two-phase domains the amplitude is independent of $\chi$ and takes the two values $A_n$ and $A_m$ corresponding to the two coexisting phases $n$ and $m$. The volume fractions occupied by each phase are shown in Fig. 4 as functions of the interaction parameter $\chi$.

On the whole, we can state that allowance for the possibility of the coexistence of phases with different superlattice symmetries significantly alters the phase diagram of the system from its form in Ref. 5, where this effect is ignored.

We close this section with a discussion of the validity of the approximations made in the study. We note first of all that the expression for the free energy (3) is written in the ''infinitely long chain'' approximation, i.e., when the wave vector satisfies $q \gg l/bL^{1/2}$, where $L$ is the average number of

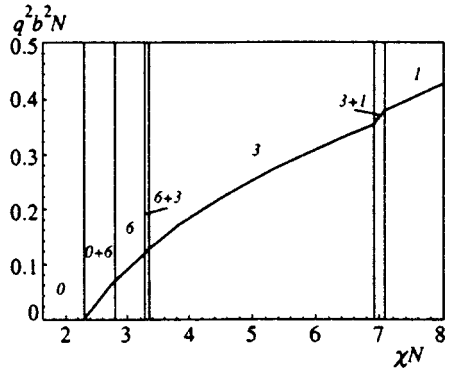



FIG. 2. Wave vector **q** of a periodic superlattice vs temperature (interaction parameter $\chi$) for $f=0.3$. The numbers have the same significance as in Fig. 1.

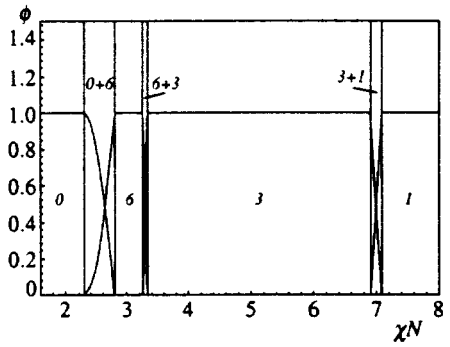



FIG. 4. Volume fractions $\phi$ occupied by each phase for $f=0.3$. The numbers have the same significance as in Fig. 1.

monomers in the molecule. If this condition does not hold, the curve in Fig. 2 will not begin with $q=0$ but with the quantity $q_{\min}\approx 1/(bL^{1/2})$. In our situation, where the chains contain a large number of blocks, $q_{\min}$ is small, and the finite-length effect can be disregarded. The finiteness of the chain lengths is also essential in the analysis of effects involving the separation of the system into macroscopic phases with different average densities of monomers of given types $A$ and $B$. Such separation is possible in a system with a small number of blocks $n_{bl}=L/N$, because the number of monomers of given type can differ for different chains by an amount of order $Nn_{bl}^{1/2}=(NL)^{1/2}$. The corresponding variation of the density of monomers of a given type for a system with an average density of chains $\rho/L$ is estimated as

$$\Delta\rho_A\approx(\rho/L)(NL)^{1/2}=\rho/n_{bl}^{1/2}.$$

For small $n_{bl}$ and a large interaction parameter $\chi$ conditions are favorable for the spatial separation of such chains into different macroscopic phases. In the opposite limit of large $n_{bl}\gg 1$ the variation of the density of monomers of two (or more) coexisting phases can be disregarded.

The influence of a variation of the densities of coexisting phases on the phase diagram of a diblock ($n_{bl}=2$) copolymer has been investigated previously[9] in the mean-field approximation. The authors of Ref. 9 ignore the contribution of the nonlocal term to the free energy of the two-phase system, but take into account the contribution of this term to the free energy of each phase. According to the discussion after (3) regarding the physical significance of nonlocality, this approximation is tantamount to the possibility of the exchange of molecules with different block lengths within each phase, but prohibits such exchange between different coexisting phases. Such a peculiar equilibrium can be only be established at small times; in the thermodynamic limit both channels of relaxation to equilibrium must be taken into account. On the other hand, it is shown in Appendix B that the variation of the average densities of monomers in each phase can be neglected in the case of multiblock copolymers with an average molecular length $L\gg N$, and the only effect that leads to a qualitative change in the phase diagram from the case of a monodisperse molt of diblock copolymers[1] is the nonlocality effect.

In deriving the free-energy expression (3), we have also disregarded the dependence of $\lambda$ and $g$ on the wave vectors. This approximation is admissible because in the domain of validity of the Landau expansion (3) the characteristic values of the wave vector of the superlattice are small in comparison with the reciprocal Gaussian block length. Consequently, the allowance for the corresponding dependences of the coefficients of the Landau expansion in Ref. 5 is a refinement that patently exceeds the computational accuracy.

Fluctuation corrections to the mean-field theory have been investigated[10] for the case of a symmetrical copolymer with $f=1/2$. It was shown that these corrections are small of order the parameter $1/N^{1/4}$ and result in the generation of a superlattice from the isotropic phase with a magnitude of the wave vector that is proportional to this parameter but is still nonzero.

## 4. CONCLUSION

We have shown in the self-consistent field approximation that as the temperature is varied, a melt of multiblock copolymers undergoes a succession of first-order transitions into microphase-separated states with a wave vector of finite magnitude, which varies continuously from zero to the characteristic block length $aN^{1/2}$; see Fig. 2. We note that a similar conclusion in Refs. 2 and 3 has been made on the assumption of a third-order phase transition, which actually describes the spinodal decomposition of the spatially homogeneous state of the system at $\chi=\chi_c$. It is readily shown that the free energy of our investigated two-phase states lies below the energy of the single-phase spatially inhomogeneous state. Consequently, a thermodynamic transition with the formation of a microphase-separated phase takes place through the formation of equilibrium nuclei of a new phase in the old phase. As the interaction parameter $\chi$ increases, the volume of these nuclei increases until the new phase occupies the entire volume of the system; see Fig. 4. We therefore have a finite temperature interval in which both phases coexist.

We note that from the standpoint of the general theory of phase transitions a melt of Markovian copolymers comprises a system with a frozen-in, or ''quenched'' randomness. It is well known that such systems exhibit a localization or glass type of behavior in the case of solids having a random distribution of interactions that is fixed in three-dimensional space. A fundamental distinction of polymer systems is the fact that only a one-dimensional sequence of monomers along the chain is random, whereas the chain can assume arbitrary spatial conformations and exist in any domain of the system space. The presence of such translational invariance is conducive to the formation of a regular three-dimensional structure in polymers with a random sequence of monomers along the chain. We shall not discuss the phenomenon of ordinary thermodynamic fluctuations, which lead to the breakdown of long-range crystal order on large scales in a process analogous to liquid-crystal systems.[11]

Our general expression (3) used for the free-energy functional is not bound to any specific choice of model of a Markovian copolymer, for which the parameters of the Ginzburg–Landau functional are given by expressions (5). Consequently, the main results of the present study are

equally valid for a multiblock copolymer of arbitrary structure.

## APPENDIX A: MINIMIZATION OF THE FREE ENERGY

We minimize the free energy (10) with respect to the parameters of both phases $n$ and $m$. In the event of transition from the isotropic phase with $n=0$ we have $q_0=0$, and the magnitude of the wave vector in the new phase is given by the expression

$$q_m^4=k(1-\phi)A_m^2/2a^4. \tag{A1}$$

We now consider the case of transition between phases with different superlattice symmetries with $n \neq 0$ and $m \neq 0$. The conditions for the minimum of the free energy with respect to the wave vectors $q_n$ and $q_m$ of these superlattices have the respective forms

$$\begin{aligned} &\frac{k\phi A_n^2}{2a^4q_n^4}+\frac{2k(1-\phi)A_m^2}{a^4(q_n^2+q_m^2)^2}=1,\\ &\frac{k(1-\phi)A_m^2}{2a^4q_m^4}+\frac{2k\phi A_n^2}{a^4(q_n^2+q_m^2)^2}=1. \end{aligned} \tag{A2}$$

A solution of these equations is readily found:

$$q^4\equiv q_n^4=q_m^4=k[\phi A_n^2+(1-\phi)A_m^2]/2a^4. \tag{A3}$$

We note that Eq. (A1) can be regarded as a special case of Eq. (A3), since $A_0=0$ in the isotropic phase. If the equality of the wave vectors in the coexisting phases with $n \neq 0$ and $m \neq 0$ is taken into account, along with Eq. (A3), the conditions for the minimum of the free energy (3) with respect to $A_n$ and $A_m$ assume the form

$$\begin{aligned} &2(\chi-\chi_c+3a^2q^2)-\lambda_nA_n/2+g_nA_n^2/6=0,\\ &2(\chi-\chi_c+3a^2q^2)-\lambda_mA_m/2+g_mA_m^2/6=0, \end{aligned} \tag{A4}$$

and the minimization of the free energy with respect to $\phi$ yields the equation

$$\begin{aligned} &(\chi-\chi_c+3a^2q^2)A_n^2-\lambda_nA_n^3/6+g_nA_n^4/24\\ &\quad=(\chi-\chi_c+3a^2q^2)A_m^2-\lambda_mA_m^3/6+g_mA_m^4/24. \end{aligned} \tag{A5}$$

Substituting the combination $\chi-\chi_c+3a^2q^2$ from the two equations (A4) into the right- and left-hand sides of Eq. (A5), respectively, we find equations for the amplitudes $A_n$ and $A_m$:

$$\begin{aligned} &-3\lambda_nA_n+g_nA_n^2=-3\lambda_mA_m+g_mA_m^2,\\ &-2\lambda_nA_n^3+g_nA_n^4=-2\lambda_mA_m^3+g_mA_m^4. \end{aligned} \tag{A6}$$

The most important consequence of these equations is the fact that neither of the amplitudes depends on $\phi$ and $\chi$. We arrive at the conclusion that the amplitudes of the two coexisting phases do not vary with the temperature.

In the case of transition from the Lamellar to the hexagonal phase, $\lambda_1=0$, and the solution of Eq. (A6) has the form

$$\begin{aligned} A_1&=\frac{\lambda_3}{g_3}\left[\frac{15}{2}(7+3\sqrt{6})\right]^{1/2}\simeq 10.37\,\frac{\lambda_3}{g_3},\\ A_3&=\frac{\lambda_3}{g_3}\left(6+\frac{3}{2}\sqrt{6}\right)^{1/2}\simeq 9.67\,\frac{\lambda_3}{g_3}. \end{aligned} \tag{A7}$$

and for transition from the hexagonal to the cubic phase the amplitudes $A_3$ and $A_6$ are found by solving Eqs. (A6) numerically:

$$\begin{aligned} A_3&\simeq 3.48\,\frac{\lambda_3}{g_3},\\ A_6&\simeq 3.17\,\frac{\lambda_3}{g_3}. \end{aligned} \tag{A8}$$

Solving the second of the equations (A4) for $q$, we find the temperature dependence of the superlattice wave vector; see Eq. (11) in the main text of the article. Equation (12), which together with (11) describes the dependence of the volume fraction of new phase $\phi$ on the interaction parameters, is determined by solving Eq. (A3) for $\phi$.

## APPENDIX B: ESTIMATION OF THE VARIATION OF THE AVERAGE DENSITY IN COEXISTING PHASES

We show that the variation of the average density in coexisting phases can be ignored for polymer molecules containing a large number of blocks, $n_{bl}=L/N \gg 1$. The physical picture of the absence of this effect is discussed in Sec. III. To describe such a variation of density on macroscopic scales, it is necessary to include the zeroth harmonic in the expansion of the order parameter (6) in each phase:

$$\psi_{\mathbf{q}}|_n=\Delta_n\delta(\mathbf{q})+\frac{A_n}{\sqrt{n}}\sum_{k=1}^{n}[\delta(\mathbf{q}-\mathbf{q}_k)+\delta(\mathbf{q}+\mathbf{q}_k)]. \tag{B1}$$

The amplitudes $\Delta_n$ and $\Delta_m$ characterize the variation of the average density in each of the coexisting phases $n$ and $m$ and are related by the condition of invariance of the total number of molecules in the system

$$\phi\Delta_n+(1-\phi)\Delta_m=0, \tag{B2}$$

which can be used to parametrize these amplitudes by a single parameter $\Delta$:

$$\Delta_n=(1-\phi)\Delta, \quad \Delta_m=-\phi\Delta. \tag{B3}$$

Here $\phi$ and $1-\phi$ are the volume fractions occupied by phases $n$ and $m$, respectively. As mentioned in Sec. III, allowance for the finite length of the polymer chain ''removes'' the singularity of the nonlocal term in the free energy (3) for zero wave vector $\mathbf{q}=0$ (Ref. 3). Consequently, substituting the order parameter for each of the coexisting phases in the form (B1) and (B3) into the functional (3), we obtain the expression for the free energy of the two-phase system

$$\frac{\mathscr{F}_{nm}}{TV}=F_{nm}+\delta F(\Delta),\quad F_{nm}\equiv\frac{\mathscr{F}_{nm}}{TV}\Big|_{L\to\infty},$$
$$\delta F(\Delta)=-\lambda\phi(1-\phi)(A_n^2-A_m^2)\Delta+\frac{k}{4}\frac{\phi^2(1-\phi)^2\Delta^4}{a^2q_{\min}^2}+\ldots, \tag{B4}$$

where the free energy $F_{nm}$ in the limit $L\to\infty$ is calculated in (10), $q_{\min}^2=1/b^2L$, and the expansion of $\delta F$ is written in the principal approximation with respect to $\Delta$ and $L$. Minimizing $\delta F$ with respect to the parameter $\Delta$ and using the expression (A3) for $q$ derived in Appendix A, we obtain the estimates

$$F_{nm}\sim k^{1/2}[\phi A_n^2+(1-\phi)A_m^2]^{3/2},$$
$$\delta F\sim(Lk)^{-1/3}\lambda^{4/3}[\phi(1-\phi)]^{2/3}|A_n^2-A_m^2|^{4/3}. \tag{B5}$$

The variation of the average density in the coexisting phases can be ignored if the following inequality holds:

$$F_{nm}\gg\delta F. \tag{B6}$$

For transition from the isotropic to an ordered phase it is necessary to set $A_n=0$, $A_m\approx\lambda/g$, and $1-\phi\ll1$. Inequality (B6) therefore assumes the form

$$1-\phi\gg(N/L)^{2/5}|1-2f|^{6/5}. \tag{B7}$$

In the limit of infinitely long molecules, which is in fact the model treated in the present article, inequality (B6) is always satisfied, and the variation of the average density in this limit can therefore be ignored. In the transition from one ordered phase ($n$) to another ordered phase ($m$), inequality (B6) assumes the form

$$1-\phi\ll(L/N)^{1/2}|1-2f|^{-3/2} \tag{B8}$$

and is always satisfied for sufficiently long molecules (with $L\geqslant N$).


[1] L. Leibler, Macromolecules **13**, 1602 (1980).
[2] E. I. Shakhnovich and A. M. Gutin, J. Phys. (France) **50**, 1843 (1989).
[3] S. V. Panyukov and S. I. Kuchanov, Zh. Eksp. Teor. Fiz. **99**, 659 (1990) [Sov. Phys. JETP **72**, 368 (1990)].
[4] C. D. Sfatos, E. I. Shakhnovich, and A. M. Gutin, Phys. Rev. E **51**, 4727 (1995).
[5] H. Angerman, G. Brinke, and I. Erukhimovich, Macromolecules **29**, 3255 (1996).
[6] G. Fredrickson, S. Milner, and L. Leibler, Macromolecules **25**, 6341 (1992).
[7] C. Burger, W. Ruland, and A. Semenov, Macromolecules **23**, 3339 (1990).
[8] S. V. Panyukov and S. I. Kuchanov, JETP Lett. **54**, 501 (1991).
[9] I. Entkhimovich and A. Dobrynin, Macromol. Symp. **81**, 253 (1994).
[10] S. B. Panyukov and I. I. Potemkin, JETP Lett. **64**, 197 (1996).
[11] P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, 2nd ed., Clarendon Press, Oxford (1993) [Russ. transl. of 1st ed. (first author only), Mir, Moscow (1982)].


Translated by James S. Wood

# Resonant Josephson tunneling through S-I-S junctions of arbitrary size

I. A. Devyatov and M. Yu. Kupriyanov[1)]

*Institute for Nuclear Physics Research, M. V. Lomonosov Moscow State University, 119899 Moscow, Russia*

(Submitted 6 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 342–352 (July 1997)

The Josephson tunneling current in S-I-S structures where the main current transport channel is resonant tunneling through an isolated localized state is calculated using the Bogolyubov–de Gennes equations. It is shown that the efficiency of equilibrium Josephson resonant tunneling is determined only by the ratio of the width of the resonance level to the absolute value of the order parameter for the superconducting electrodes with arbitrary relationships among the system parameters. © *1997 American Institute of Physics.* [S1063-7761(97)03007-2]

## 1. INTRODUCTION

Experimental studies of Josephson HTSC junctions with a semiconducting oxide spacer layer have led to the observation of the ''long-range proximity effect.'' [1–10] It involves the existence of a significant critical current $J_c$ in structures with spacer layer thicknesses $d \geqslant 100$ nm. Experimental $J_c(d) \propto \exp(-d/\xi)$ curves have yielded temperature independent coherence lengths $\xi$ that are considerably longer (10–50 nm) in these materials than in the superconducting oxides (1–3 nm). This effect has been explained[11] in terms of a model based on the assumption that the main mechanism for transport of the normal and superconducting currents in these structures is resonant tunneling through localized states in a low energy tunnel barrier. It is found that the characteristic scale length for the reduction in the superconducting current with increasing spacer thickness was determined by the effective penetration depth $\xi = \alpha$ of quasiparticles into the spacer. As opposed to the coherence length in a superconductor, $\xi_0 = \hbar v_F / 2\Delta$, where $v_F$ is the velocity at the Fermi surface and $\Delta$ is the absolute value of the order parameter for the superconducting electrons, the characteristic length $\alpha$ is independent of the parameters responsible for the superconducting properties of the electrodes and its relatively large values are a consequence of the low energy barrier of the spacer material,

$$\alpha = \frac{1}{\sqrt{2m(V-\mu)}} \approx 100-1000 \ \text{Å}, \tag{1}$$

as well as of a double gain in the exponent owing to the resonant character of the current transport (for the normal current). In Eq. (1) $m$ is the effective mass of the charge carriers, $\mu$ is the chemical potential, and $V$ is the potential of the bottom of the conduction band of the spacer material.

On the other hand, it is known[12] that the nonequilibrium properties of the weak bonds may depend on the effective coherence length $\xi_i$ determined by both the superconducting properties of the electrodes and by the spacer parameters.

In this paper resonant Josephson tunneling in S-I-S structures is analyzed in a physically clear approach based on the Bogolyubov–de Gennes equations and three characteristic lengths describing the penetration of superconductivity into these structures are found to exist. Besides depending on the length $\alpha$ (Eq. (1)), which actually determines the effective penetration depth of normal quasiparticles into the subbarrier region, the solutions depend on the parameters $\xi_i$ and $\xi_b$:

$$\xi_i = \alpha \frac{V-\mu}{E} = \frac{1}{E}\sqrt{\frac{V-\mu}{2m}}, \quad \xi_b = 2\alpha \ln\left(\frac{V-\mu}{E}\right), \tag{2}$$

where $E$ is the energy of the bound Andreev state responsible for transport of the superconducting current. In an earlier theoretical analysis[11,13] of resonant Josephson tunneling through a dielectric barrier it was assumed implicitly that the barrier is thin on the scale of the lengths $\xi_i$ and $\xi_b$. In this paper we examine the case of an arbitrary relationship among $d$, $\xi_i$, and $\xi_b$ and refine the ranges of applicability of the results obtained in the earlier work.[11,13] It is shown that in the limit $\Gamma \gg \Delta$, where $\Gamma$ is the effective width of the resonance level, the direct and Andreev reflection channels are equally effective and the expression for the resonant Josephson current is the same as the analogous formula for a microscopic short circuit. In the opposite limit of a narrow resonance level, $\Gamma \ll \Delta$, the anomalous proximity effect is accompanied by a reduction in the critical current by a factor of $\Delta/\Gamma$ owing to the mismatch of the resonance channels for direct and Andreev reflection of the quasiparticles. Then it is found that for structures with an arbitrary length $d$, as opposed to short ($d \ll \xi_i, \xi_b$) structures, not just one, but three channels for Andreev penetration by quasiparticles may operate during transport of a superconducting current. Nevertheless, their interference may lead to exact cancellation of the terms that depend on $\xi_i$ and $\xi_b$ in the formula for the superconducting current, and this yields the same value for the superconducting current as the one calculated for short junctions.

## 2. MODEL FOR THE JUNCTION

We shall assume that the density of localized states in the spacer layer is low, so that the interaction among quasiparticles belonging to different states is unimportant, their effective interaction with the electrons is negligibly small, and they are uniformly distributed, both over the volume of

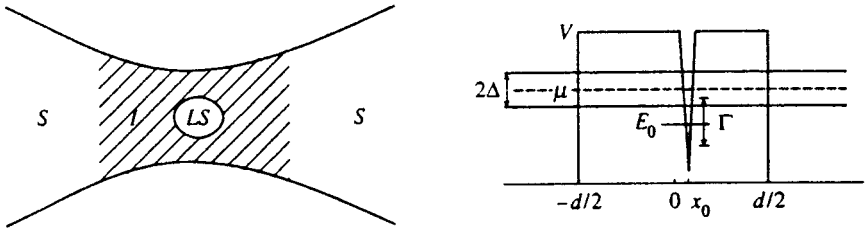



FIG. 1. Schematic and energy diagram of the structure being discussed here.

the spacer and with respect to energy (at least, in a region of order $T_c$ near the chemical potential $\mu$). To retain the physical clarity of this picture, we restrict ourselves below to considering the one dimensional problem in a single mode regime, i.e., to studying processes in an S-c-S structure with tunnel conductivity in the constriction region and with a potential barrier of the form (see Fig. 1)

$$V(x)=V\Theta(|x|-d/2)-(B/2m)\,\delta(x-x_0). \tag{3}$$

Here $x_0$ is the coordinate of the localized state described by a $\delta$-function potential with a single allowed energy state with an energy of $E_0=B^2/8m$.

### 3. TRANSMISSION COEFFICIENTS AND THE ENERGY OF THE BOUND STATES

Under the above assumptions the magnitude of the order parameter in superconductors can be regarded as independent of the spatial coordinates and the processes taking place in this system can be described using the 1D Bogolyubov–de Gennes equations (here and in the following $\hbar=1$ and $k_B=1$):

$$\left\{\left(-\frac{1}{2m}\frac{\partial^2}{\partial x^2}-\mu+V(x)\right)\sigma_z+\hat{\Delta}(x)\right\}\Psi=E\Psi,$$

$$\hat{\Delta}(x)=\begin{pmatrix}0 & \Delta e^{i\varphi/2}\\ \Delta e^{-i\varphi/2} & 0\end{pmatrix}\Theta(|x|-d/2). \tag{4}$$

Here $\sigma_z$ is the third Pauli matrix, $\varphi$ is the phase of the order parameter for the superconducting electrons, $E$ are the energy eigenvalues, and $\Psi$ is a spinor composed of the $u$- and $v$- functions of Bogolyubov and de Gennes. The solution of Eqs. (4) in the neighborhood of the spacer and superconducting electrodes can be written in the form of a superposition of plane waves transmitted and reflected from the boundaries of the junction and localized state. By matching these plane waves and their derivatives, we have found the direct $t^N$ and Andreev $t^A$ transmission coefficients for the quasiparticles:[12–15]

$$|t^A|=\frac{k_+}{k}\frac{\tilde{D}}{Q}\frac{|\zeta|}{\Delta}\,|(E-|\zeta|)e^{i\varphi/2}a_+-(E+|\zeta|)e^{-i\varphi/2}a_-|,$$

$$|t^N|=\frac{k_+}{k}\frac{\tilde{D}}{Q}\frac{|\zeta|}{\Delta}\,|e^{i\varphi/2}b_+-e^{-i\varphi/2}b_-|, \tag{5}$$

where

$$Q=(\cos\,\alpha-\cos\,\beta)-2(E/\Delta)^2\,\cos\,\beta$$
$$-2i(E/\Delta^2)|\zeta|\sin\,\beta,\quad \zeta=\sqrt{\Delta^2-E^2}. \tag{6}$$

The bound states $|E|<\Delta$ are determined from the condition $Q=0$, which yields a dispersion equation of the form[12]

$$E^2=\Delta^2\,\cos^2\left(\frac{|\beta|\pm\alpha}{2}\right),\quad \cos\,\alpha=\tilde{R}+\tilde{D}\cos\,\varphi,$$

$$\sin\,\beta=\mathrm{Im}(a_+a_-^*)\tilde{D},\quad |\beta|\pm\alpha\geqslant 0. \tag{7}$$

The quantities $\tilde{D}=1/|a_+a_-^*|$ and $\tilde{R}=\mathrm{Re}(b_+b_-^*)\tilde{D}$ in Eq. (7) depend on the values of the energy that are symmetrically positioned relative to the chemical potential $\mu$ and represent transmission and reflection coefficients for the quasiparticles through the direct and Andreev channels. The parameters $a_\pm$ and $b_\pm$ are given by

$$a_\pm=\left\{[B_\pm\,\sinh(\kappa_\pm d)-\cosh(\kappa_\pm d)]+\frac{i}{2}\left[\left(\frac{\kappa_\pm}{k_+^*}-\frac{k_+^*}{\kappa_\pm}\right)\right.\right.$$
$$\times[\sinh(\kappa_\pm d)-B_\pm\cosh(\kappa_\pm d)]-\left(\frac{\kappa_\pm}{k_+^*}\right.$$
$$\left.\left.\left.+\frac{k_+^*}{\kappa_\pm}\right)B_\pm\,\cosh(\kappa_\pm x_0)\right]\right\}\frac{k_+^*}{k}, \tag{8}$$

and

$$b_\pm=\frac{1}{2k}\left\{(k_+^*-k_+)[\cosh(\kappa_\pm d)\right.$$
$$-B_\pm\,\sinh(\kappa_\pm d)]B_\pm(k_+^*+k_+)\sinh(\kappa_\pm x_0)$$
$$+\frac{i}{\kappa_+}[(\kappa_\pm^2+|k_+|^2)[B_\pm\cosh(\kappa_\pm d)-\sinh(\kappa_\pm d)]$$
$$\left.+B_\pm(\kappa_\pm^2-|k_+|^2)\cosh(\kappa_\pm x_0)]\right\}. \tag{9}$$

Here

$$k=\sqrt{2m\mu},\quad \kappa_\pm=\sqrt{2m(V-(\mu\pm E))},$$

$$k_+=\sqrt{2m(\mu+i\zeta)}\simeq k\left(1+\frac{i\zeta}{2\mu}\right),\quad B_\pm=B/2\kappa_\pm.$$

$D_\pm=|a_\pm|^{-2}$ and $R_\pm=|b_\pm/a_\pm|^2$ are, respectively, the transmission and reflection coefficients in the normal tunneling channel for quasiparticles with energies of $\pm E$. It is easy to confirm that the normalization condition $R_\pm+D_\pm=1$ is satisfied for both normal channels. When there are no localized states ($B=0$), Eq. (7) transforms to the previously[12] established dispersion equation for tunneling junctions of finite thickness.

## 4. THE JOSEPHSON CURRENT

We write the superconducting current flowing through the structure as the sum of two terms[14,15]

$$I_s = I_{\text{bound}} + I_{\text{cont}},$$

$$I_{\text{bound}} = 2e \sum_j \frac{\partial E_j}{\partial \varphi} \frac{1}{1+\exp(E_j/T)},$$

$$I_{\text{cont}} = \frac{e}{\pi} \int_0^\infty d|\zeta| \{ |t^N(\varphi, x_0)|^2 - |t^A(\varphi, x_0)|^2 - |t^N(-\varphi, -x_0)|^2 + |t^A(-\varphi, -x_0)|^2 \} \tanh \frac{E}{2T}, \tag{10}$$

which determine the contributions from the bound states $I_{\text{bound}}$ ($|E|<\Delta$) and from the continuum $I_{\text{cont}}$ ($|E|>\Delta$), respectively.

Equations (4)–(10) are valid for arbitrary relationships among the characteristic lengths $(d, \xi_i, \xi_0, \xi_b, \alpha)$ and characteristic energies $(V-\mu, \Delta, \Gamma)$ of the structure. In the following, however, we limit ourselves solely to the practically important case of low barrier transparency, i.e., to the case of a relatively thick barrier:

$$\kappa d \gg 1, \quad \kappa |d/2 \pm x_0| \gg 1, \quad \kappa = \sqrt{2m(V-\mu)} = \alpha^{-1}, \tag{11}$$

when the localized state has well determined energy eigenvalues.[16]

In this limit the expression for the transmission coefficients in the normal tunneling channels, $D_\pm$, reduces to the standard Breit–Wigner formula[16]

$$D_\pm = \frac{\Gamma_0^2}{(E \mp (E_R + \delta E_0^\pm))^2 + \Gamma_0^2 \cosh^2(\kappa_\pm x_0)},$$

$$\Gamma_0 = 2(V_0 - \mu)\sqrt{D_0},$$

$$D_0 = \frac{16 k^2 \kappa^2}{(k^2+\kappa^2)^2} \exp(-2\kappa d), \tag{12}$$

$$E_R = (V-\mu) - \frac{B^2}{8m} + \Gamma_0 \frac{k^2-\kappa^2}{2k\kappa} \cosh(\kappa x_0),$$

$$\delta E_0^\pm = \Gamma_0 \frac{k^2-\kappa^2}{2k\kappa} \exp[(\kappa - \kappa_\pm) d] \cosh(\kappa_\pm x_0), \tag{13}$$

in which $E_R$ is the renormalized energy of the localized state, $\Gamma_0$ is the width of its resonance level located inside the barrier, and $D_0$ is the normal (in the absence of a localized state) transparency of the junction.

According to Eqs. (5)–(10), the Josephson current through the structure is determined by a combination of the $a_\pm$ and $b_\pm$ which describe the coherent processes of normal and Andreev quasiparticle reflection. Thus, besides the conditions (11), which determine the thickness of the junction and the coordinate of the localized state, which has a well defined eigenenergy, and are sufficient for describing single particle tunneling processes, additional conditions arise for the energies of the localized and bound states, and the thickness and height of the barrier which separate the various regimes of coherent Josephson tunneling. Note that in all the cases examined below, the current from the bound states is the main contribution to the resonance Josephson current.

## 5. THE LIMIT OF A SHORT TRANSITION AND A HIGH BARRIER

When the inequalities

$$|E| \ll (V-\mu)\exp(-\kappa d/2), \tag{14}$$

and

$$|E_R| \ll (V-\mu)\exp(-\kappa d/2)\sqrt{\cosh(\kappa x_0)} \tag{15}$$

are satisfied (condition (14) is identical to the condition $d \ll \xi_b$), it is possible to neglect the differences in both the transmission coefficients and the phase shifts of the Bogolyubov–de Gennes functions $u$ and $v$ in the $\pm E$ channels, which makes it much simpler to calculate the superconducting properties of the structure. If we rewrite the dispersion equation $Q=0$ in the form

$$\left(\frac{E}{\Delta}\right)^2 \mathrm{Re}(a_+^* a_-) - \frac{E\xi}{\Delta^2} \mathrm{Im}(a_+^* a_-) - \frac{1}{2}[\mathrm{Re}(a_+^* a_-) + \cos\varphi + \mathrm{Re}(b_+^* b_-)] = 0 \tag{16}$$

and note that when the inequalities (14) and (15) are satisfied, the quantities in the dispersion relation (16) are given by

$$\mathrm{Im}(a_+^* a_-) = -2\frac{E}{\Gamma_0}\cosh(\kappa, x_0),$$

$$\mathrm{Re}(a_+^* a_-) = \frac{E_R^2 - E^2}{\Gamma_0^2} + \cosh^2(\kappa x_0),$$

$$\mathrm{Re}(b_+ b_-^*) = \mathrm{Re}(a_+^* a_-) - 1, \tag{17}$$

we arrive at a dispersion relation obtained previously by Beenakker and van Houten[13] without rigorously determining the conditions for its validity (they only pointed out that $d \ll \xi_0$ holds, while the actual conditions for validity of Eq. (17) are that the inequalities (14) and (15) be satisfied):

$$(\Delta^2 - E^2)(E^2 - E_R^2 - \Gamma^2/4) + \Delta^2 \Gamma_2 \Gamma_1 \sin^2(\varphi/2) + \Gamma E^2 \sqrt{\Delta^2 - E^2} = 0,$$

$$\Gamma = \Gamma_1 + \Gamma_2, \quad \Gamma_1 = \Gamma_0 \exp(-\kappa x_0), \quad \Gamma_2 = \Gamma_0 \exp(+\kappa x_0). \tag{18}$$

The dispersion relation (18) implies that, in this case, the amplitude and shape of the bound Andreev states depend only on the relationship between the magnitude $\Delta$ of the order parameter and the characteristic energy $\Gamma_0$ of the localized state, as well as on its coordinate $x_0$.

In the limit of a wide energy band of the localized state $\Gamma_0 \gg \Delta$, i.e., in the range of spacer thicknesses

$$\kappa^{-1} \ll d \ll \frac{\ln[8\kappa(V-\mu)/k\Delta]}{\kappa}, \tag{19}$$

the dispersion relation (18) and Eq. (10) yield an expression for the energy of the bound Andreev state and the Josephson current through this structure:

$$E(\varphi)_{\Delta\ll\Gamma_0}=\pm\Delta\sqrt{1-D(E_R)\sin^2(\varphi/2)},$$

$$D(E_R)=\{(E_R/\Gamma_0)^2+\cosh^2(\kappa x_0)\}^{-1}, \quad (20)$$

$$I(\varphi)_{\Delta\ll\Gamma_0}=\frac{e\Delta}{2}\,D(E_R)\,\frac{\sin\,\varphi}{\sqrt{1-D(E_R)\sin^2(\varphi/2)}}\tanh\frac{|E(\varphi)|}{2T}.$$

Equation (20) is formally the same as the dependence for a single mode S-c-S transition.[17,18] The difference is that the transparency $D(E_R)$ is the Breit–Wigner resonance transparency (12) for localized states with a different energy $E_R$.

Increases in the junction thickness $d$ are accompanied by a reduction in the width of the resonance band and in the limit $\Gamma_0\ll\Delta$, i.e., for strong failure of the inequality (19), Eqs. (10) and (18) imply that

$$E(\varphi)_{\Delta\gg\Gamma_0}=\pm\frac{\Delta\Gamma_0\sqrt{1-D(E_R)\sin^2(\varphi/2)}}{\sqrt{D(E_R)}\sqrt{E_R^2+(\Delta+\Gamma_0\,\cosh(\kappa x_0))^2}},$$

$$I(\varphi)_{\Delta\gg\Gamma_0}=\frac{e\Delta\Gamma_0}{2}\,\frac{\sqrt{D(E_R)}}{\sqrt{E_R^2+(\Delta+\Gamma_0\,\cosh(\kappa x_0))^2}}$$

$$\times\frac{\sin\,\varphi}{\sqrt{1-D(E_R)\sin^2(\varphi/2)}}\tanh\frac{|E(\varphi)|}{2T}. \quad (21)$$

It follows from Eq. (21) that in the limit $\Gamma_0\ll\Delta$ the maximum Josephson current (for $E_R=0$, $x_0=0$) has a phase dependence proportional to $\sin(\varphi/2)$ and an amplitude that is a factor of $\Delta/\Gamma_0$ smaller than in the limit $\Gamma_0\gg\Delta$. Then the bound Andreev state is compressed toward the chemical potential $\mu$. The rise in $E_R$ ''repels'' the Andreev levels and the gap boundaries and ensures that the phase dependence of the Josephson current is proportional to $\sin\varphi$.

Equations (20) and (21) imply that in both limiting cases $I(\varphi)$ differs from sinusoidal for energies of the localized states such that $E_R\leqslant\Gamma_0$. In particular, for $E_R=0$ and $x_0=0$ the resonance transparency is equal to unity and $I(\varphi)\propto\sin(\varphi/2)$ in both limiting cases. As the difference between the energy of the localized state and the chemical potential $\mu$ increases, the resonance transparency drops sharply, so that $I(\varphi)$ approaches a sinusoidal dependence quite rapidly. This means that the superconducting current averaged over the energy and coordinates of the localized states is proportional to $\sin\varphi$ over almost the entire temperature range.

An averaging procedure carried out in the three-dimensional case in calculations of the superconducting current using the Green's function formalism yielded the same result.[11]

The procedure for averaging over the coordinates and energies of the localized resonance Josephson current described in Eqs. (20) and (21) leads to a dependence of the form

$$\langle J\rangle_{E_R,x_0}\propto(J_c)_{\max}n_{E_R,x_0}\Gamma_0 S/\kappa, \quad (22)$$

where $(J_c)_{\max}$ is the maximum critical resonant tunneling current through an isolated localized state which follows from Eqs. (20) and (21) for $E_R=0$, $x_0=0$, and $\varphi=\pi$; $n_{E_{R,x_0}}$ is the concentration of localized states, $\Gamma_0$ is the size of the optimal energy band near the chemical potential $\mu$, and $S/\kappa$ is the volume of the optimal spacer layer region near $x_0=0$. Note that the suppression parameter for the averaged Josephson current, $\Gamma_{LS}$, introduced in Ref. 11 is directly proportional to the ratio of $\Delta(0)$ to $\Gamma_0$ (for a three-dimensional geometry). Thus, the predicted[11] suppression of the critical junction current as the parameter $\Gamma_{LS}$ increases is a direct consequence of Eqs. (20)–(22): as $\Gamma_{LS}$ increases, there is a reduction in the width of the resonant Andreev level and a corresponding reduction in the critical current which is described by the transition from the limit of Eq. (20) to the limit of Eq. (21).

The averaged resonant conductivity of junctions with normal edges and a single localized state on the trajectory is given by[19]

$$\frac{1}{\langle\rho\rangle}=\frac{e^2}{\pi}\,n_{E_R,x_0}\pi^2\Gamma_0 S/\kappa, \quad (23)$$

where $e^2/\pi$ is the quantum mechanical unit of conductivity. An analysis of resonance quasiparticle tunneling in S-Sm-S structures has shown[20] that superconductivity in the edges does not have an asymptotic effect (for $eV\gg\Delta$) on the conductivity of the structure. At the same time, as Eqs. (20) and (21) imply, the maximum critical current depends on the ratio $\Delta/\Gamma_0$ and equals $e\Delta$ for $\Gamma_0\gg\Delta$ and $e\Gamma_0$ in the opposite limit.

Combining Eqs. (22) and (23), we see that in the limit $\Gamma_0\gg\Delta$ the product $\langle J\rangle_{E_{R,x_0}}\langle\rho_n\rangle_{E_{R,x_0}}$ depends weakly on the spacer parameters (to within a proportionality coefficient which is not written down in Eq. (20)) and is determined mainly by the maximum critical junction current $e\Delta$, while in the opposite limit $\Gamma_0\ll\Delta$ we have

$$\langle J\rangle_{E_R,x_0}\langle\rho_n\rangle_{E_R,x_0}\propto\Gamma_0\propto 1/\langle\rho_n\rangle_{E_R,x_0}.$$

This is a possible explanation for the dependence

$$J_c\rho_n\propto 1/\rho_n$$

(the so-called scaling law) observed in experiments with HTSC junctions.[1–3]

## 6. THE LIMIT OF INTERMEDIATE SPACER THICKNESSES AND BARRIER HEIGHTS

In experiments, especially with HTSC structures, conditions (14) and (15) may be violated, while condition (11) for effective resonant single-particle tunneling continues to be satisfied. Let us first examine the case where the junction length is still small on the length scale of the quasiparticle dephasing[12] for all values of the bound state energies $|E|\leqslant\Delta$, i.e.,

$$\xi_b\ll d\ll\xi_i, \quad (24)$$

while the eigenenergy of the localized state and its coordinate equal zero, i.e., $E_R=0$ and $x_0=0$. In this case an analytic solution of the dispersion equation (7) is possible. Note that condition (14) can fail only in the ''narrow resonance line'' limit, i.e., $\Delta\gg\Gamma_0$, since in the opposite limit $\Delta\ll\Gamma_0$ condition (14) is automatically satisfied.

Thus, in the limit of a narrow resonance line, $\Delta\gg\Gamma_0$, for relatively thick junctions condition (14) may fail at least for

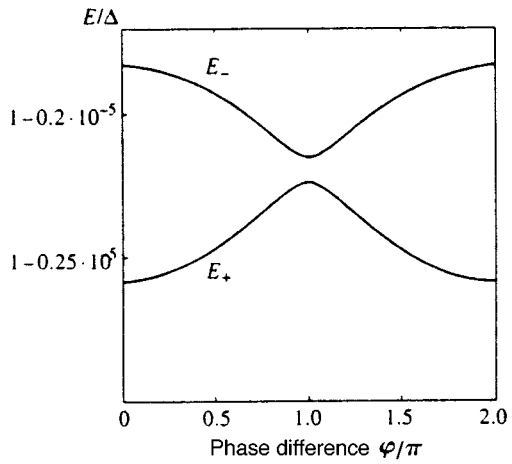



FIG. 2. Split Andreev levels for $E \approx \Delta$, $d/\alpha = 12$, $\Delta/\mu = 2.5 \times 10^{-3}$, and $(V-\mu)/\mu = 0$.

Andreev levels with an energy on the order of $\Delta$. However, for $E_R = 0$ the solution (21) is ''pressed down'' against the chemical potential $\mu$ and has amplitude $\Gamma_0$, so that condition (14) for the solution (21) is satisfied for arbitrary junction thicknesses. At the same time, we note that for a localized state eigenenergy $E_R \geqslant \Gamma_0$, the solution (21) corresponds to Andreev levels with an energy $E$ approaching $\Delta$. Then Eqs. (20) and (21) imply a crossover in the form of the resonance curves of the critical current with $1/E_R$ and $1/E_R^2$ when $E_R \approx \Delta \geqslant \Gamma_0$ (first noted in Ref. 13). Since condition (14) fails for $E_R \approx \Delta$ in the relatively thick junctions which satisfy condition (24), this crossover will not occur in these junctions.

When condition (24) is satisfied for $|E| \approx \Delta$, two new roots of the dispersion relation (7) will appear. These two new solutions are direct analogs of the ''split'' Andreev levels examined in Ref. 12 for the case of a long S-I-S junction. The difference is that the effective reflection coefficient $\tilde{R}$ now has an anomalous sign: $\tilde{R} \simeq -1$. (In the case of a long S-I-S junction $\tilde{R} \cong 1$.) Thus, the angle $\alpha$ is close to $\pi$:

$$\alpha = \pi + \alpha_0, \quad \alpha_0 = 2\sqrt{(\delta/2 + \cos^2(\varphi/2))\tilde{D}} \ll 1,$$

$$\delta \cong ((\kappa_+ - \kappa_-)d)^2/2 \ll 1, \quad \tilde{D} \simeq \{2(V-\mu)/\Delta\}^2 D_0. \quad (25)$$

The expression for the dephasing angle $\beta$ has the same form as in the case of a long S-I-S junction:[12]

$$\beta \simeq 2(\kappa_+ - \kappa_-)/k \ll 1.$$

Then the dispersion relation (7) yields an expression for the split Andreev levels near $\Delta$:

$$E_\pm^2 = \Delta^2\{1 - [\beta/2 \pm \sqrt{\tilde{D}(\delta/2 + \cos^2(\varphi/2))}]^2\}. \quad (26)$$

Equation (26) transforms to the expression for the split Andreev levels of a long S-I-S junction when $\cos(\varphi/2)$ is replaced by $\sin(\varphi/2)$ and $\tilde{D}$ by $D$. Since it follows from the first inequality of Eq. (24) that $\beta/2 > \sqrt{\tilde{D}}$, neither of the solutions (26) reaches the boundary $E = \Delta$ (Fig. 2).

The steepness of the bands (26) in $E(\varphi)$ determines the current flowing through each bound state in accordance with Eq. (10):

$$I_\pm = \text{sign } E \sin\varphi(e\Delta\tilde{D}/2)[1 \pm \beta/\sqrt{4\tilde{D}(\delta/2 + \cos^2(\varphi/2))}]. \quad (27)$$

Both of the currents (27) are substantially higher than the forward current, which does not involve localized tunneling states and is given by the Ambegaokar–Baratoff formula,[21] and have a phase dependence proportional to $\sin(\varphi/2)$. These currents are lower than the current from the Andreev level (21) only by the preexponential factor $(\kappa/k)(\Delta/(V-\mu))$. However, as in the case of the long S-I-S junction, the currents (27) have different signs and cancel each other under equilibrium conditions, yielding an overall amplitude of the same order as the forward tunneling current $4e\Delta(k/\kappa)^2 D_0$, which is exponentially smaller than the current (21). Thus, the appearance of additional solutions to the dispersion equation (7) when conditions (14) and (15) are no longer satisfied does not affect the equilibrium resonance Josephson current through a junction.

## 7. THE LIMIT OF LOW BARRIER HEIGHT AND ARBITRARY JUNCTION THICKNESS

As noted in the previous section, a further reduction in the ratio $\Delta/(V-\mu)$ and (or) an increase in the junction thickness $\kappa d$, which leads to failure of the second inequality in Eq. (24), does not affect the expression for the Andreev level near the chemical potential $\mu$ (21) (for $E_R \approx \Gamma_0$). Numerical analysis of Eq. (7) shows that, as before, two split roots exist near $E \approx \Delta$, but they become flatter and have a phase dependence proportional to $\cos\varphi$. Thus, as before, their presence has no effect on the equilibrium resonance Josephson current of the junction determined (21) by the Andreev level.

## 8. CONCLUSION

The above analysis of resonance Josephson tunneling through a single isolated state in a spacer layer reveals the existence of two additional characteristic lengths in the system, $\xi_i$ and $\xi_b$ (see Eq. (2)) which separate the different regimes for transport of Cooper pairs through the structure. Nevertheless, the interesting feature of resonance Josephson tunneling was that, despite the substantially different dynamics for transport of Cooper pairs through the structure in the different regimes (the appearance of additional bound Andreev states), the overall equilibrium current through the structure is determined only by the ratio of the order parameter of the superconducting electrodes, $\Delta$, to the width of the resonance level, $\Gamma_0$, in all regimes and for arbitrary spacer thicknesses. The results of this paper can be used to establish precisely the limits of applicability of earlier calculations of the resonant Josephson current[11,13] and to explain, in a natural way, the predicted[13] suppression of the resonant Josephson current compared to single-particle tunneling in the limit $\Delta \gg \Gamma_0$, as well as the $J_c\rho_n \propto 1/\rho_n$ dependence observed in experiments with HTSC junctions.[1–3]

This work was supported by the Program on Modern Problems in Solid State Physics and the INTAS-RFBR 95-1305 project.

[1)]E-mail: idev@rsfq.npi.msu.su

[1] U. Kabasawa, Y. Tanitani, T. Fukazawa *et al.*, Jpn. J. Appl. Phys., Part 1 **30**, 1670 (1991).
[2] Y. Kozono, M. Kasai, Y. Kanke *et al.*, Physica C **185–189**, 1919 (1991).
[3] S. Tanutani, T. Fukazawa, O. Kabasawa *et al.*, Appl. Phys. Lett. **58**, 2707 (1991).
[4] Yu. M. Boguslavskii and J. Gao, Physica C **194**, 268 (1992).
[5] J. Gao, Yu. M. Boguslavskii, B. B. G. Klopman *et al.*, J. Appl. Phys. **72**, 575 (1992).
[6] M. Kasai, Y. Kanke, T. Oho *et al.*, J. Appl. Phys. **72**, 5344 (1992).
[7] I. I. Vengrus, M. Yu. Kupriyanov, O. V. Singirev *et al.*, JETP Lett. **60**, 381 (1994).
[8] T. Satoh, M. Hidaka, M. Yu. Kupriyanov *et al.*, IEEE Trans. Appl. Supercond. **5**, 2612 (1995).
[9] M. Yu. Kupriyanov and J. S. Tsai, IEEE Trans. Appl. Supercond. **5**, 2531 (1995).
[10] M. Siegel, R. Dommel, C. Horstmann, and A. I. Braginskii, in *Extended Abstracts of International Conference on Superconductor Electronics*, Nagoya, Japan (1995), p. 141.
[11] I. A. Devyatov and M. Yu. Kupriyanov, JETP Lett. **59**, 200 (1994).
[12] G. Wendin and V. S. Shumeiko, Phys. Rev. B **53**, R6006 (1996).
[13] C. W. J. Beenakker and H. van Houten, in *Single-Electron Tunneling and Mesoscopic Devices*, H. Koch and H. Lubbig, eds., Springer-Verlag, Berlin, Heidelberg (1992), p. 175.
[14] B. J. van Wees, K.-M. H. Lenssen, and C. J. P. M. Harmans, Phys. Rev. B **44**, 470 (1991).
[15] P. F. Bagwell, Phys. Rev. B **46**, 12573 (1992).
[16] H. Knauer, J. Rihter, and P. Siedel, Phys. Stat. Sol. **44**, 303 (1979).
[17] C. W. J. Beenakker and H. van Houten, Phys. Rev. Lett. **66**, 3056 (1991).
[18] A. Furusaki, H. Takayanagi, and M. Tsukuda, Phys. Rev. Lett. **67**, 132 (1991).
[19] A. I. Larkin and K. A. Matveev, Zh. Éksp. Teor. Fiz. **93**, 1030 (1987) [Sov. Phys. JETP **66**, 580 (1987)].
[20] A. V. Tartakovskiĭ and M. V. Fistul', Zh. Éksp. Teor. Fiz. **94**, 353 (1988) [Sov. Phys. JETP **67**, 1935 (1988)].
[21] V. Ambegaokar and A. Baratoff, Phys. Rev. Lett. **10**, 486 (1963).

Translated by D. H. McNeill

# Effect of interparticle interactions on radiative lifetime of photoexcited electron–hole system in GaAs quantum wells

L. V. Kulik, A. I. Tartakovskiĭ, A. V. Larionov, E. S. Borovitskaya, and V. D. Kulakovskiĭ

*Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia*

(Submitted 19 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 353–361 (July 1997)

The paper reports on an investigation of changes in the photoluminescence linewidth and lifetime of excitons and electron–hole plasma over a wide range of densities between $3\times10^7$ and $3\times10^{12}$ cm$^{-2}$ at a temperature of 77 K in GaAs/AlGaAs quantum wells. The roles played by thermal ionization of excitons at low densities of nonequilibrium carriers, exciton–exciton and exciton–electron collisions, and ionization of excitons at high pumping power densities have been studied. © *1997 American Institute of Physics.* [S1063-7761(97)03107-7]

## 1. INTRODUCTION

Radiative recombination of excitons in quasi-two-dimensional semiconducting structures is quite different from the three-dimensional case. The lower dimensionality of the system leads to radical changes in the interaction between excitons and electromagnetic waves. Owing to the translational symmetry of a bulk crystal, this interaction results in formation of stationary excitonic polaritons, which can decay only through phonon scattering or conversion on the crystal surface.[1,2] In the case of excitons in quantum wells, the translational symmetry in the direction perpendicular to the quantum well plane is broken, which results in a very fast (of order 10 ps) decay of excitons with very small in-plane quasimomenta ($k<k_0=n\omega_x/c$). Here $\hbar\omega_x$ is the exciton energy, $n$ is the refraction index, and $c$ is the speed of light. Excitons with $k>k_0$ do not recombine.[2]

The cause of the fast recombination of excitons with $k<k_0$ is the phase coherence of the excitonic states. The loss of coherence due to either localization of excitons, or scattering by phonons, electrons or other quasiparticles leads to a sharp increase in the electron lifetime.[1] Partial ionization of excitons at higher temperatures also leads to an increase in the excitonic system lifetime.[3]

In the present work, we have studied the effect of interparticle interactions in the excitonic system in GaAs/AlGaAs quantum wells on the luminescence linewidth and lifetime over a wide range of densities of nonequilibrium carriers, including the region of the transition from excitonic gas to electron–hole plasma. Experiments have been conducted at a relatively high temperature of 77 K, when excitons and electron–hole plasma coexist in equilibrium and the effect of exciton localization on potential irregularities is negligible. Under these conditions, it is possible to reliably determine both the total density of photoexcited carriers and the system composition, which allows us to analyze on a quantitative level the effect of exciton–electron collisions on the decay of excitonic states and on the radiative annihilation of excitons, and also to study the radiative recombination time in a quasi-two-dimensional system in the region of high densities, where the transition from excitons to electron–hole plasma occurs.

## 2. EXPERIMENTAL TECHNIQUE

We selected for our experiments an undoped GaAs/$Al_{0.18}Ga_{0.82}$As heterostructure grown by the MBE technique and containing a single quantum well of width $L=5$ nm. Excitons were generated by a pulsed picosecond R6G dye laser operating at a wavelength of 590 nm with a 70-ps pulse with a repetition rate of 4 MHz. The sample was placed in a cryostat. Pumping radiation was conducted to the sample and luminescence was fed from the cryostat via an optic fiber with a diameter of 0.6 mm, adjacent to the sample surface (within 0.5 mm). In order to prevent the spread of nonequlibrium carriers from the optically excited area, we used samples with 0.5-mm mesas selectively etched on their surfaces. Luminescence was detected by a photomultiplier tube operating in the time-correlated photon-counting mode. The density $N$ of nonequilibrium carriers in the quantum well at high pumping powers, when a dense $e-h$ plasma was produced, was determined using two methods, namely, by analyzing the luminescence line shape[4] and by deriving it from the pumping power density (under conditions of the experiment, the lifetime of nonequilibrium carriers was always much longer than both the laser pulse width and the width of the time gate during which luminescence was detected). The values of $N$ determined by the two methods agreed within 10%, which indicates that nonequilibrium carriers were effectively contained in the quantum well. At lower pumping power densities, when the excitonic line dominated in the luminescence spectrum, the concentration was derived from the pumping power density under the assumption that the fraction of carriers contained in the quantum well was constant as a function of the pumping power density.

## 3. EXPERIMENTAL RESULTS

Figure 1 shows luminescence spectra of a GaAs/AlGaAs quantum well recorded over a wide range of pumping densities at an ambient temperature of 77 K. For comparison, an excitonic spectrum recorded at 4.2 K is shown by a dashed line in Fig. 1a. This curve demonstrates that the full width at half maximum (FWHM) $\Delta_x$ of the exciton line at liquid helium temperatures is 1.4 meV. This linewidth is due to local-

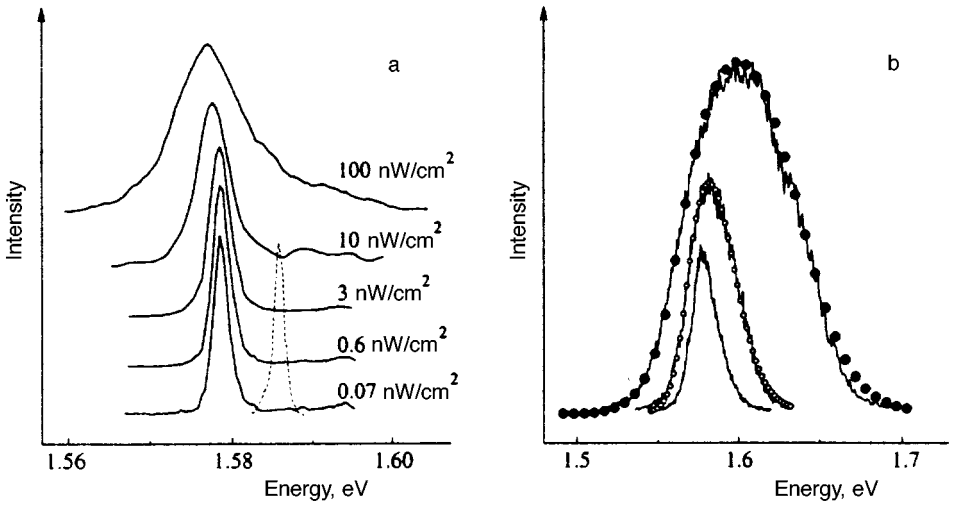



FIG. 1. Luminescence spectra of a GaAs/AlGaAs quantum well at 77 K. For comparison, the dashed line in Fig. 1a plots the exciton luminescence spectrum at 4.2 K.

ization of excitons at potential fluctuations caused by the inhomogeneities in the quantum well width and the content of Al in AlGaAs.[5] Inhomogeneous broadening of the exciton line of about 1 meV is typical of high-quality quantum wells with $L=5$ nm.[5] As the temperature is increased to 77 K at a low pumping power density, the excitonic spectral line broadens to $\Delta_x=2.2$ meV. In the range of low pumping powers, the linewidth is almost constant for photogenerated carrier densities of up to $N\approx 10^9$ cm$^{-2}$. This leads us to a conclusion that the increase in $\Delta_x$ in the temperature range between 4 K and 77 K is mainly caused by the increase in the exciton state damping due to the exciton–phonon scattering, i.e., it is a manifestation of the increase in the homogeneous linewidth of the excitonic line.[1,6]

Figure 1 shows that the exciton line monotonically broadens with the pumping power density $W$ for $W>3$ nW/cm$^2$. This broadening is an indication of additional exciton damping due to collisions among particles, primarily exciton–electron collisions. The exciton binding energy in the quantum well is comparable to $kT$ at 77 K, and the excitonic gas is highly ionized in the range of densities up to fairly high values. Besides, the exciton–exciton interaction (involving two neutral particles) is notably weaker than the exciton–electron interaction.

At pumping power densities $W>100$ nW/cm$^2$, the carrier concentration is higher than the critical value for the Mott transition from the excitonic gas to the $e-h$ plasma. Figure 1b clearly shows that the shape of the recombination line in this case is in a good agreement with calculations based on the plasma approximation,[3] and the density and temperature of the $e-h$ plasma can be derived from the shapes of experimental curves.[7] Our fits of calculations to the experimental line shapes indicate that the temperature of the $e-h$ plasma increases from 90 K at $N=5\times 10^{11}$ cm$^{-2}$ to 150 K at $N=2\times 10^{12}$ cm$^{-2}$. Figure 1 also clearly demonstrates that there are no peaks in the behavior of the linewidth in the density range corresponding to the transition from the excitonic gas to an $e-h$ plasma. This should have been expected, because the excitonic line broadening just below this transition due to interaction between particles is approximately equal to the excitonic Rydberg, which is, in turn, comparable near this transition to the Fermi energy in the $e-h$ plasma.[8]

Figure 2 displays a decay of luminescence from the quantum well, $I(t)$, at various pumping power densities. The luminescence decay time constant $\tau$ falls monotonically with the density of carriers in the quantum well. In a general case, when the radiation lifetime $\tau_r$ of the excitonic system recombination is density-dependent, the time constant $\tau$ is related to $\tau_r$ by the equation

$$\tau=\eta\tau_r/(1+\eta d\tau_r/dt), \tag{1}$$

where $\eta=\tau_{nr}/(\tau_r+\tau_{nr})$ is the luminescence quantum efficiency and $\tau_{nr}$ is the nonradiation lifetime. Equation (1) is derived from the relations

$$dI/dt=-I/t, \tag{2}$$

$$I=N/\tau_r, \tag{3}$$

$$dN/dt=-N(1/\tau_r+1/\tau_{nr}). \tag{4}$$

Measurements of the quantum efficiency as a function of electron–hole pair density in the quantum well at 77 K are plotted in Fig. 3. In the range of high densities ($N=10^{10}$–$10^{11}$ cm$^{-2}$) the quantum efficiency is constant. Measurements taken at lower temperatures indicate that $\eta$ is

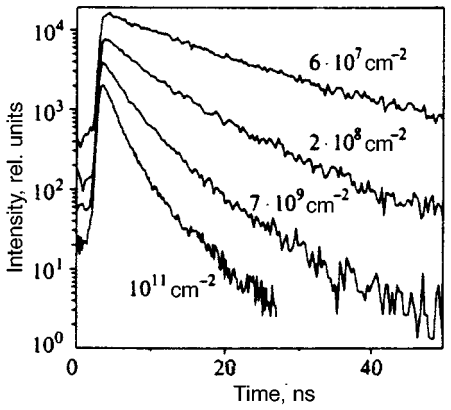



FIG. 2. Luminescence decay curves for the quantum well at different densities.

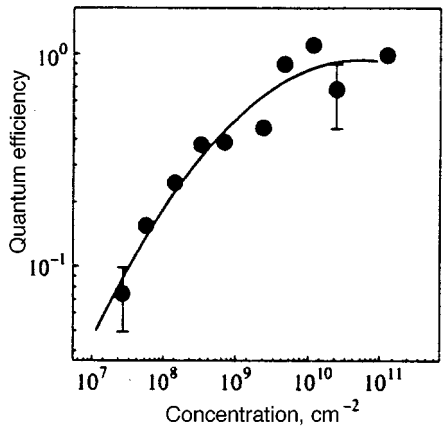



FIG. 3. Quantum efficiency as a function of $e-h$ pair density in the quantum well at 77 K.

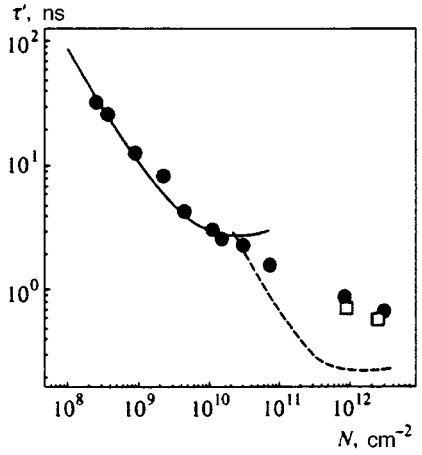



FIG. 5. Lifetime $\tau'$ as a function of $e-h$ pair concentration. Measurements are plotted by full circles, calculations by Eqs. (1), (9)–(12) by the solid line. The dashed line shows the calculated $e-h$ plasma lifetime at 77 K. Open squares show calculations at $N=10^{12}$ cm$^{-2}$, $T=150$ K and $N=3\times10^{12}$ cm$^{-3}$, $T=280$ K.

also constant in this density range as the temperature drops down to 2 K, although the lifetime in this range drops more than threefold. Therefore, in this range of densities we assume $\eta\sim1$. Figures 2 and 3 show that $\tau$ increases and the quantum efficiency $\eta$ drops as the density falls below $10^{10}$ cm$^{-2}$, so the radiation lifetime becomes comparable to $\tau_{nr}$ in this density range. In the range $N<10^8$ cm$^{-2}$, the quantum efficiency $\eta$ is lower than 0.1, i.e., the nonradiational channel of recombination dominates.

#### 4. DISCUSSION

The exciton line FWHM as a function of the density of $e$-$h$ pairs is given in Fig. 4. In the range $N<10^{11}$ cm$^{-2}$ the function $\Gamma(N)$ is linear. It can be approximated using the expression[9]

$$\Delta_x(N_e)=\Delta_{x0}+\Gamma_e(N_e), \tag{5}$$

where

$$\Gamma_e(N_e)=g_e\pi R a_x^2 N_e,$$

$R$ and $a_x$ are the exciton binding energy and Bohr radius, respectively, $N_e$ is the number of decoupled $e-h$ pairs, and $g_e$ is the constant of electron–exciton interaction. The density of free electrons can be determined from the equality between chemical potentials of excitons and free electrons and holes in equilibrium:

$$\mu_x=\mu_e+\mu_h, \tag{6}$$

where $\mu_{x,e,h}$ are the chemical potentials of excitons, electrons, and holes, respectively.

It follows from the approximation that $g_e=9.5$. This value is in agreement with earlier estimates of $g_e$ based on the four-wave mixing measurements:[9] $g_e=10.2$. The effect of exciton–electron collisions on the width of the luminescence line has been quantitatively estimated using the formalism developed by Feng and Spector.[8] The homogeneous linewidth is[8]

$$\Gamma_e=\frac{4\hbar^2}{\pi M}\int_0^\infty dk k^2 Q f\left(\frac{2m_e+m_h}{m_e+m_h}k\right),$$

where $M=m_e(m_e+m_h)/(2m_e+m_h)$, $m_e(m_h)$ is the electron (hole) effective mass, and $f$ is the Fermi distribution function. The scattering cross section in this case is expressed as

$$Q=4\pi\left(\frac{M}{m_e}\right)^2 k^{-1}\int_0^\pi d\theta K^{-2}\left(\left[1+\left(\frac{m_e K}{4m_h}\right)^2\right]^{-3/2}-\left[1+\left(\frac{K}{4}\right)^2\right]^{-3/2}\right)^2,$$

where $K=2ka_x\sin(\theta/2)$ and $\theta$ is the scattering angle. As a result, we have derived the constant of the exciton–electron interaction $g_e=12.4$, which is in a fair agreement with the experimental data.

Figure 5 shows the function $\tau'(N)=\eta\tau(N)$ obtained by processing the data plotted in Figs. 2 and 3. It follows from Eq. (2) that $\tau'(N)$ equals the radiation lifetime divided by the factor $1+\eta d\tau_r/dt$, which is, as will be shown below, approximately equal to 2. Figure 5 clearly shows that in the range $N=3\times10^7-10^9$ cm$^{-2}$ the time $\tau'(N)$ decreases by more than one order of magnitude in inverse proportion to

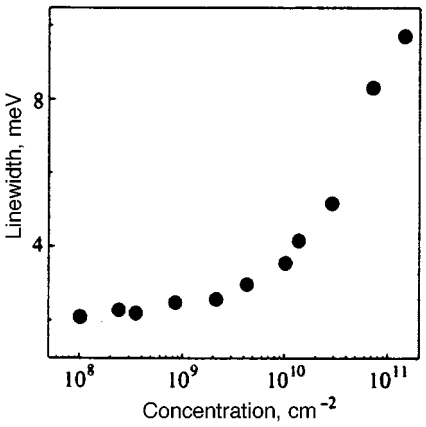



FIG. 4. FWHM of the exciton line versus density of $e-h$ pairs.

the density. At higher excitation densities this dependence is weaker. The increase in the radiation lifetime at low $N$ is caused by ionization of excitons.

The fraction of nonionized excitons in the photoexcited system is

$$\alpha=\frac{N_x}{N}=1-\frac{K}{2N\sqrt{(1+4N/K)}}, \tag{7}$$

where

$$K=\frac{N_eN_h}{N_x}=\frac{m_em_hkT}{\pi\hbar^2m_x}\exp\left(-\frac{R}{kT}\right), \tag{8}$$

and $m_x$ is the exciton mass. Since the exciton Rydberg in the quantum well (11.5 meV) is comparable to $kT$ at 77 K, the ionization degree of excitons at low densities is high and $\alpha \propto 1/N$ (Fig. 5). Luminescence due to free electrons and holes is negligible, so the radiation lifetime of the two-dimensional system can be expressed by

$$\tau_r(N)=\tau_x(T)/\alpha(N). \tag{9}$$

Here $\tau_x(T)$ is the exciton radiation lifetime. It is related to the reduction lifetime $\tau_0$ for excitons with $k<k_0$ by

$$\tau_x(T)=3m_xkT\tau_0/\hbar^2k_0^2. \tag{10}$$

In the range of densities $N>10^{10}$ cm$^{-2}$, when the exciton line FWHM increases (Fig. 4), the effect of collisions among particles on $\tau_x$ should be also taken into account. In a general case, the time $\tau_x$ is related to the homogeneous broadening $\Gamma_h$ of the luminescence line as follows:[2,6]

$$\tau_x\propto\frac{\hbar\Gamma_h}{1-\exp(-\hbar\Gamma_h/kT)}\,\tau_0. \tag{11}$$

At low carrier concentrations $\hbar\Gamma_h\ll kT$, the lifetime satisfies $\tau_x \propto kT$ and is independent of $N$. However, the contribution of exciton–electron collisions becomes important at higher densities, when the exciton damping increases and becomes comparable to $kT$. In particular, estimates based on Eq. (11) yield an increase in $\tau_x$ by nearly half at $N=3\times10^{11}$ cm$^{-2}$ when $\Gamma$ increases to 8 MeV as derived from the luminescence line FWHM.

In addition, note that at high densities the exciton wave function is modified by interparticle interaction, which leads to a change in the exciton oscillator strength $f_x$, and therefore in $\tau_0$ (since $\tau_0 \propto 1/f_x$, see Ref. 10). In fact, an electron or hole, either free or bound in an exciton, can be scattered only to an unoccupied cell of the $k$-space. Therefore the scattering rate of carriers decreases at high density. The modification of the exciton wave function results in a smaller oscillator strength since it satisfies $f_x \propto |\Psi(r=0)|^2$, hence longer $\tau_0$ and $\tau_x$.

The contribution of interparticle interaction to the oscillator strength can be taken into account in perturbation theory with the Coulomb interaction treated as a perturbation. A similar technique was used by Schmitt–Rink *et al.*,[10] who studied the many-body effects on excitonic absorption. To first order, the oscillator strength is given by

$$f_x\propto|\Psi_0(r=0)|^2\left[1-\pi a_x^2N_x-(\pi^2-\pi)a_x^2N_e+\frac{\pi\ln 2}{2\sqrt{3}}(k_0a_x)^2\frac{\mu}{m_x}\right], \tag{12}$$

where

$$k_0=\frac{4\pi N_ee^2}{\epsilon_0kT},$$

$\mu$ is the exciton reduced mass, and $\epsilon_0$ is the material dielectric permittivity. The ratio between $N_x$ and $N_e$ is determined by Eq. (7). This result has been obtained in the limit $k_0a_x\ll1$ and $R>kT$, i.e., Eq. (12) is valid for $N_e<10^{11}$ cm$^{-2}$. The second condition corresponds to the case $R=11.5$ meV and $kT=6.6$ meV.

The first two terms in Eq. (12) are due to the exchange interaction between excitons, and between an exciton and free carriers, respectively. The third term is the contribution of recombination of a free electron (hole) with a hole (electron) bound in an exciton. This term is positive, i.e., it decreases the exciton lifetime, and its contribution is significant at $N\sim10^{10}$ cm$^{-2}$.

The solid trace in Fig. 5 corresponds to the function $\tau'(N)$ calculated by Eqs. (1), (9)–(12) with due account of exciton ionization, exciton–electron collisions, and changes in $f_x$. The only adjustable parameter is the radiation time $\tau_0$. It was selected to fit the calculations to experimental data in the range of low densities, where interparticle interaction can be neglected. The value $\tau_0=24$ ps derived from this fitting is in a fair agreement with the calculations ($\tau_0=16$ ps for a AlGaAs/GaAs quantum well with $L=50$ Å) based on Ref. 2.

The calculation of $\tau'(N)$ shown in Fig. 5 by the solid trace is in a good agreement with measurements for carrier densities of up to $N\approx7\times10^{10}$ cm$^{-2}$. At higher densities, the calculations of $\tau'(N)$ are notably different from measurements, which indicates that in this region excitonic correlations are no longer dominant and the lifetime of $e-h$ pairs should be calculated in the plasma approximation.

The lifetime of dissociated $e-h$ pairs in the quantum well can be calculated similarly to that of excitons.[2] The inverse lifetime of a pair can be expressed as follows:

$$\tau_{eh}^{-1}=\frac{1}{N}\frac{2\pi}{m_0c_0\sqrt{\epsilon_0}}e^2mE_{eh}\sum\mu_{vc}\left|\int\psi_c\psi_v\,dV\right|^2\times\frac{1}{\hbar^2}\frac{1}{2\pi^2}\int n_kn_{k'}\delta(\mathbf{k}-\mathbf{k}')\,d^2k\,d^2k', \tag{13}$$

where $n_k$ is the fermion distribution function, and $\psi_{c(v)}$ is the envelope of the wave function in the conductance (valence) band. It follows from Eq. (13) that the usually accepted density dependence of the $e-h$ pair lifetime $\tau \propto n^{-2}$ is valid only in the case of nondegenerate fermion distribution, whereas in the limit $N\pi\hbar^2/m\gg kT$ of degenerate fermion statistics, the lifetime of $e-h$ pairs becomes independent of the density and approaches a constant which equals 0.24 ns for the parameters of the quantum well in question. This time is about an order of magnitude longer than the lifetime of

excitons for $k<k_0$, but considerably shorter than the exciton lifetime at 77 K: $\tau_x=0.8$ ns. The measured and calculated radiative lifetimes in the region of high densities are compared in Fig. 5, where the calculated dependence is indicated by a dashed line. Even at very high densities, $N>10^{12}$ cm$^{-2}$, the lifetime of the $e-h$ plasma calculated for $T=77$ K is notably shorter than measured. This is not surprising because, as was noted above, the temperature of the photoexcited $e-h$ plasma in a quantum well is higher than 200 K at such densities. At so high temperatures, holes are nondegenerate, which increases the radiative lifetime of the $e-h$ pairs. The two points indicated by open squares in Fig. 5 correspond to calculations for $N=10^{12}$ cm$^{-2}$ at $T=150$ K and for $N=3\times10^{12}$ cm$^{-2}$ at $T=280$ K, i.e., at temperatures derived from luminescence spectra of the $e-h$ plasma. These calculations are in good agreement with experimental values.

It is obvious that the lower boundary of the region where the plasma approximation can be used in calculating $\tau'$ is determined by the emergence of exciton-like correlations in the plasma. Excitonic correlations lead to a larger overlap between the electron and hole wave functions, hence to a shorter lifetime. Figure 5 indicates that this transition occurs near the density $N\propto10^{11}$ cm$^{-2}$, which corresponds to the nondimensional parameter $r_s=1/\pi a_x^2 N\approx3$, in agreement with the expected value for the exciton-to-plasma transition in a dense electron–hole system.

## 5. CONCLUSION

Under conditions of equilibrium between excitons and electron–hole plasma, we have studied the effect of interparticle interaction in a GaAs/AlGaAs quantum well on the width of the luminescence line and carrier lifetime. At low excitation densities, the role of thermal ionization of excitons has been considered. At higher excitation levels, including those corresponding to the region of the transition from the exciton gas to electron-hole plasma, the effects of exciton–electron and exciton–exciton collisions and ionization of excitons have been analyzed.

The authors acknowledge helpful discussions with Prof. A. Forchel (Würzburg University, Germany). The work was supported by INTAS (Grant No. 94-2112) and Russian Program *Physics of Solid-State Nanostructures*.

[1] E. Hanamura, Phys. Rev. B **38**, 1228 (1988).
[2] D. S. Citrin, Phys. Rev. B **47**, 3832 (1993).
[3] V. Srinivas, J. Hryniewicz, Y. J. Chen *et al.*, Phys. Rev. B **46**, 10193 (1992).
[4] P. T. Landsberg, Solid-State Electron. **10**, 513 (1967).
[5] L. C. Andreane, F. Tassoni, and F. Bassani, Solid State Commun. **77**, 641 (1991).
[6] J. Feldmann, G. Peter, E. O. Göbel *et al.*, Phys. Rev. Lett. **59**, 1337 (1987).
[7] B. Deveaud, F. Clerot, N. Roy *et al.*, Phys. Rev. Lett. **67**, 2355 (1991).
[8] Y.-P. Feng and H. N. Spector, J. Phys. Chem. Solids **48**, 593 (1987).
[9] A. Honold, L. Schultheis, J. Kuhl *et al.*, Phys. Rev. B **40**, 6442 (1989).
[10] S. Schmitt-Rink, D. S. Chemla, and D. A. Miller, Phys. Rev. B **32**, 6601 (1985).

Translation was provided by the Russian Editorial office.

# Restoration of selection rules in nonadiabatic resonant inelastic x-ray scattering

F. Gel'mukhanov and T. Privalov

*Institute of Automation and Electrometry, Siberian Branch, Russian Academy of Sciences, 630090 Novosibirsk, Russia*

H. Ågren

*Institute of Physics and Measurement Technology, Linköping University, S-58183 Linköping, Sweden*

(Submitted 13 September 1996)
Zh. Éksp. Teor. Fiz. **112**, 37–49 (July 1997)

Recently a new effect in the Raman scattering of x-ray radiation has been predicted theoretically and discovered in experiments, the effect of restoration of the selection rules for the scattering tensor under strong electron–vibrational interaction. We propose a fairly simple model for describing this effect, a model that allows for an exact solution and takes into account the real vibrational structure of the molecule and electron–vibrational interaction. © *1997 American Institute of Physics.* [S1063-7761(97)00407-1]

## 1. INTRODUCTION

Systematic experimental studies of resonant inelastic x-ray scattering (RIXS), or resonant Raman x-ray scattering, began about a decade ago. This work became possible primarily because of a new generation of high-intensity sources of polarized synchrotron radiation. The high power of these sources radiation makes it possible for a monochromator to ''cut out'' a fairly strong narrow line from the smooth broadband spectrum. Thus it became possible to tune x-ray radiation in frequency and monitor the degree of polarization.

The physics of RIXS can be explained in following manner. An initial x-ray photon excites the target (atom, molecule, or solid). There are two channels, radiative and nonradiative, into which this highly excited intermediate state can decay emitting a spontaneous photon or an Auger electron, respectively. There can in turn be radiative RIXS[1,2] and nonradiative RIXS (or the Auger resonance Raman effect),[3,4] depending on what final particle (an x-ray photon or an Auger electron) is registered by the spectrometer.

The cross section of the Auger resonance Raman effect in the soft x-ray range is considerably larger than the radiative RIXS cross section, since the Coulomb interaction, responsible for the Auger decay, is stronger than the electromagnetic interaction. This is the reason why at present radiative RIXS has lower spectral resolution than nonradiative RIXS. On the other hand, radiative RIXS spectra are simpler and hence more informative. The reason is the dipole nature of the interaction between radiation and target, in accordance with the selection rules for the scattering tensor.[5–12] As shown in Refs. 5–12, the selection rules for the scattering tensor cause the spectral shape of the radiative RIXS cross section to depend strongly on the initial-photon frequency $\omega$. The case where $\omega$ is below the threshold $I$ of ionization of a core electron was considered in Refs. 5–9, and the theory of RIXS for $\omega > I$ was developed in Refs. 10 and 12. Theoretical[5,6,10–13] and experimental[7–9] studies of radiative RIXS by molecules[5,6,10] and solids[11,13] demonstrate the effectiveness of these selection rules in determining the symmetry of occupied and vacant electronic states. In the above papers RIXS was studied for adiabatic electron–vibrational transitions (in the Born–Oppenheimer approximation). It is well known, however, that the adiabatic approximation fails in the presence of two or several closely lying electronic states. In this case the selection rules for the scattering tensor may be violated, since according to the Jahn–Teller theorem the symmetry of the electron subsystem is lowered because of strong electron–vibrational interaction, which mixes the closely lying electronic states of different symmetry. Obviously, in the case of RIXS by symmetric molecules or solids, this effect is more the rule than the exception. The only exception is homonuclear diatomic molecules, in which there is only a symmetric vibrational mode, which does not mix electronic states of opposite parity. Experimental data on RIXS by $O_2$ ($\omega < I$) molecules (Ref. 8) and $N_2$ ($\omega > I$) molecules (Ref. 9) do indeed demonstrate that the selection rules hold rigorously for the scattering tensor. Violation of the selection rules for the x-rayscattering tensor was first discovered in experiments with the $C_{60}$ molecule (Ref. 7), the benzene molecule (Ref. 14) and the $CO_2$ molecule (Ref. 15).

A general theory of RIXS that allows for electron–vibrational interaction and an arbitrary spectral distribution of the excitation radiation has recently been developed in Refs. 15 and 16. The case of broadband excitation was considered earlier.[17] An important result was obtained in Refs. 15 and 16 by theoretical means and was confirmed in experiments. The effect consists in the restoration of the selection rules for the scattering tensor when the detuning $\Omega$ of the frequency $\omega$ of the excitation narrowband radiation from the electron–vibrational absorption band corresponds to a certain nonadiabatic electron transition. Note that the selection rules are again violated for extremely large detunings.[5,16] Before formulating the goal of the present research we list the main theoretical results of Refs. 15 and 16. The effect of restoration of selection rules is described in these papers at three levels of rigor.

(1) A rigorous theory with a corresponding nonempirical calculation of the $CO_2$ molecule allowing for the nonadiabatic interrelationship of the electron and nuclear subsystems. The positive side of this approach, the rigor, is at the same time a drawback in view of extreme complexity.

 1063-7761/97/070020-07$10.00  

(2) A rigorous proof of the fact that the ratio of the intensities of the forbidden and allowed transitions tends to zero for large detunings $\Omega$. Since this asymptotic result is valid only for large detunings, the corresponding description allows for no conclusions about the shape of the spectral RIXS band in the intermediate range of values of $\Omega$.

(3) A qualitative description of the spectral dependence of the relative intensity of the forbidden transition via a five-level model. Here the vibrational structure of the molecule is taken into account by two parameters: the electron–vibrational interaction parameter $\lambda$, and the effective width of the corresponding electron–vibrational band in the absorption spectrum. A drawback is the poor accuracy in allowing for the vibrational structure of the molecule.

Thus, the main goal of our research is to select a fairly simple model that takes into account the real vibrational structure of the molecule and at the same time allows for a rigorous solution of the problem of symmetry restoration in radiative RIXS. Since here we study only radiative RIXS, for the sake of brevity we use only the abbreviation RIXS.

## 2. RIXS AND ELECTRON–VIBRATIONAL INTERACTION

### 2.1. Differential and integral RIXS cross sections

The differential cross section of RIXS into the solid angle $dO$,

$$\frac{d^2\sigma(\omega',\omega)}{d\omega' dO} = r_0^2 \omega \omega'^3 \sum_f |F_f|^2 \delta(\omega' + \omega_{f_0} - \omega), \qquad (1)$$

is expressed in terms of the classical electron radius $r_0 = \alpha^2 \simeq 2.82 \times 10^{-13}$ cm ($\alpha = 1/137$) and the scattering amplitude given by the well-known Kramers–Heisenberg formula[18]

$$F_f = \sum_i \frac{\mathscr{D}'_{fi} \mathscr{D}_{i0}}{\omega - \omega_{i0} + i\Gamma}, \quad \mathscr{D} = \mathbf{e}\cdot\mathbf{d}, \quad \mathscr{D}' = \mathbf{e}'\cdot\mathbf{d}. \qquad (2)$$

Here it is assumed that $kR \ll 1$, where $R$ is the size of the molecule. This condition is met for soft x-ray radiation and not very long molecules. The case $kR \gtrsim 1$ was studied in Refs. 5, 10, and 11. In the present paper we use the atomic system of units ($\hbar = m = e = 1$); $\omega, \mathbf{e}, \mathbf{k}$ and $\omega'$, $\mathbf{e}'$, $\mathbf{k}'$ are the frequencies, polarization vectors, and wave vectors of the initial and final photons, respectively; $\mathbf{d}$ is the operator of the dipole moment of the molecule; and $\omega_{ij} = \mathscr{E}_i - \mathscr{E}_j$ is the frequency of the resonant transition between the molecular states $i$ and $j$. The half-width $\Gamma$ at the half-maximum of the absorption line is assumed constant for simplicity. In Eq. (1) we ignored the broadening $\Gamma_f$ of the final state $|f\rangle$, since it is small compared to $\Gamma$. The scattering amplitude (2) describes the absorption of the initial photon followed by the transition of the molecule from the ground state $|0\rangle$ to the x-ray excited intermediate state $|i\rangle$ followed by spontaneous decay to the final state $|f\rangle$. The delta function in Eq. (1) reflects the conservation of energy in the scattering process and describes the Stokes shift of the emission line,

$$\omega' = \omega - \omega_{f0}. \qquad (3)$$

The presence of two spectral peaks given by (2) and (3) in the cross section (1) is the reason for Stokes doubling of the RIXS line.[5,19,20] Recently the existence of this effect has been confirmed in the experiment of Aksela *et al.*[4]

The differential cross section (1) describes the scattering of a monochromatic beam of x-ray photons by a molecule. In practice, the spectral distribution $\Phi(\omega_1 - \omega, \gamma)$ centered at $\omega$ has a finite width $\gamma$. Thus, if we wish to describe the experiment correctly, we must use the convolution

$$\sigma(\omega',\omega) = \int d\omega_1 \frac{d^2\sigma(\omega',\omega_1)}{d\omega' dO} \Phi(\omega_1 - \omega, \gamma) \qquad (4)$$

of the RIXS cross section (1) and the spectral function $\Phi(\omega_1 - \omega, \gamma)$ normalized to unity. Note that the finite width $\gamma$ of the spectral function violates the linear dependence on $\omega$ of the RIXS line given by Eq. (3).[20,21]

Below we will also need the integral cross section (the area of the corresponding electron–vibrational RIXS band)

$$\sigma(\omega) = \int d\omega' \sigma(\omega',\omega) = \int d\omega_1 \sigma_0(\omega_1) \Phi(\omega_1 - \omega, \gamma), \qquad (5)$$

which is expressed in terms of the integral cross section

$$\sigma_0(\omega) = \int d\omega' \frac{d^2\sigma(\omega',\omega)}{d\omega' dO} = r_0^2 \omega \overline{\omega'}^3 \sum_f |F_f|^2 \qquad (6)$$

in the event of monochromatic excitation ($\gamma = 0$). Here with a high accuracy we can consider $\overline{\omega'}$ as the center of gravity of the corresponding emission line.

### 2.2. The electron–vibrational interaction model

To determine the amplitude (2) and the integral cross sections (5) and (6) we choose a reasonable model reflecting the main features of the problem. The simplest molecule in which there is an antisymmetric vibration connecting the electronic states of opposite parity is a linear triatomic molecule $XY_2$ (e.g., $CO_2$). We assume that the initial photon excites the $K$-electrons of the Y atoms to an unoccupied molecular orbital (or a quasistationary state above the ionization threshold of the $K$-level) of definite parity ($u$ or $g$). We also assume that the molecule in the $K$-excited state remains linear. In accordance with the symmetry of the molecule, the intermediate delocalized $K$-hole electronic state may be either even-parity ($\Psi_g$) or odd-parity ($\Psi_u$). These states are quasidegenerate, i.e., their electron energies are practically the same,

$$E \equiv E_g \simeq E_u, \qquad (7)$$

since the wave functions of the $1s$-electrons belonging to different Y atoms overlap only weakly. For instance, for the $CO_2$ molecule, $|E_g - E_u| \simeq 0.005$ eV. This high degeneracy

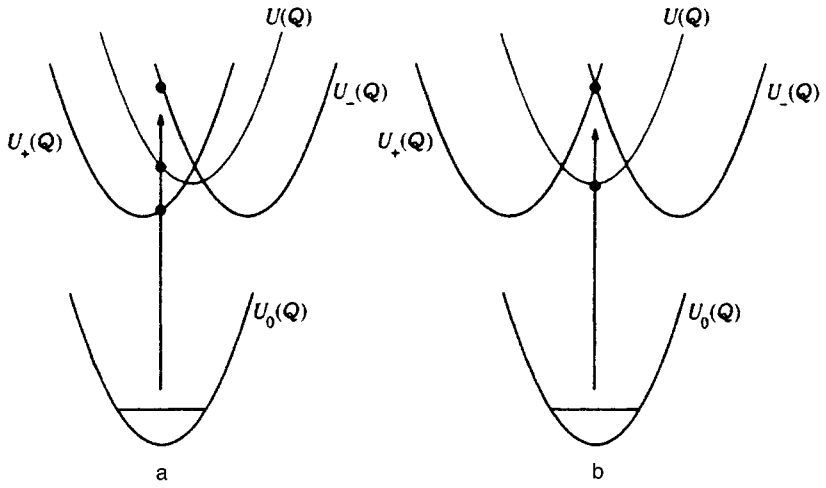



FIG. 1. The shift of the potential curves $U_{\pm}(Q)$ of the states (13) for the antisymmetric mode, caused by the electron–vibrational interaction. Here $Q \equiv Q_u$. The nonadiabatic potentials $U_{\pm}(Q)$ for the antisymmetric mode of the $K$-excited state is determined by the expression for $H_{\pm}$ in Eq. (15) without the contribution of the harmonic potential of the symmetric mode. The light curve depicts the potential curve $U(Q)$ of the adiabatic $K$-excited state. (a) The adiabatic potential curves of the ground ($U_0(Q)$) and $K$-excited ($U(Q)$) states have different equilibrium distances. (b) The ground and $K$-excited states have the same vibrational frequencies and the same equilibrium distances.

of the x-ray excited states is the main reason for strong electron–vibrational interaction. In the two-level approximation, the total Hamiltonian of the molecule,

$$H=\begin{pmatrix} H_0 & V \\ V & H_0 \end{pmatrix}, \tag{8}$$

in the intermediate ($K$-excited) state is the sum of the adiabatic Hamiltonian

$$H_0 = E + h \tag{9}$$

and the operator $V$ responsible for the nonadiabatic coupling of the degenerate $K$-hole electronic states $\Psi_g$ and $\Psi_u$. The Hamiltonian

$$h=\sum_{j=g,u}\left(\frac{P_j^2}{2M_j}+\frac{1}{2}M_j\omega_j^2Q_j^2\right), \tag{10}$$

which describes nuclear motion in the harmonic approximation, allows for two normal modes, the symmetric mode with frequency $\omega_g$ and normal coordinate $Q_g$, and the antisymmetric mode with frequency $\omega_u$ and normal coordinate $Q_u$. For simplicity we ignore the contribution of bending vibrations in Eq. (10). The coefficients $M_g$ and $M_u$ have the dimensions of mass and for the $XY_2$ molecules can be expressed in terms of the masses $M_X$ and $M_Y$ of the atoms X and Y as follows:

$$M_g=\frac{M_Y}{2}, \quad M_u=\frac{M_Y(2M_Y+M_X)}{2M_X}.$$

The orthogonal transformation[17]

$$U=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \tag{11}$$

which diagonalizes the Hamiltonian (1) ($\tilde{H}=U^+HU$)

$$\tilde{H}=\begin{pmatrix} H_+ & 0 \\ 0 & H_- \end{pmatrix}, \quad H_{\pm}=E+h\pm V, \tag{12}$$

lifts the degeneracy (7) of the ''delocalized'' states $\Psi_g$ and $\Psi_u$ and establishes a transition to the localized electron base functions

$$\Psi_1=\frac{1}{\sqrt{2}}(\Psi_g-\Psi_u), \quad \Psi_2=\frac{1}{\sqrt{2}}(\Psi_g+\Psi_u), \tag{13}$$

where $\Psi_1$ is the eigenfunction of the Hamiltonian $H_+$, and $\Psi_2$ is the eigenfunction of the Hamiltonian $H_-$.

According to the Jahn–Teller theorem, the results (12) and (13) have a profound physical meaning: the interaction $V$ lifts the degeneracy of the initial adiabatic states $\Psi_g$ and $\Psi_u$ by lowering the symmetry of the system. Since the initial states $\Psi_g$ and $\Psi_u$ are of opposite parity, they can be coupled only by an interaction $V$ with an antisymmetric vibrational mode. Thus, we can write $V$ in the linear approximation in $Q_u$ as follows:[22,17]

$$V=\lambda Q_u. \tag{14}$$

Plugging this into the expression (12) for $H_{\pm}$, we see that $H_{\pm}$ is again reduced to a sum of two harmonic Hamiltonian:

$$H_{\pm}=E+\frac{P_g^2}{2M_g}+\frac{1}{2}M_g\omega_g^2Q_g^2+\frac{P_u^2}{2M_u}+\frac{1}{2}M_u\omega_u^2Q_u^2,$$

$$Q_u^{\pm}=Q_u\pm\frac{\lambda}{M_u\omega_u^2}. \tag{15}$$

In the expression for $H_{\pm}$ we ignored the term $-\lambda^2/(2M_u\omega_u^2)$, which is quadratic in $\lambda$, since allowing for this term takes us beyond the linear approximation (14). The nonadiabatic approximation (14) shifts the equilibrium distance for the antisymmetric mode by $\mp\lambda/M_u\omega_u^2$ for the Hamiltonians $H_{\pm}$, respectively (Fig. 1).

### 3. PARTIAL AMPLITUDES AND THE SCATTERING CROSS SECTIONS IN THE SYMMETRY-FORBIDDEN AND SYMMETRY-ALLOWED SCATTERING CHANNELS

For definiteness we assume that the ground electronic state is an even-parity one ($\Psi_g^0$). In this case, in the absence of electron–vibrational interaction, the final electronic state $\Psi_f$ must also be an even-parity one ($\Psi_{f_g}$), since a transition to an odd-parity final state ($\Psi_{f_u}$) is symmetry-forbidden.[5] The nonadiabatic interaction (14) ''opens'' the forbidden scattering channel into the final state $\Psi_{f_u}$. Thus, in addition

to the amplitude $F_{f_g}$ and the integral cross section $\sigma_{0g}(\omega)$ of the allowed scattering channel, there is a finite amplitude $F_{f_u}$ and a finite integral cross section $\sigma_{0u}(\omega)$ for the forbidden scattering channel. Then the differential and integral scattering cross sections (Eqs. (1) and (6)) are given by sums of two terms:

$$\frac{d^2\sigma(\omega',\omega)}{d\omega' dO}=\left(\frac{\omega'}{\overline{\omega'}}\right)^3\sum\ [\sigma_{0f_g}(\omega)\,\delta(\omega'+\omega_{f_g0}-\omega)$$
$$+\sigma_{0f_u}(\omega)\,\delta(\omega'+\omega_{f_u0}-\omega)],$$
$$\sigma_0(\omega)=\sigma_{0g}(\omega)+\sigma_{0u}(\omega), \tag{16}$$

where

$$\sigma_{0\alpha}(\omega)=\sum\ \sigma_{0f_\alpha}(\omega),\quad \sigma_{0f_\alpha}(\omega)=r_0^2\omega\overline{\omega^3}|F_{f_\alpha}|^2,$$
$$\alpha=g,u,$$
$$\omega_{f_\alpha 0}=E_{f_\alpha}-E_0+\omega_{f_g}\left(m_{f_g}+\frac{1}{2}\right)+\omega_{f_u}\left(m_{f_u}+\frac{1}{2}\right)$$
$$-\frac{1}{2}(\omega_{0g}+\omega_{0u}). \tag{17}$$

Here $\omega_{0\alpha}$ and $\omega_{f_\alpha}$ are the vibrational frequencies of the ground and final electronic states, respectively, and the index $\alpha=g$, $u$ denotes the parity of the electronic state. The sums in Eqs. (6) and (17) mean summation over the vibrational levels ($m_{f_u}=0,1,2,\ldots$) of the final state. Using the localized representation (13), which diagonalizes the total molecular Hamiltonian, and the Frank–Condon approximation, we arrive at the following expressions for the amplitudes (2) of the symmetry-allowed ($F_{f_g}$) and symmetry-forbidden ($F_{f_u}$) scattering channels:

$$F_{f_\alpha}=\frac{1}{2}\,\mathscr{D}_{gu}\mathscr{D}'_{\beta f_\alpha}\sum_{\mathbf{m}}\ (\pm\langle 0,\mathbf{o}|1,\mathbf{m}\rangle G_{\mathbf{m}}\langle 1,\mathbf{m}|f_\alpha,\mathbf{m}_f\rangle$$
$$+\langle 0,\mathbf{o}|2,\mathbf{m}\rangle G_{\mathbf{m}}\langle 2,\mathbf{m}|f_\alpha,\mathbf{m}_f\rangle),$$
$$G_{\mathbf{m}}=\frac{1}{\Omega-m_g\omega_g-m_u\omega_u+i\Gamma},$$
$$\Omega=\omega-\left(\nu_0+\frac{1}{2}(\omega_g-\omega_{0g}+\omega_u-\omega_{0u})\right). \tag{18}$$

According to the dipole selection rules, in the matrix element $\mathscr{D}'_{\beta f_\alpha}$ of the dipole moment of the emission transition the parity $\beta$ of the $K$-excited state is opposite to the parity $\alpha$ of the final state (if $\alpha=g,u$, then $\beta=u,g$). In the scattering amplitude (18) the ''plus'' corresponds to an even-parity final state ($\alpha=g$) and the ''minus'' to an odd-parity final state ($\alpha=u$); the frequency $\nu_0=E-E_0$ is the difference between the equilibrium values of the electron energies of the $K$-excited state, $E$ (Eq. (7)), and the ground state, $E_0$. The Frank–Condon factor between the vibrational wave functions of the ground state $|0,\mathbf{o}\rangle$ and the $K$-excited state $|i,\mathbf{m}\rangle$ ($i=1,2$) is denoted by $\langle 0,\mathbf{o}|i,\mathbf{m}\rangle$, while $\langle i,\mathbf{m}|f_\alpha,\mathbf{m}_f\rangle$ stands for the Frank–Condon factor between the vibrational wave functions of the $K$-excited, $|i,\mathbf{m}\rangle$, and final, $|f,\mathbf{m}_f\rangle$, states. The vectors $\mathbf{o}=(0,0)$, $\mathbf{m}=(m_g,m_u)$, and $\mathbf{m}_f=(m_{f_g},m_{f_u})$ stand for the set of quantum numbers of the vibrational levels of the symmetric and antisymmetric normal modes of the ground, $K$-excited, and final states, respectively. Here we assume that the medium is kept at room temperature, at which only the lower vibrational levels of the ground electronic state with frequencies $\omega_{0g}$ and $\omega_{0u}$ are occupied ($m_{0g}=m_{0u}=0$) in moderate-sized molecules. The matrix elements of the electron dipole moment,

$$\mathscr{D}_{gu}=\langle\Psi_g^0|\mathscr{D}|\Psi_u\rangle,$$
$$\mathscr{D}'_{uf_g}=\langle\Psi_u|\mathscr{D}'|\Psi_{f_g}\rangle,\quad \mathscr{D}'_{gf_u}=\langle\Psi_g|\mathscr{D}'|\Psi_{f_u}\rangle, \tag{19}$$

describe the processes of absorption ($\mathscr{D}_{gu}$) and emission ($\mathscr{D}'_{uf_g}$ and $\mathscr{D}'_{gf_u}$), where $\Psi_\alpha^0$, $\Psi_\alpha$, and $\Psi_{f_\alpha}$ are the electron wave functions of the ground, $K$-excited, and final states of parity $\alpha$.

In contrast to the differential cross section (16) and the scattering amplitude (18), in the partial integral cross section $\sigma_{0\alpha}(\omega)$ (Eq. (17)) we can sum over the vibration levels of the final state:

$$\sigma_{0\alpha}(\omega)=\frac{1}{2}r_0^2\omega\overline{\omega^3}\mathscr{D}_{gu}^2\mathscr{D}'_{\beta f_\alpha}\sum_{\mathbf{m}}\Bigg[(|\langle 0,\mathbf{o}|1,\mathbf{m}\rangle|^2$$
$$+|\langle 0,\mathbf{o}|2,\mathbf{m}\rangle|^2)|G_{\mathbf{m}}|^2\pm 2\ \mathrm{Re}\sum_{\mathbf{m}_1}\ |\langle 0,\mathbf{o}|1,\mathbf{m}\rangle$$
$$\times G_{\mathbf{m}}|\langle 1,\mathbf{m}|2,\mathbf{m}_1\rangle G_{\mathbf{m}_1}^*\langle 2,\mathbf{m}_1|0,\mathbf{o}\rangle\Bigg]. \tag{20}$$

Note that the partial integral cross section (20) does not coincide with the absorption cross section. According to the optical theorem, the absorption cross section is $\Sigma\sigma_{0\alpha}(\omega)$, where we mean summation over all the final state of the molecule and the final photon (the corresponding analysis for RIXS can be found in Ref. 23). Obviously,

$$\langle 1,\mathbf{m}|2,\mathbf{m}_1\rangle=\delta_{m_g,m_{1g}}\langle 1,m_u|2,m_{1u}\rangle, \tag{21}$$

since the nonadiabaticity parameter $V$ (Eq. (14)) does not change the vibrational wave functions of the symmetric mode. The antisymmetric modes $|1,m_u\rangle$ and $|2,m_u\rangle$ of the electronic states (13) differ only in their equilibrium distances, which are shifted in relation to each other by $2\lambda/(M_u\omega_u^2)$ (Eq. (15)). The expression for the Frank–Condon factor (21) is given below (see Eq. (28)).

In the total integral cross section $\sigma_0(\omega)$ (Eq. (16)) it is advisable to distinguish the cross section $\sigma_{0g}(\omega)$ of the allowed scattering channel and a parameter $\chi(\omega)$ equal to the relative intensity of the forbidden scattering channel:

$$\sigma_0(\omega)=\sigma_{0g}(\omega)(1+\chi(\omega)),\quad \chi(\omega)=\frac{\sigma_{0u}(\omega)}{\sigma_{0g}(\omega)}. \tag{22}$$

The restoration of the selection rules ($\chi(\omega)=0$ and $F_{f_u}=0$) in the adiabatic limit ($\lambda=0$) follows directly from Eqs. (18) and (20), since in this limit the vibrational wave functions of the electronic states, Eqs. (13), coincide ($|1,\mathbf{m}\rangle=|2,\mathbf{m}\rangle$).

### 3.1. Restoration of the selection rules at large detunings

Let us examine the important limiting case of large detunings $\Omega$ of the initial photon from the electron–vibrational absorption band under consideration,

$$(\Omega^2+\Gamma^2)^{1/2}\gg\Delta\omega, \tag{23}$$

where the effective width of the band is $\Delta\omega$. In this limit of large detunings or large natural widths $\Gamma$ of the absorption line, the quantity $G_{\mathbf{m}}$ specified in Eq. (18) is independent of $\mathbf{m}$ and can be taken outside the sum in the expression for the scattering amplitude $F_{f_\alpha}$ in Eq. (18). Allowing for the completeness condition

$$\sum_{\mathbf{m}} |i,\mathbf{m}\rangle\langle i,\mathbf{m}|=1, \quad i=1,2,$$

we arrive at the following asymptotic expressions for the scattering amplitude and the relative intensity $\chi(\omega)$ of the forbidden scattering channel:

$$F_{f_u}=0, \quad F_{f_g}=\mathscr{D}_{gu}\mathscr{D}'_{uf_g}\frac{\langle 0,\mathbf{o}|f_g,\mathbf{m}_f\rangle}{\Omega+i\Gamma}, \quad \chi(\omega)=0, \tag{24}$$

which demonstrates that the scattering channel into the odd-parity final state in the limit (12) is again parity-forbidden ($F_{f_u}=0$ and $\chi(\omega)=0$) in the limit (23). This result is rigorous and model-independent.[15,16] The case of large $\Gamma$ in broadband excitation was also examined be Cederbaum.[17] Equation (24) implies that the carrier frequency $\omega$ of the excitation radiation provides a means of actively controlling the selection rules in nonadiabatic transitions.

### 3.2. Analysis of the selection rules in the time-dependent representation

To clarify the physical meaning of the phenomenon, it is advisable to write the scattering amplitude (18) in the time-dependent representation:

$$\begin{aligned}F_{f_\alpha}&=\int_0^\infty dt\ F_{f_\alpha}(t)\\&=\frac{i}{2}\mathscr{D}_{gu}\mathscr{D}'_{uf_g}\sum_{\mathbf{m}}\int_0^\infty dt\ \exp(i(\Omega-m_g\omega_g\\&\quad-m_u\omega_u+i\Gamma)t)[\pm\langle 0,\mathbf{o}|1,\mathbf{m}\rangle\langle 1,\mathbf{m}|f_\alpha,\mathbf{m}_f\rangle\\&\quad+\langle 0,\mathbf{o}|2,\mathbf{m}\rangle\langle 2,\mathbf{m}|f_\alpha,\mathbf{m}_f\rangle].\end{aligned} \tag{25}$$

This expression can be interpreted as follows. Because of photoabsorption at time $t=0$, the molecule goes into the $K$-excited state. According to quantum theory, it is impossible to say at what time $t$ the molecule will emit a photon and go into the final state. What the theory does state is that the spontaneous transition probability amplitude $F_{f_\alpha}(t)$ is different for different times $t$. The integral (25) over the emission times $t$ sums all the partial amplitudes $F_{f_\alpha}(t)$ into the total scattering amplitude $F_{f_\alpha}$. Obviously, not all times $t$ of emission of the final photon are equally probable, so there is a strong correlation between the emission and absorption acts in Raman scattering. The primary reason for this correlation is the finite lifetime ($\Gamma^{-1}$) of the $K$-excited state. Indeed, Eq. (25) and the physical picture of the process show that the delay time $t$ between the absorption and emission acts cannot be much longer than the lifetime $\Gamma^{-1}$ of the intermediate state. The second mechanism of the correlation between absorption and emission differs qualitatively from the first and is due to the phase difference

$$\varphi=\Omega t \tag{26}$$

between the ''instantaneous'' amplitude $F_{f_\alpha}(0)$ and the amplitude $F_{f_\alpha}(t)$ of emission of the final photon at time $t$. To be more exact, this phase is $(\Omega-m_g\omega_g-m_u\omega_u)t$. Because of the interference of the partial scattering amplitudes $F_{f_\alpha}(t)$, their contribution to the total scattering amplitude $F_{f_\alpha}$ is significant only for delay times $t\lesssim|\Omega|^{-1}$. Thus,

$$\tau_c=(\Omega^2+\Gamma^2)^{-1/2} \tag{27}$$

can be interpreted as the duration of the RIXS process, or the correlation time (the effective delay time) between the absorption and emission acts.

Restoration of the selection rules (24) for short correlation times $\tau_c\ll\Delta\omega^{-1}$ (see Eq. (23)), corresponding to instantaneous RIXS, means that in the course of the Raman process (the time interval $\tau_c$) the molecule is unable to perform an antisymmetric vibration ($\tau_c\ll\omega_u^{-1}$) capable of mixing the electronic states with different parities.

### 3.3. Identical potential surfaces for the ground and *K*-excited states

Calculations can be made simpler if the initial potential surfaces of the ground and $K$-excited states coincide (Fig. 1b). In this case $\omega_0\equiv\omega_{0g}=\omega_g=\omega_u$ and the Frank–Condon factors for the symmetric mode disappear, in view of which

$$\langle 0,0|i,m_g\rangle=\delta_{m_g,0}.$$

The fact that the remaining Frank–Condon factors for the antisymmetric mode are off-diagonal is due solely to the shift in the equilibrium distances in (15) ($\mp\lambda/M_u\omega_u^2$ for the states $\Psi_1$ and $\Psi_2$, respectively):

$$\begin{aligned}&\langle 0,0|i,m\rangle=\frac{(\pm\eta)^m}{\sqrt{m!}}\exp\left(-\frac{\eta^2}{2}\right), \quad \eta=\frac{\lambda}{\sqrt{2M_u\omega_u^3}},\\&\langle 1,m|2,m_1\rangle=e^{-2\eta^2}\\&\qquad\times\begin{cases}(2\eta)^{m-m_1}\left(\dfrac{m_1!}{m!}\right)^{1/2}L_{m_1}^{m-m_1}(4\eta^2),\\ \quad m>m_1,\\(-2\eta)^{m_1-m}\left(\dfrac{m!}{m_1!}\right)^{1/2}L_m^{m_1-m}(4\eta^2),\\ \quad m<m_1.\end{cases}\end{aligned} \tag{28}$$

Here $L_n^\alpha(x)$ is the generalized Laguerre polynomial, and $i=1,2$. In the first equation in (28) the ''plus'' corresponds to state $\Psi_1$ and the ''minus'' to state $\Psi_2$.

The integral cross section $\sigma_0(\omega)$ in Eq. (22) is expressed in terms of the integral RIXS cross section of the allowed transition,

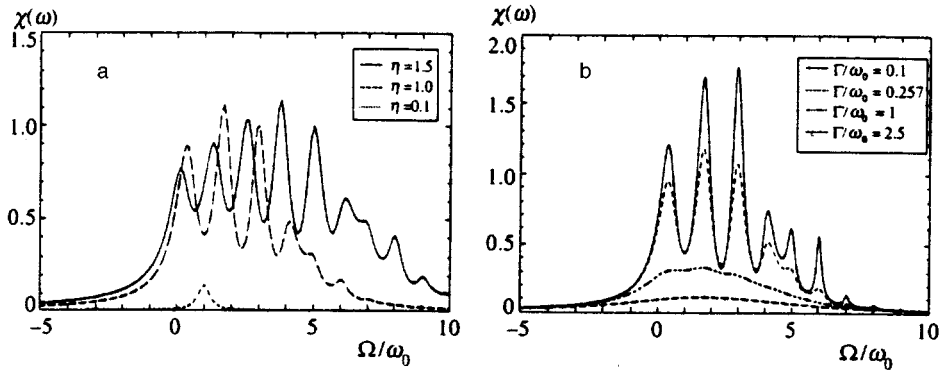



FIG. 2. Dependence of the relative intensity of the forbidden transition, $\chi(\omega)$ (Eq. (30)), on the frequency $\omega$ of the excitation radiation: (a) for different values of the nonadiabaticity parameter $\eta$ (Eq. (28)) ($\Gamma/\omega_0=0.257$), and (b) for different ratios of the natural absorption linewidth to the vibration frequency, $\Gamma/\omega_0$ ($\eta=1$).

$$\sigma_{0g}(\omega)=r_0^2\omega\overline{\omega'^3}\mathscr{D}_{gu}^2\mathscr{D}_{uf_g}'^2 e^{-\eta^2}\sum_{m,m_1=0}^{\infty}\beta_{mm_1}[|G_m|^2 + \mathrm{Re}(G_m G_{m_1}^*)], \tag{29}$$

and the parameter

$$\chi(\omega)=\left(\frac{\mathscr{D}_{gf_u}'}{\mathscr{D}_{uf_g}'}\right)^2\frac{\Sigma_{m,m_1=0}^{\infty}\beta_{mm_1}(|G_m|^2-\mathrm{Re}(G_m G_{m_1}^*))}{\Sigma_{m,m_1=0}^{\infty}\beta_{mm_1}(|G_m|^2+\mathrm{Re}(G_m G_{m_1}^*))}, \tag{30}$$

which describes the frequency dependence of the intensity of the forbidden transition. By taking into account the properties of generalized Laguerre polynomials,[24] we can easily see that the coefficients $\beta_{mm_1}$ are symmetric:

$$\beta_{mm_1}=\frac{e^{-2\eta^2}}{m_1!}(2\eta^2)^{m_1}L_m^{m_1-m}(4\eta^2), \quad \beta_{mm_1}=\beta_{m_1m}. \tag{31}$$

Clearly, $\mathrm{Re}(G_m G_{m_1}^*)$ in (29) is the term describing the interference of the channels of scattering through different intermediate states.[25] The expression (30) for the function $\chi(\omega)$ clearly shows that the interference of these scattering channels plays an important role in the formation of the parity-forbidden electron–vibrational band.

The restoration of the selection rules in parity ($\chi(\omega)\to 0$) for the scattering tensor at large detunings (see the condition (23)) follows directly from Eq. (30) for $\chi(\omega)$, since in this limit

$$|G_m|^2-\mathrm{Re}(G_m G_{m_1}^*)\to 0.$$

The spectral dependence of the relative intensity $\chi(\omega)$ of a forbidden transition (Eq. (30)) shows that the characteristic frequency scale on which function $\chi(\omega)$ decays is determined by the effective width $\Delta\omega$ of the absorption spectrum $\sigma_{0g}(\omega)$ (see Refs. 15 and 16). This result, which agrees with the estimate (23), justifies the use of the simplified five-level model adopted in Refs. 15 and 16. The dependence of the spectral shape of $\chi(\omega)$ on the nonadiabaticity parameter $\eta$ (Eq. (30)), depicted in Fig. 2a, shows that this function tends to zero in the adiabatic limit $\eta\to 0$. The narrowing of $\chi(\omega)$ to a single resonance with a natural linewidth $\Gamma$ is due to the specific features of the adopted model. Indeed, in the model we are using the potential surfaces of the ground and $K$-excited states coincide as $\eta\to 0$. Therefore, in this limit we have only a transition from the lower vibration level of the ground state to the lower vibrational level of the $K$-excited state (a 0–0 transition). Figure 2b depicts the spectral dependence of the function $\chi(\omega)$ for different ratios of the natural width of the x-ray absorption line and the vibrational frequency, $\Gamma/\omega_0$. The decrease in $\chi(\omega)$ with increasing $\Gamma$ was explained above in discussing the correlation time (27).

## 4. CONCLUSION

Due to electron–vibrational interaction, the RIXS spectrum symmetric molecules and solids acquires a parity-forbidden electron–vibrational band. We have studied the dependence of the ratio $\chi(\omega)$ (Eq. (22)) of the total intensities of the parity-forbidden and parity-allowed transitions in a linear triatomic molecule $XY_2$ on the duration $\tau_c$ (Eq. (27)) of the RIXS process. According to Eq. (27), the duration of the resonant scattering process depends on $\Omega$ and $\Gamma$. Being a constant quantity for a given spectral transition in the molecule, the natural linewidth $\Gamma$ is a passive parameter, not making it possible to actively vary $\chi(\omega)$. The extent to which the electron selection rules are violated, characterized by $\chi(\omega)$, can be actively controlled by another parameter, $\Omega$, by varying the difference between the initial-photon frequency $\omega$ and the center of the band of the electron transition being considered. We have found that the selection rules are strongly violated ($\chi(\omega)\sim 1$) in the region of maximum photoabsorption $0\lesssim|\Omega|\lesssim\Delta\omega$ and then are restored ($\chi(\omega)\simeq 0$) when the detuning of the initial-photon frequency reaches the wings of the absorption line: $|\Omega|\gg\Delta\omega$.

The adopted model of the RIXS effect with electron–vibrational interaction has made it possible to establish the spectral dependence of the extent of violation of the selection rules, $\chi(\omega)$, and to show that $\chi(\omega)$ is determined by the product ($\tau_c\Delta\omega$) of the duration of the RIXS process and the effective absorption linewidth (see Eqs. (23) and (27)). Figure 2a shows that when electron–vibrational interaction is fairly strong ($\eta\sim 1$), the width of the function $\chi(\omega)$ is given by the effective width $\Delta\omega$ of the absorption line. A decrease in the electron–vibrational interaction parameter $\eta$ (Eq. (28)) leads to decrease in the function $\chi(\omega)$, which narrows to a single resonance with a width $\Gamma$ (Fig. 2a). For $\Gamma/\omega_0$ small and $\eta\sim 1$, the spectral dependence $\chi(\omega)$ is of a clearly resonant nature. An increase in $\Gamma/\omega_0$ smears out this resonance

(Fig. 2b). The dependence of $\chi(\omega)$ on $\omega$ smooths out at $\Gamma/\omega_0 \gtrsim 1$, which justifies the use of the five-level model for studying the effect.[15,16]

We briefly discuss our simplifying assumptions. By using the above model in studying the RIXS process we wish only to establish the main qualitative laws governing the effect of restoration of the selection rules in the RIXS process. Note that the model can be applied directly only to linear triatomic molecules $XY_2$ that remain linear in the intermediate and final electronic states. The model cannot be applied directly to the case of hard x-ray radiation, in which dipole selection rules are violated.[5] Our model allows for symmetry breaking in the intermediate highly excited electronic state, and the main result of our research, Eq. (18), is based on this realistic assumption. Here we did not touch on the random quasidegeneracy of the final states (possible at least in principle), which can also lead to violation of selection rules. Since we do not claim that the model is universal, we can only state that it yields good results when applied to the $CO_2$ molecule, studied earlier by more rigorous methods.[16] In our numerical calculations of the Frank–Condon factors we employed the harmonic approximation. This approximation violates no principles and, being extremely accurate, is widely used in calculating x-ray transitions far for the dissociation thresholds of molecules.

The present work was done with partial financial support from the Russian Fund for Fundamental Research (Grant No. 96-0219826) and the Swedish Natural Science Research Council. The authors are grateful to S. G. Rautian and A. M. Shalagin for useful remarks.

[1] J.-H. Guo, P. Glans, P. Skytt, N. Wassdahl, J. Nordgren, Y. Luo, H. Ågren, Y. Ma, T. Warwick, P. Heimann, E. Rotenberg, and J. D. Denlinger, Phys. Rev. B **52**, 10 681 (1995).
[2] P. Cowan, in *Resonant Anomalous X-ray Scattering: Theory and Applications*, G. Materlik, C. J. Sparks, and K. Fisher (eds.), North-Holland, Amsterdam (1994), p. 449.
[3] Z. F. Liu, G. M. Bancroft, K. H. Tan, and M. Schachter, Phys. Rev. Lett. **72**, 621 (1994).
[4] S. Aksela, E. Kukk, H. Aksela, and S. Swensson, Phys. Rev. Lett. **74**, 2917 (1995).
[5] F. Gel'mukhanov and H. Ågren, Phys. Rev. A **49**, 4378 (1994).
[6] Y. Luo, H. Ågren, and F. Gel'mukhanov, J. Phys. B **27**, 4169 (1995).
[7] Y. Luo, H. Ågren, F. Gel'mukhanov, J.-H. Guo, P. Skytt, N. Wassdahl, J. Nordgren, Phys. Rev. B **52**, 14 479 (1995).
[8] P. Glans, K. Gunnelin, P. Skytt, J.-H. Guo, N. Wassdahl, J. Nordgren, H. Ågren, and F. Kh. Gel'mukhanov, T. Warwick, E. Rotenberg, Phys. Rev. Lett. **76**, 2448 (1996).
[9] P. Glans, P. Skytt, K. Gunnelin, J.-H. Guo, and J. Nordgren, J. Electron Spectrosc. Relat. Phenom. **82**, 193 (1996).
[10] F. Kh. Gel'mukhanov, L. N. Mazalov, and N. A. Shklyaeva, Zh. Éksp. Teor. Fiz. **69**, 1971 (1975) [Sov. Phys. JETP **42**, 1001 (1975)].
[11] F. Kh. Gel'mukhanov, L. N. Mazalov, and N. A. Shklyaeva, Zh. Éksp. Teor. Fiz. **71**, 960 (1976) [Sov. Phys. JETP **44**, 504 (1976)].
[12] F. Gel'mukhanov and H. Ågren, J. Phys. B **29**, 2751 (1996).
[13] Y. Ma, Phys. Rev. B **49**, 5799 (1994).
[14] P. Skytt, J.-H. Guo, N. Wassdahl, J. Nordgren, Y. Luo, and H. Ågren, Phys. Rev. A **52**, 3572 (1995).
[15] P. Skytt, P. Glans, J.-H. Guo, K. Gunnelin, C. Såthe, J. Nordgren, F. Kh. Gel'mukhanov, A. Cesar, and H. Ågren, Phys. Rev. Lett. **77**, 5035 (1996).
[16] A. Cesar, F. Gel'mukhanov, Y. Luo, H. Ågren, P. Skytt, P. Glans, J.-H. Guo, K. Gunnelin, and J. Nordgren, J. Chem. Phys. **106**, 3439 (1997).
[17] L. S. Cederbaum, J. Chem. Phys. **103**, 562 (1995).
[18] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 3rd ed., Nauka, Moscow (1989), p. 255 [English translation: Pergamon Press, Oxford (1991)].
[19] F. Gel'mukhanov and H. Ågren, Phys. Lett. A **193**, 375 (1994).
[20] G. B. Armen and H. Wang, Phys. Rev. A **51**, 1241 (1995).
[21] F. Gel'mukhanov and H. Ågren, Phys. Rev. A **54**, 3960 (1996).
[22] R. Englman, *The Jahn–Teller Effect*, Wiley, New York (1972).
[23] F. Gel'mukhanov, H. Ågren, M. Neeb, J.-E. Rubensson, and A. Bringer, Phys. Lett. A **211**, 101 (1996).
[24] A. Erdélyi, *Higher Transcendental Functions*, (Bateman Project) Vol. 2, McGraw-Hill, New York, (1953).
[25] F. Gel'mukhanov, L. N. Mazalov, and A. V. Kondratenko, Chem. Phys. Lett. **46**, 133 (1977).

Translated by Eugene Yankovsky

# Adsorption of cesium atoms at structural defects on sapphire surfaces

A. M. Bonch-Bruevich, T. A. Vartanyan, Yu. N. Maksimov, S. G. Przhibel'skiĭ, and V. V. Khromov

*S. I. Vavilov State Optics Institute, 199034 St. Petersburg, Russia*
(Submitted 16 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 362–370 (July 1997)

Results are presented from an experimental study of the adsorption of cesium atoms on sapphire surfaces and their photostimulated desorption. The adsorbed atoms are found to form chains on the surface which are localized near one-dimensional structural defects of the surface. One-dimensional adsorption is analyzed theoretically with different assumptions regarding the mobility of adsorbed atoms along a chain. A comparison of the theories with experimental data favors localized adsorption described in terms of a one-dimensional lattice gas model. The energy of adsorption for an isolated atom on a linear surface structural defect is 0.58 eV, while the energy of attraction between neighboring atoms in the chain is 26 meV. © *1997 American Institute of Physics.* [S1063-7761(97)03207-1]

## 1. INTRODUCTION

Studies of atomic and molecular adsorption on solid surfaces are a classical and widely used method for determining the energy and structural properties of surfaces. In this method the experimentally measured adsorption isotherms (or isobars) are compared with a theoretical analysis of equations of state for the adsorbed phase that have been obtained with different initial assumptions regarding the nature of the adsorption. This kind of comparison is effective and unique only in the simplest cases, i.e., for adsorption on energetically uniform surfaces when no lateral interactions exist, etc. In more complicated cases, additional experimental data on the processes taking place during adsorption are needed. An understanding of one of the basic properties of the test surfaces, namely, the inhomogeneity of their adsorptive characterstics, requires either complicated quantum mechanical calculations or the use of delicate and expensive experimental techniques, most of which are applicable only to electrically conducting surfaces. In this paper the nature of the adsorption centers is evaluated on the sole basis of the thermodynamic characteristics of the processes involved in the desorption of cesium atoms from sapphire surfaces. A combination of two methods is used here to study the surface properties of sapphire: one, which we have discovered previously,[1] uses nonthermal photostimulated desorption of alkali metals from single crystal sapphire surfaces, and the other is laser thermal desorption.

An analysis of the experimental data showed that with saturation adsorption the Cs atoms occupy a small fraction of the surface of a single crystal, while lateral interactions among the atoms make a significant contribution to the energetics of adsorption. This seeming contradiction is resolved by the fact that adsorption occurs at centers which form agglomerations that can, in general, have the attributes of clusters, islands, or filaments. A qualitative analysis of the cluster and island adsorption models shows that they cannot be reconciled with the experimental data. Using the reliably established existence of steps on the surface of ionic crystals,[2] in particular on sapphire,[3] and their role in increasing the adsorption energy,[4] we interpret the experimental results as a manifestation of one-dimensional adsorption in the form of filaments. This sort of adsorption can be described in terms of a one-dimensional lattice gas model, modified to take into account the contact between the adsorbed and bulk phases. The good agreement between the theoretical curves and the experimental data allows us to estimate both the adsorption energy for atoms at a step and the interaction energy of neighboring adsorbed atoms.

## 2. EXPERIMENTAL APPARATUS AND RESULTS

All the experiments were done with sapphire slabs cut from single crystal $\alpha$-$Al_2O_3$ parallel to the (0001) plane. The mechanically polished and chemically cleaned surface of the crystal slab was turned inward toward a sealed-off vacuum system with a cryogenic pump. A droplet of metallic cesium was placed in a special extension and served as a source of test particles. All the experiments were conducted under dynamic equilibrium conditions between the bulk and surface phases. The design of the cell made it possible to vary independently the temperatures of the test surface and the metal droplet. This made it possible to study the temperature and density characteristics of the adsorption, while avoiding simultaneous condensation of metal on the test surface. The temperature dependence of the photodesorption and thermal desorption of cesium atoms from the sapphire surface was studied.

Photodesorption was triggered by pulsed light from a ruby laser, whose wavelength, $\lambda = 694$ nm, lies in an absorption band of the adsorbed atoms. To detect the desorbed particles a 2-cm-diameter probe beam resonant with the strong $\lambda = 894.3$ nm absorption line of the free atoms was directed parallel to the test surface, almost touching it. In this geometry the change in the absorption of the probe beam was measured when a pulsed flux of desorbed atoms arrived in it; this made it possible to estimate the number of atoms and their average ejection velocity based on their time of flight across the cross section of the detector beam. Note that the presence of stronger, but constant in time, absorption by

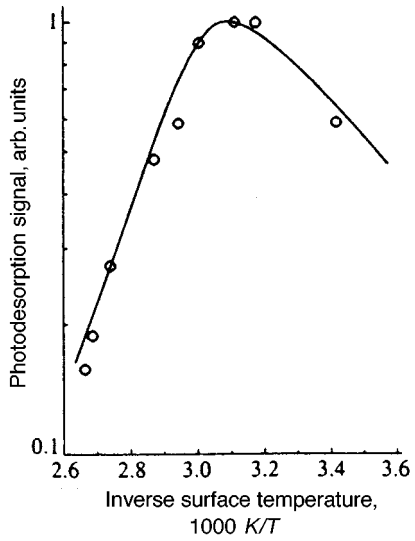



FIG. 1. The signal produced by the photodesorption of cesium atoms from a sapphire surface as a function of the reciprocal surface temperature for a constant concentration $N=3\cdot10^{10}$ cm$^{-3}$ of cesium atoms in the gaseous phase. Driver radiation wavelength $\lambda=694$ nm, pulse length 0.2 ms, and intensity 350 W/cm$^2$. The points are experimental data and the smooth curve, a calculation using Eqs. (2), (3), and (4).

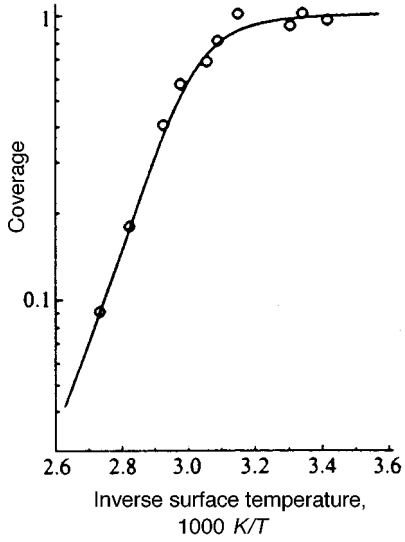



FIG. 2. The signal produced by thermal desorption of cesium atoms from a sapphire surface as a function of the reciprocal surface temperature for a constant density $N=3\cdot10^{10}$ cm$^{-3}$ of cesium atoms in the gaseous phase. The driver radiation had wavelength $\lambda=694$ nm, pulse length 0.2 ms, and intensity $10^5$ W/cm$^2$. The points are experimental data and the smooth curve is from a calculation using Eq. (4) with $E_i=0.026$ eV and $E_a=0.58$ eV. Within this temperature range a calculation using Eq. (1) with $E_a=0.95$ eV is indistinguishable from the smooth curve on the scale of the figure.

atoms in the gas phase does not interfere with these measurements. The average velocity of the thermally desorbed atoms corresponded to 700 K, and of the photodesorbed atoms, to 500 K. The number of desorbed atoms was determined from the amplitude of the change in the intensity of the probe beam. Here we took the resonant absorption cross section of the atoms to be $\sigma=3\cdot10^{-12}$ cm$^{-2}$, which corresponds to a Doppler mechanism for broadening of the atomic line at the temperatures indicated above. The distributions of the desorbed atoms with respect to velocity and ejection angle were not recorded in the present experiments. Similar measurements for an Na/$\alpha$-$Al_2O_3$ system[1] show, however, that the velocity distribution of the ejected atoms is Maxwellian to high precision with a temperature $T=700-900$ K, depending on the wavelength of the exciting radiation, while the angular distribution is close to diffuse.

The experimentally measured dependence of the number of photodesorbed atoms on the substrate temperature for a fixed concentration of gaseous phase atoms is shown in Fig. 1. The photodesorption quantum yield at room temperature was $(5\pm2)\cdot10^{-6}$.

A linear dependence of the number of desorbed atoms on the incident radiation power, which is typical of the photodesorption process described above, is observed for intensities below 10 kW/cm$^2$. At higher intensities a thermal desorption phenomenon is observed with a threshold rise in the number of desorbed atoms, which is soon replaced by saturation owing to complete desorptoin of adatoms during the time the laser pulse acts. The number of atoms desorbed in this regime is determined only by the surface concentration of the adsorbate, and we have used this to measure the latter as a function of substrate temperature.

The measured surface concentration of adsorbed atoms is shown in Fig. 2. At low temperatures the surface concentration as a function of temperature undergoes saturation, reaching a level on the order of $10^{13}$ cm$^{-2}$. The estimated systematic error in our determination of this quantity is $\pm50\%$, and the scatter in the experimental points shown in Figs. 1 and 2 characterizes the relative error in the experiment, which is $\pm10\%$. Since our value for the surface concentration is substantially lower than the maximum with dense packing, $10^{15}$ cm$^{-2}$, it is obvious that adsorption does not take place over the entire surface, but only at adsorption centers whose density determines the maximum adsorbate concentration.

### 3. DISCUSSION OF RESULTS

Given that the detected saturation density of adsorbed Cs atoms is much lower (Fig. 2) than the density with dense packing, it is natural to assume the possibility of adsorption at centers. In principle, adsorption at centers can be analyzed in terms of two models: cluster and Langmuir. However, the first model, different versions of which have been examined in detail elsewhere,[5] cannot explain the invariability of the characteristics of the photoprocesses and adsorption energy with covering density, while the maximum numbers of particles in a cluster can be large. In addition, in all the experiments the surface temperature of the sapphire was somewhat higher than the temperature of the metallic cesium droplet which created the equilibrium atomic cesium vapor pressure in the volume. Under these conditions clusters cannot be stable if we make the obvious assumption that the binding energy of the atoms in them is lower than the binding energy in the large droplet of metal.

These difficulties are eliminated by the Langmuir model, so it is natural to attempt to interpret the experimental data

using such a model. Let $n_0$ be the surface density of centers. Then the surface density of adsorbed atoms is given by

$$n=\frac{n_0}{1+\exp(E_a/T_*-E_a/T)}, \tag{1}$$

where $T_*$ is determined by the density of particles in the gaseous phase and represents the temperature at which half the adsorption centers are filled. For $T>T_*$, the surface density falls off exponentially with rising temperature, while for $T<T_*$ it rapidly reaches its maximum value $n_0$. The best agreement between Eq. (1) and the experimental data is obtained for $T_*=336$ K and $E_a=11000$ K.

Despite the qualitative agreement between the theoretical and experimental data on thermal desorption (Fig. 2), the estimate for $E_a$ obtained from it cannot be made consistent with the data on photodesorption. In fact, the photodesorption of atoms from the surface of wideband dielectrics can be explained in the following way:[1] in the first stage, because of the transparency of sapphire in the range of the absorption bands of the adatoms, selective excitation of the optical electron of an adsorbed atom takes place without significant perturbation of the electron and phonon subsystems of the substrate. Later, the electronic excitation of the adatom is quenched rapidly and the energy of the photon is redistributed statistically among the vibrational degrees of freedom of the adsorbed atom and the substrate atoms closest to it. Here some fraction of the adatoms may obtain enough energy to overcome the attractive potential to the surface, $E_a$, and be desorbed from the surface.

A study of the time of flight spectra of the desorbed atoms showed that they are well described by an elementary model[1] in which a uniform distribution of the photon energy among a fixed number of degrees of freedom of the adsorption complex $z$ is assumed. This model leads to a Maxwellian distribution of the desorbed atoms with respect to their ejection velocity with an effective temperature that depends linearly on the photon frequency $\omega$ and is independent of the intensity of the exciting radiation. Taking $T$ as the surface temperature, we can write the effective temperature of the adsorption complex after absorption of a photon and quenching of the resulting excited electron state as

$$T_e=T+2\hbar\omega/z. \tag{2}$$

The quantum yield of this process can then be written in the form

$$f=\sqrt{\frac{E_a}{T_e}}\exp\left(-\frac{E_a}{T_e}\right). \tag{3}$$

As noted above, the temperature dependence of the photodesorption signal, which is proportional to the product $nf$ (Fig. 2), and the experimentally determined quantum yield at room temperature, $f_0=5\cdot10^{-6}$ is inconsistent with the estimate of $E_a$ from the thermodesorption data if Eq. (1) is used for $n$. Here we notice at once that the value of $E_a=0.95$ eV determined in this way is considerably higher than the value $E_a=0.7$ eV given in Ref. 1 for adsorption of Na on sapphire and is even greater than the binding energy of an atom in metallic cesium.

The disagreement between the point-center models and the experimental data means that it is necessary to consider adsorption in models of reduced dimensionality. It is natural to assume that one dimensional adsorption is realized in the form of chains lying along the bases of the steps which always exist on a real crystalline surface.[3] This assumption, based on the well known increase in the adsorption energy near the base of a step,[4] means that it is necessary to include the interaction between neighboring adsorbed atoms, since, in general, saturation of adsorption may occur with dense packing. In addition, two alternative situations must be considered: localized and delocalized adatom sites on the adsorption line. As will be shown below, including lateral interactions and choosing a localized adsorption model can provide an adequate desciption of the experimental results.

#### 4. ONE-DIMENSIONAL LOCALIZED ADSORPTION MODEL

Adsorption on linear chains of centers can be described theoretically in terms of a one-dimensional lattice gas model. Then the lateral interaction is determined by a single parameter, the nearest-neighbor interaction energy $E_i$. The thermodynamic parameters of this type of model can be calculated exactly using the methods employed in the theoretical analysis of the Ising model.[6] This model differs from the traditional model of a lattice gas in that the lattice system has contact with a bulk phase and an occupied lattice site has a nonzero energy. These differences, expressed through the chemical potential $\mu$ and the binding energy $E_a$ of its bond with a lattice site, show up as the energy $B$ of a spin in an external magnetic field in the Ising model:

$$2B=\mu+E_a+E_i$$

and, thereby, restrict the exact solubility of the model to the one-dimensional case alone. Without dwelling on the details of the derivation, which reduces basically to redefining the notation and making corrections to Eq. (16.77) of Huang's book, here we give only the final result for the surface density of adsorbed atoms,

$$n=\frac{n_0}{2}\left[1+\frac{\sinh[(E_a+E_i)/2T-(E_a+E_i)/2T_*]}{\sqrt{\sinh^2[(E_a+E_i)/2T-(E_a+E_i)/2T_*]+\exp(-E_i/T)}}\right]. \tag{4}$$

Equation (4) transforms to the Langmuir formula (1) if we neglect the lateral interactions, setting $E_i=0$. For $T>T_*$ the degree of filling of the adsorption centers is low and both formulas reduce to a simple exponential dependence,

$$n\propto\exp\left(\frac{E_a}{T}\right), \tag{5}$$

which coincides with the temperature dependence of the degree of coverage for unlocalized forms of adsorption. In the earlier studies of the adsorption of sodium atoms on sapphire surfaces,[1] just this sort of dependence of the degree of coverage on the surface temperature was observed, which made

it impossible to reach any definite conclusions about the existence or nature of any adsorption centers. During adsorption of Cs from a vapor with considerably greater elasticity at room temperature, saturation is observed with a substantially lower degree of coverage, and this is evidence of the presence of adsorption centers.

The theoretical curves constructed from Eq. (4) for different $E_i$ are essentially the same within the range of temperatures that was studied, provided the value of $E_a$ for each value of $E_i$ is chosen so that at the temperature $T_*$ corresponding to a coverage of 0.5, their slopes corrrespond to that of the experimental curve. In order to determine the model parameters uniquely, it is necessary to turn to the results of the photodesorption experiment.

It should be noted that in this model atoms occupying a different position in the chains have different adsorption energies. Single atoms have an adsorption energy $E_a$. The atoms at the ends, with only one neighbor, are bound with an energy of $E_a+E_i$. Finally, atoms lying within the chains have two neighbors each and, therefore, a binding energy of $E_a+2E_i$. In terms of the theoretical model the relative fractions of each type of atom can be calculated exactly. We shall not give the corresponding results here, since for high degrees of coverage most of the atoms belong to the last type and in comparisons with experiment the contributions of atoms with coordination numbers less than 2 can be neglected. Using the measured photodesorption quantum yield $f=5\cdot 10^{-6}$ at the lower limit $T_l=293$ K of the temperature range, we find the missing relationship between the parameters with Eq. (3):

$$\frac{E_a+2E_i}{T_l+\hbar\omega/z}=13.5. \tag{6}$$

Varying the sole parameter $E_i$ when condition (6) is satisfied, we can obtain good agreement between the theoretical curve and the experimental photodesorption data (Fig. 2). The best approximation is obtained for $E_i=300$ K, from which we find $E_a=6700$ K, $2\hbar\omega/z=250$ K, and $z=165$.

## 5. NONLOCALIZED ONE-DIMENSIONAL ADSORPTION MODEL

We examined the alternative possibility, delocalized one dimensional adsorption, in terms of a model of classical free motion of particles along adsorption lines with a pairwise lateral interaction between nearest neighbors. The repulsion of adatoms as they come close to one another and their attraction as they move apart were taken into account in the interaction.

Here we present only the results of the derivation of an adsorption equation of state. The temperature dependence of the one-dimensional concentration $c$ of adatoms for a constant volume concentration $N$ of the vapor is given by

$$c=T\frac{\partial s(\mu,T)}{\partial\mu}, \tag{7}$$

where the dependence of $s$ on the arguments $\mu$ and $T$ is expressed by the formula

$$\exp[(\mu-E_a)/T]\int_0^\infty \exp[-sx-U(x)/T]dx=\lambda, \tag{8}$$

in which $U(x)$ is the lateral interaction potential, $\lambda$ is the de Broglie wavelength of an atom with the thermal momentum, and $\mu=T\ \ln(N\lambda^3)$ is the atomic chemical potential.

These results yield the well-known Hill-de Boer equation of state[8] in the case of low coating densities and lateral attractive potentials that are small compared to the thermal energy.

We have analyzed the results of this model numerically over a wide range of parameters for a lateral potential in the form of a solid core with a rectangular well. A comparison of calculations based on this model with the experimental data showed that agreement between them is not satisfactory for realistic values of the model parameters. It turned out that in the model saturation of the adsorption is not fast enough as a function of surface temperature. This happens because when the lateral adatom attraction energies are not too large, their thermal motion inhibits adsorption saturation (because of the lateral repulsion). Better agreement is attained when the motion of the adatoms along the adsorption line is retarded. In the limit of complete freezing out of this motion, we return to the localized one-dimensional adsorption model discussed above.

## 6. CONCLUSION

The different approaches to describing adsorption on nonuniform surfaces and accounting for the interaction between adsorbed particles are being actively discussed in the modern literature on surfaces.[7] In this paper we have attempted to interpret the results of an experimental study of the adsorption of cesium atoms on sapphire in which effects associated with the energy inhomogeneity of the sapphire surface with respect to the adsorption of cesium atoms, as well as effects associated with the interaction of adsorbed atoms among themselves, have been observed.

A theoretical description has been provided in terms of a one dimensional lattice gas model. This model is especially attractive in that its thermodynamic characteristics can be calculated exactly using the methods employed in the theory of the Ising model. We note here that, although it cannot be solved exactly, the two dimensional model can be analyzed qualitatively and yields results that differ strongly from the experimental data. In fact, in two dimensional systems there is always a phase transition which manifests itself as a sharp jump in the coating density over a very narrow temperature range. In one dimensional systems, on the other hand, there is no phase transition,[9] in agreement with the smooth rise in the coating density observed experimentally as the surface temperature is lowered. From the standpoint of the general concepts of surface structure, there is no doubt of the existence of steps on real crystalline surfaces. Steps have been observed directly on an $\alpha$-$Al_2O_3$ surface using a scanning tunnelling microscope.[3] Furthermore, according to quantum mechanical calculations the adsorption energy at the edge of a step is higher than on the terraces.[4] The decoration of steps with metallic clusters, which is widely used in electron microscopy, also is related to the preferential deposition of at-

oms at the edges of steps. These facts suggest that steps are the most likely candidates for the linear structural defects which we have observed in our photodesorption and thermal desorption experiments. It should be emphasized that, as we have shown here, data on the temperature dependence of the degree of coating are not, in themselves, sufficient for a unique determination of the adsorption energy when lateral interactions are taken into account. Only the additional photodesorption data made it possible to establish that neighboring atoms attract one another and to find that their interaction energy $E_i = 300$ K or 0.026 eV. The adsorption energy $E_a$ for an isolated atom is 0.58 eV and the binding energy of an atom in the middle of a chain, $E_a + 2E_i$, equals 0.63 eV. If lateral interactions were neglected, then the Langmuir formula (1) would give an estimate $E_a = 0.95$ eV for the adsorption energy that is far too high and exceeds the binding energy of the atoms in metallic cesium. The lattice gas model assumes that adsorption centers, near which the adsorbed atoms spend most of their time, are positioned more or less regularly along the steps. A transition from one adsorption center to another involves overcoming an energy barrier whose height is, in any case, greater than the mean thermal energy of the atoms.

The attraction between neighboring atoms in a chain reduces the magnitude of the barrier, but, according to the above estimates of $E_i$ does not destroy it. The sign and magnitude of the lateral interaction between the atoms indicates that it is of van der Waals type. The large number of degrees of freedom of the substrate that are involved in the photodesorption process is related to the many-particle nature of the adsorption bond of the alkali metals on sapphire surfaces. In fact, the orbital radius of the weakly bound valence electron in the alkali metals is much greater than the distance between ions in the compact crystalline lattice of sapphire. In a study of the photodesorption of sodium ions from sapphire surfaces[1] we found $z = 120$. The increase in $z$ to 65 in the case of cesium can be related naturally to the fact that the radius of the outer electron shell is larger in cesium than in sodium. Evidently, the lower adsorption energy of cesium atoms (compared to sodium) on a sapphire surface is also related to this circumstance. Of course, cesium atoms are also adsorbed on terraces as well as on the steps, but the energy of adsorption on terraces is so much lower than on the steps that the terraces do not contribute signficantly to the overall density of adsorbed atoms. At the same time, photodesorption should be accompanied by efficient ejection onto the terraces of atoms adsorbed on the steps. The appearance of excess adsorbed atoms on the terraces, compared to the equilibrium amount, with an effective temperature higher than the surface temperature must cause a signficant change in the kinetics of all surface processes. We believe that these questions merit special study.

This work was carried out as part of the State Scientific Program on Surface Atomic Structures and partially supported by grants from the International Science Foundation (Grant No. R3K300) and the Russian Fund for Fundamental Research (Grant No. 96-02-16963a).

[1] A. M. Bonch-Bruevich, T. A. Vartanyan, A. V. Gorlanov, Yu. N. Maksimov, S. G. Przhibel'skiĭ, and V. V. Khromov, Zh. Éksp. Teor. Fiz. **97**, 1077 (1990) [Sov. Phys. JETP **70**, 604 (1990)].
[2] M. C. Gallagher, M. S. Fyfield, J. P. Cowin, and S. A. Joyce, Surf. Sci. **339**, L909 (1995).
[3] W. Mahoney, D. M. Schefer, A. Partil, R. P. Andres, and R. Reifenberg, Surf. Sci. **316**, 383 (1994).
[4] A. Zangwill, *Physics at Surfaces*, Cambridge Univ. Press, New York, 1988.
[5] Yu. K. Tovbin, *Theory of Physical and Chemical Processes at a Gas-Solid Interface*, CRC Press, Boca Raton, FL, 1991.
[6] K. Huang, *Statistical Mechanics*, Wiley, New York, 1963.
[7] A. J. Ramirez-Pastor and G. Zgrablich, Surf. Sci. **341**, 249 (1995).
[8] I. O. Protod'yakonov and S. V. Siparov, *The Mechanics of the Adsorption Process in Gas-Solid Systems* [in Russian], Nauka, Leningrad (1985).
[9] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Part 1, 3rd ed., Pergamon, Oxford, 1980.

Translated by D. H. McNeill

# Topological characteristics of electronic spectra of single crystals

I. A. Dynnikov[1)]

*M. V. Lomonosov Moscow State University, 119899 Moscow, Russia*

A. Ya. Maltsev[2)]

*L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia*

(Submitted 27 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 371–378 (July 1997)

The paper considers the topological characteristics of dispersion functions $\epsilon_s(\mathbf{p})$ in energy bands in single crystals related to classical electron trajectories in uniform magnetic fields. Specifically, the topological properties of open trajectories in $\mathbf{p}$-space on various energy levels within one energy band and related physical effects are described. © *1997 American Institute of Physics.* [S1063-7761(97)03307-6]

This paper is a continuation of an earlier publication[1] and treats physical effects related to the results recently obtained in three-dimensional topology. Initially, the problem statement was based on the physical effects of open quasiclassical electron trajectories in metal single crystals in a uniform magnetic field. Let us briefly review the situation.

Under conditions such that the quasiclassical approximation applies, one can introduce a one-particle electron distribution function $f(\mathbf{x},\mathbf{p},t)$ with respect to position and quasimomenta within a single energy band which satisfies the Boltzmann equation:

$$f_t + v^i_{\rm gr}(\mathbf{p})\frac{\partial f}{\partial x^i} + F^i(\mathbf{x},\mathbf{p},t)\frac{\partial f}{\partial p_i} = I[f(\mathbf{x},\mathbf{p},t)], \tag{1}$$

where $I[f]$ is the collision integral.

The group velocity $\mathbf{v}_{\rm gr}(\mathbf{p})$ is the gradient of the dispersion function $\epsilon_s(\mathbf{p})$ in the given energy band. This function is triply periodic in $\mathbf{p}$ with periods equal to the vectors of the reciprocal lattice. We are interested in the case

$$\mathbf{F}(\mathbf{x},\mathbf{p},t) = \frac{e}{c}\mathbf{v}_{\rm gr}(\mathbf{p})\times\mathbf{B} + \tilde{F}(\mathbf{x},t) \tag{2}$$

for strong magnetic field $\mathbf{B}$ (the criterion is $\omega_B \gg \tau^{-1}$, where $\tau$ is the time of the electron free path). In this case, the motion of electrons in quasimomentum space is determined to lowest order in $B$ by the system of equations

$$\frac{d\mathbf{p}}{dt} = \frac{e}{c}[\nabla\epsilon_s(\mathbf{p})\times\mathbf{B}]. \tag{3}$$

The integral trajectories derived from Eq. (3) are the intersection of planes perpendicular to the magnetic field $\mathbf{B}$ with surfaces of equal energies, $\epsilon(\mathbf{p}) = \text{const}$. An important role is played by open trajectories, which lie in the plane perpendicular to the magnetic field in a strip of finite width and pass through it. The specific role of such trajectories was analyzed in detail long ago.[2–4] In studies of metal conductivity in strong uniform magnetic fields determined by properties of time-independent solutions of Eq. (1) when $\tilde{\mathbf{F}}$ is a weak uniform electric field, striking manifestations of such trajectories were considered, and a contribution of such trajectories to the electrical conductivity tensor which is nonvanishing as $B\to\infty$ and leads to a finite conductivity not only along the magnetic field vector, but in all directions perpendicular to the average direction of such trajectories (which we denote by $\eta$) in an arbitrarily strong magnetic field. In this case, levels near the Fermi energy are most important, so the properties of the conductivity tensor are controlled by the topology of trajectories on the Fermi surface. This property was utilized in experimental studies of Fermi surfaces of metals reported in Refs. 5–8 and reviewed in detail in the monograph by Lifshits *et al.*[9]

Novikov[10–13] formulated the topological problem concerning classification of all possible open trajectories, which can be obtained by crossing between surfaces of constant energy of triply periodic arbitrary function and a plane in $\mathscr{R}^3$. In particular, he conjectured that any open trajectory should lie in a rectilinear strip of finite width, passing through it. This issue has been studied extensively.[14–20] Zorich[14] and Dynnikov[15–19] obtained results which led to a detailed description of the general case discussed below.

Suppose there is an arbitrary triply periodic function $\epsilon(\mathbf{p})$ with periods equal to vectors of the reciprocal lattice, whose values range between $\epsilon_{\min}$ and $\epsilon_{\max}$. Then,[16,18] in the case of a general magnetic field orientation (i.e., when the plane $\Pi(\mathbf{B})$ perpendicular to the magnetic field does not contain vectors of the reciprocal lattice), open trajectories exist either over a continuous energy interval $\epsilon_1 \leqslant \epsilon(\mathbf{p}) \leqslant \epsilon_2$, where $\epsilon_1 > \epsilon_{\min}$ and $\epsilon_2 < \epsilon_{\max}$, or on a single energy level $\epsilon_0$ (in particular, it is impossible that no energy level has open trajectories). In the former case, all open trajectories are in rectilinear strips of finite width in planes perpendicular to $\mathbf{B}$ and pass through these strips. The orientation vector $\eta$ of these strips is unique for all energy levels and open trajectories and is defined by the intersection between the plane $\Pi(\mathbf{B})$ and a certain (*a priori* unknown) integer plane $\Gamma$ (i.e., one generated by two vectors of the reciprocal lattice). Moreover, there is a stability zone around the field direction $\mathbf{B}$, which is a region of finite measure with a piecewise smooth boundary on the unit sphere $\mathscr{S}^2$ such that for all general orientations of $\mathbf{B}$ in this region a similar situation obtains, and the orientation of the plane $\Gamma$ is the same, although the direction $\eta$, interval boundaries $\epsilon_1$ and $\epsilon_2$, and the measure of open trajectories on each level are, generally speaking, different. On the boundary $\Omega$ of the stability zone, $\epsilon_1$ and $\epsilon_2$ are equal but

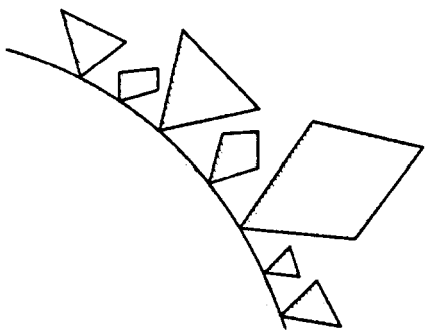


FIG. 1. Boundary of a stability zone with adjacent smaller zones.

the open trajectories are still in strips of finite width and their direction is defined, as in other cases, by the intersection between the plane $\Pi(\mathbf{B})$ and the integer plane $\Gamma$.

In the second case, however, when open trajectories exist only on a single energy level, they may, generally speaking, be not in a strip of a finite width (the magnetic field direction does not belong in this case to a boundary of any stability zone); moreover, for magnetic field directions with irrationality 3 [i.e., $\Pi(\mathbf{B})$ does not contain any vector of the reciprocal lattice], open trajectories do not attain an asymptotic direction, and the mean projection of the group velocity $\mathbf{v}_{\rm gr}(\mathbf{p})$ on $\Pi(\mathbf{B})$ on each of the associated closing components of such trajectories is zero.

The unit sphere $\mathscr{S}^2$ of magnetic field directions can be either a single stability zone or contain an infinite number of such zones, and in the latter case any two stability zones have no more than one common boundary point. The average direction $\eta$ of open trajectories corresponding to the common point should belong to the two integer planes $\Gamma_1$ and $\Gamma_2$ corresponding to these stability zones, i.e., it is defined by their intersection line. Thus, the direction of the magnetic field $\mathbf{B}$ corresponding to the common point of two stability zones $\Omega_1$ and $\Omega_2$ is such that $\Pi(\mathbf{B})$ contains the integer vector $\Gamma_1 \cap \Gamma_2$. The resulting pattern is far from trivial, namely, the unit sphere $\mathscr{S}^2$ contains an infinite number of zones with piecewise smooth boundaries, and an infinite number of smaller zones contact it at an infinite number of points[19] (Fig. 1).

The union of all the stability zones densely fills the sphere $\mathscr{S}^2$, but it has been hitherto unknown whether their total measure equals the measure of $\mathscr{S}^2$. The complement of the union of all zones on the sphere is an everywhere dense set, whose measure has not been determined yet. To general orientations of $\mathbf{B}$ from this set, the second of the situations described above applies.

This pattern provides a comprehensive description of the behavior of open trajectories for field directions of irrationality 3. Now let us discuss additional effects in the cases when the magnetic field alignment is such that $\Pi(\mathbf{B})$ contains vectors of the reciprocal lattice. In this case, two different situations are possible, namely,

(a) $\Pi(\mathbf{B})$ is generated by two vectors of the reciprocal lattice (irrationality 1);

(b) $\Pi(\mathbf{B})$ contains one vector (with an arbitrary integer factor) of the reciprocal lattice (the case of irrationality 2).

In case (a) any open trajectory is periodic (irrespective of whether the direction of $\mathbf{B}$ belongs to any stability zone or not). Thus, it is always in a strip of a finite width with an integer direction $\boldsymbol{\eta}$ and passes through the strip. If this direction is within any stability zone, the previously discussed trajectories corresponding to a general field alignment continue to exist, however, additional open trajectories directed along integer vectors in $\Pi(\mathbf{B})$ may emerge. The remarkable property of such trajectories is that each trajectory is stable against variations of $\mathbf{B}$ perpendicular to the integer vector $\boldsymbol{\eta}_\alpha$ [such that $\Pi(\mathbf{B}+\delta\mathbf{B})$ contains $\boldsymbol{\eta}_\alpha$] and vanishes for other variations in $\mathbf{B}$ (whereas the contribution of trajectories of general orientation persists for arbitrary variations in $\mathbf{B}$).

It also follows from the topological analysis that the directions $\eta_\alpha$ corresponding to the ''complementary'' trajectories also belong to the appropriate integer plane $\Gamma$. Hence, if $\Pi(\mathbf{B})$ does not coincide with $\Gamma$, these directions coincide with general directions, being intersection lines between $\Gamma$ and $\Pi(\mathbf{B})$. Like trajectories of a general orientation, the complementary trajectories occupy a finite phase volume, and their contribution to transport effects can be measured separately in experiments owing to their properties under variations in $\mathbf{B}$, as discussed above. If $\Pi(\mathbf{B})$ coincides with $\Gamma$ (correctly defined for a given stability zone), which is an exceptional case, the directions $\boldsymbol{\eta}_\alpha$ can be defined by any integer vectors in $\Pi(\mathbf{B})$. If there is more than one, an exceptional situation is realized, namely, the conductivity along all three axes remains finite as $B \to \infty$ for this specific field alignment (this property is lost for any variation in the field alignment). Note also that such a situation can be realized in the case of magnetic fields aligned with a rational direction outside all stability zones.[3)]

Zorich[14] (see also Ref. 15) proved that for an energy level of general orientation, for any rational direction of $\mathbf{B}$, there is a neighborhood such that for all directions of $\mathbf{B}$ of this neighborhood all open trajectories on this level (if any) are in a strip of a finite width and pass through it.

In case (b), if the direction of $\mathbf{B}$ belongs to one stability zone, periodic open orbits can also occur. They are absolutely identical to those described above, and are directed along an integer vector belonging to $\Pi(\mathbf{B})$ and aligned with the direction of trajectories of the general orientation. They have the same properties and produce no new effects in comparison to case (a), except that $\Pi(\mathbf{B})$ may be outside the integer plane $\Gamma$. Generally speaking, open trajectories in case (b) (if $\mathbf{B}$ does not belong to any stability zone) may not belong to a rectilinear strip of a finite width.[20] In this case, they also belong to a single energy level and correspond to a finite phase volume. Note, however, that in case (b) such trajectories, nonetheless, always have an asymptotic direction[18] defined as

$$\lim_{T\to\infty} \frac{1}{T} \frac{\mathbf{p}(T)-\mathbf{p}(0)}{|\mathbf{p}(T)-\mathbf{p}(0)|},$$

where $\mathbf{p}(t)$ is the trajectory radius-vector in the $\mathbf{p}$-space.

Now let us discuss in detail the topological situation in the case when $\mathbf{B}$ belongs to a stability zone. The function $\epsilon(\mathbf{p})$ can be defined on a three-dimensional torus $\mathscr{B}=\mathscr{T}^3$ (in the first Brillouin zone) obtained by factorizing the Euclid-

ean space $\mathscr{E}=\mathscr{R}^3$ in terms of the reciprocal lattice vectors. Integral trajectories of the flux generated in the quasimomentum space by the uniform magnetic field $\mathbf{B}=(B_1,B_2,B_3)$ are defined on each energy level by level surfaces of the form

$$\omega=B_1dp_x+B_2dp_y+B_3dp_z,$$

limited on this level from the three-dimensional torus $\mathscr{T}^3$. For fields $\mathbf{B}$ with irrational directions (irrationality 3), closed and open trajectories in $\mathscr{T}^3$ correspond to closed and open trajectories in $\mathscr{R}^3$, but if $\Pi(\mathbf{B})$ contains vectors of the reciprocal lattice, $\mathscr{R}^3$ may contain periodic open trajectories corresponding to closed trajectories in $\mathscr{T}^3$. Given a fixed alignment of the magnetic field, let us remove from each energy level in $\mathscr{B}$ the parts which consist of nonsingular trajectories closed in $\mathscr{E}$. The boundary of the resulting manifold consists of singular orbits closed in $\mathscr{E}$, which can be covered in $\mathscr{B}$ with disks in the plane $\Pi(\mathbf{B})$. There is a nontrivial topological theorem[16,17] which asserts that, for those directions of $\mathbf{B}$ which are in a stability zone or on its boundary, the resulting manifold is a union of nonintersecting tori in $\mathscr{B}$. Two-dimensional tori on different energy levels also do not intersect.

Two types of two-dimensional tori can emerge. The simplest one is a ''needle'' or an $N$-type torus defining a region in $\mathscr{B}$ and topologically homologous to zero in $\mathscr{B}$. The embedding of such a torus in $\mathscr{B}$ looks like a cylinder or tube and can be continuously deformed to a closed curve in $\mathscr{B}$. In the extended zone in $\mathscr{E}$, the covering of such a torus is an infinite periodically deformed (''rippled'') cylinder corresponding to a certain vector $\eta_\alpha$ of the reciprocal lattice. Longitudinal cross sections of such a cylinder made by the plane $\Pi(\mathbf{B})$ yield trajectories open in $\mathscr{E}$. It is obvious that tori of this type can occur only if $\Pi(\mathbf{B})$ contains vectors of the reciprocal lattice and disappear after a variation in $\mathbf{B}$ violating this condition.

Another option is a ''membrane'' or an $M$-type torus, which is not homologous to zero in $\mathscr{B}$, i.e., it does not bound any region in $\mathscr{B}$. In the extended Brillouin zone, its covering is a periodically deformed integer plane (''rippled'' plane), whose intersection with the plane $\Pi(\mathbf{B})$ defines open trajectories of a general direction discussed above. All such planes corresponding to all tori of $M$-type are parallel in $\mathscr{E}$ since they do not intersect with each other, and their common homology class is the integer plane $\Gamma(\mathbf{B}_0)$ described above. The closure of an open trajectory lying on one of the $M$-type tori in $\mathscr{B}$ covers this torus, except its initial holes, with winding number $\alpha(\mathbf{B})$. Thus, such a trajectory can be treated as quasiperiodic, so any open trajectory for a direction of $\mathbf{B}$ in a stability zone is either periodic or quasiperiodic. The number of $M$-type tori on a given energy level is even, so their total homology class equals zero, and (for internal points of the stability zone) they are locally stable against variations in the direction of $\mathbf{B}$.

The topological scheme described above permits an easy investigation of general properties of solutions to Eq. (1) in the limit of strong uniform magnetic field $\mathbf{B}$ with a direction within a stability zone. Specifically, in the limit of high magnetic field, when the condition $\omega_B \gg \tau^{-1}$ is satisfied, the evolution of a one-particle distribution function is determined by the motion of electrons on the quasiclassical trajectories described above (note that the flux generated by the magnetic field in the quasimomentum space conserves the phase volume element $d^3p$). Moreover, in measuring parameters averaged over a certain time interval $T$ (this averaging will be denoted by $\langle\ldots\rangle_T$), we will obtain values calculated with the distribution function averaged over time on each trajectory (in the case when a trajectory forms an irrational winding on a torus with holes, this implies constancy on the torus). The projection of the time average $\langle\mathbf{v}_{gr}\rangle$ of the group velocity $\mathbf{v}_{gr}(\mathbf{p})$ on the $\Pi(\mathbf{B})$ plane, as directly follows from Eq. (3), is zero on closed trajectories and nonzero on open trajectories lying in rectilinear strips of finite widths and passing across them. In the second case $\langle\mathbf{v}_{gr}\rangle$ is perpendicular to the trajectory direction (although, generally speaking, $\langle\mathbf{v}_{gr}\rangle$ is not in the plane perpendicular to $\mathbf{B}$). The $\Pi(\mathbf{B})$ projection of the average velocity on trajectories not covered by strips of finite widths but with an asymptotic direction $\boldsymbol{\eta}$ (which is possible for directions of $\mathbf{B}$ with irrationality 2 outside all stability zones) is also nonzero and perpendicular to $\boldsymbol{\eta}$. The $\Pi(\mathbf{B})$ projection of a velocity averaged over any connected closure of trajectories without an asymptotic direction (when the direction of $\mathbf{B}$ has irrationality 3 and is outside all stability zones) is always zero, as follows from topological considerations.

Thus, it becomes obvious that any of the tori with holes described above (when $\mathbf{B}$ is in a stability zone) yields in the case of a distribution function constant over its trajectories a nonvanishing contribution to the component of the electron current $\mathbf{j}(\mathbf{x},\mathbf{p},t)$ perpendicular to $\mathbf{B}$. The condition $\mathbf{j}(\mathbf{x},\mathbf{p},t)\perp\boldsymbol{\eta}$ is always satisfied, which allows one to determine experimentally the asymptotic trajectory direction as a vector product $\mathbf{j}\times\mathbf{B}$. Note that for distribution functions constant on each energy level, for example, in the case of thermal equilibrium and in the absence of external force $\tilde{\mathbf{F}}(\mathbf{x},t)$ [Eq. (2)], the total contribution to $\mathbf{j}(\mathbf{x},\mathbf{p},t)$ of all such tori in any energy interval $(\epsilon,\epsilon+d\epsilon)$ is zero because all energy levels in $\mathscr{T}^3$ are homologous to zero [it is easy to prove that the average of the group velocity components $\mathbf{v}_{gr}(\mathbf{p})=\nabla\epsilon(\mathbf{p})$ over the region $\epsilon\leqslant\epsilon(\mathbf{p})\leqslant\epsilon+d\epsilon$ ($d\epsilon$ is small) reduces to values of basic 2-forms in two-dimensional cohomologies of the torus $\mathscr{T}^3$ $(dp_y\wedge dp_z, dp_z\wedge dp_x, dp_x\wedge dp_y)$ on the surface $\epsilon(\mathbf{p})=\epsilon$ with unknown factors]. For example, if $f_0(\mathbf{p})$ is a Fermi distribution of electrons in the absence of the external force $\tilde{\mathbf{F}}(\mathbf{x},t)$ and we are seeking stationary distributions in a weak uniform electric field $\mathbf{E}$,[2,4] the system can be described in terms of a linear response to an external perturbation, hence the conductivity tensor. In this case, when the electric field is switched on, electrons of energy levels on half of all tori acquire energy and move to higher levels ($\langle\mathbf{v}_{gr}\rangle\mathbf{E}>0$), and their places are occupied by particles from filled lower levels, whereas particles of the second half of all the tori lose energy ($\langle\mathbf{v}_{gr}\rangle\mathbf{E}<0$), and their states are occupied by particles from higher levels. As a result, a new distribution with a higher concentration of particles on the first half of the tori and a lower concentration on the second half is formed, so that the described process could be compensated for by the collision integral. The equilibrium current density $\mathbf{j}$ characterizes the response to an external field, and in the limit

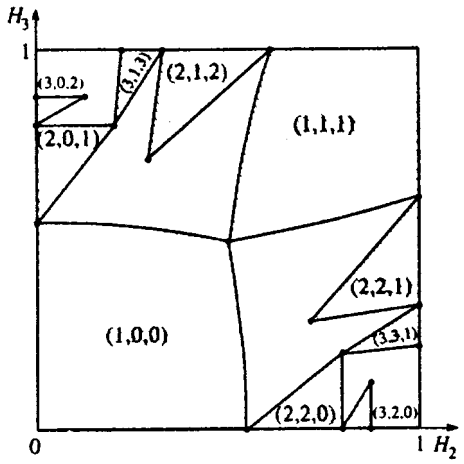



FIG. 2. Largest stability zones for the surface described by the equation $\cos 2\pi x_1+\cos 2\pi x_2+\cos 2\pi x_3=0$ calculated numerically. Digits in parentheses are indices of integer planes $\Gamma$.

$B\to\infty$, in which the distribution function is constant on the trajectories, the condition $\mathbf{j}\perp\boldsymbol{\eta}$ always holds, and the induced current is zero if $\mathbf{E}\|\boldsymbol{\eta}$ ($\langle\mathbf{v}_{gr}\rangle\mathbf{E}\equiv 0$).

The investigation of metal conductivity based on the topological scheme described in this paper was discussed in detail in Ref. 1. Here we only wish to emphasize that, in studies of metal conductivity, only energy levels near the Fermi surface are important, so only effects of trajectories on these levels can be observed in experiments. In particular, only those stability zones whose energy intervals $[\epsilon_1(\mathbf{B}),\epsilon(\mathbf{B})]$ (corresponding to open trajectories for a given direction of $\mathbf{B}$) contain the Fermi energy $\epsilon_F$ can be detected in experiment. The observed stability zones (in combination with the corresponding integer planes $\Gamma$) are topological characteristics of the Fermi surface, but not of the electronic spectrum branch as a whole. Their boundaries, generally speaking, do not coincide with the boundaries of the stability zone for a given dispersion relation, because $\epsilon_F$ can be outside the interval $[\epsilon_1(\mathbf{B}),\epsilon_2(\mathbf{B})]$ near the boundary of the stability zone, and the experimentally determined zones, generally speaking, do not form an everywhere dense set on the unit sphere $\mathscr{S}^2$, otherwise the situation can be described in terms of the topological scheme discussed in the paper. Figure 2 shows a fraction of the unit sphere with the largest zone for the surface

$$\cos 2\pi x_1+\cos 2\pi x_2+\cos 2\pi x_3=0$$

calculated using numerical techniques.

In order to study topological characteristics of a dispersion function as a whole, one should create within a single band nonequilibrium quasimomentum distributions of electrons of more general shapes. Such experiments can be conducted, for example, in semiconductors, where a certain number of electrons with a definite quasimomentum distribution can be supplied to an empty band, or electrons can be withdrawn from a full band to generated holes in it. Measurements of parameters described above, like $\langle\mathbf{j}(\mathbf{x})\rangle_T$ in the limit $B\to\infty$, would allow one to determine the characteristics of the energy spectrum discussed above, namely,

(1) stability zones $\Omega_\alpha$ with piecewise smooth boundaries and respective integer planes $\Gamma_\alpha$;

(2) integer directions $\boldsymbol{\eta}_\alpha$ defining directions of complimentary trajectories lying in strips aligned with $\boldsymbol{\eta}_\alpha$ and passing through them.

If it is possible to create distribution functions which are nonzero in narrow energy intervals $(\epsilon,\epsilon+d\epsilon)$, where $d\epsilon$ is small, then, in the case of a large stability zone where the difference $\epsilon_2(\mathbf{B})-\epsilon_1(\mathbf{B})$ is much greater than $d\epsilon$, one can also determine with an appropriate accuracy the functions $\epsilon_1(\mathbf{B})$ and $\epsilon_2(\mathbf{B})$, which determine the energy interval in which open trajectories exist for a given alignment of magnetic field.

In conclusion, the authors express their gratitude to S. P. Novikov for formulation of the problem and constant interest in this work.

One of the authors (I.A.D.) acknowledges the support by the Russian Fund for Fundamental Research (Project No. 96-01-0104), the other author (A.Ya.M.) is grateful for the support from the Russian Fund for Fundamental Research (Project No. 96-01-01623) and KFA Forschungszentrum Jülich GmbH.

[1)]E-mail: dynnikov@nw.math.msu.su
[2)]E-mail: maltsev@itp.ac.ru
[3)]Through the fault of the second co-author of Ref. 1, it contains an error, namely the statement that the number of directions $\boldsymbol{\eta}_\alpha$ is finite and no greater than the genus $g$ of the surface under discussion. This statement holds only for a fixed orientation of $\mathbf{B}$, but for variable directions of $\mathbf{B}$ the number of such vectors in $\mathscr{R}^3$ can be infinite. This error, however, does not affect the main conclusions of Ref. 1.

[1] S. P. Novikov and A. Ya. Maltsev, JETP Lett. **63**, 855 (1996).
[2] I. M. Lifshits, M. Ya. Azbel', and M. I. Kaganov, Zh. Éksp. Teor. Fiz. **31**, 63 (1956) [Sov. Phys. JETP **4**, 41 (1957)].
[3] I. M. Lifshits and V. G. Peschanskii, Zh. Eksp. Teor. Fiz. **35**, 1251 (1958) [Sov. Phys. JETP **8**, 875 (1959)].
[4] I. M. Lifshits and V. G. Peschanskii, Zh. Éksp. Teor. Fiz. **38**, 188 (1960) [Sov. Phys. JETP **11**, 137 (1960)].
[5] N. E. Alekseevskii and Yu. P. Gaidukov, Zh. Éksp. Teor. Fiz. **35**, 554 (1958) [Sov. Phys. **8**, 383 (1959)].
[6] N. E. Alekseevskii and Yu. P. Gaidukov, Zh. Éksp. Teor. Fiz. **36**, 447 (1959) [Sov. Phys. JETP **9**, 311 (1959)].
[7] N. E. Alekseevskii and Yu. P. Gaidukov, Zh. Éksp. Teor. Fiz. **37**, 672 (1959) [Sov. Phys. JETP **10**, 481 (1960)].
[8] Yu. P. Gaidukov, Zh. Éksp. Teor. Fiz. **37**, 1281 (1959) [Sov. Phys. JETP **10**, 913 (1960)].
[9] I. M. Lifshits, M. Ya. Azbel', and M. I. Kaganov, *Electron Theory of Metals*, Consultants Bureau, New York, (1973).
[10] S. P. Novikov, Usp. Mat. Nauk **37**, 3 (1982).
[11] S. P. Novikov, Trudy MIAN **166**, 201 (1984).
[12] S. P. Novikov, in *Proceedings of the Conf. on Topological Methods in Mathematics* dedicated to the 60th birthday of J. Milnor, June 15–22, 1991, S.U.N.Y. Stony Brook (1993).
[13] S. P. Novikov, in *Proceedings of Conf. on Geometry*, December 15–26, 1993, Tel Aviv University (1995).
[14] A. V. Zorich, Usp. Mat. Nauk **39**, 235 (1984).
[15] I. A. Dynnikov, Usp. Mat. Nauk **47**, 161 (1992).
[16] I. A. Dynnikov, Usp. Mat. Nauk **48**, 179 (1993).
[17] I. A. Dynnikov, Mat. zametki **53**, 57 (1993).
[18] I. A. Dynnikov, submitted to Advances in Soviet Math. (1997).
[19] I. A. Dynnikov, Surfaces in 3-Torus: Geometry of Plane Sections, Proceedings of ECM2, BuDA (1996).
[20] S. P. Tsarev, private communication (1992–93).

Translation was provided by the Russian Editorial office.

# Relativistic wave equation for the bound states of a system of interacting particles

A. I. Agafonov and É. A. Manykin

*Kurchatov Institute, 123182 Moscow, Russia*
(Submitted 24 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 50–62 (July 1997)

A method for obtaining the relativistic wave equation for the bound states of a system of interacting charged particles without consideration of spin is proposed. An expansion of the wave function of the system in a complete basis of orthonormal wave functions of vacuum states for each type of particle is used in this equation. It is shown that this equation contains two types of solutions for a proton + electron system. The first type corresponds to Bohr bound states. Exact expressions are obtained for the energy and Bohr radius of the ground state with consideration of the finite mass of the particles. An influence of the energy of translational motion of the system as a whole on the structure of the atomic levels in the laboratory frame is predicted. This effect is due to the finite value of $m/M$, and leads to removal of the degeneracy of the levels with respect to orbital angular momentum $l$, and partial removal of the degeneracy with respect to its projection. The second type of solution represents states of the system with binding energy $E_b=M+m-\sqrt{|M^2-m^2|}$ and an exponential wave function damping radius equal to the Compton wavelength of the electron. A complete description of this state requires consideration of the electronic vacuum polarization. © *1997 American Institute of Physics.* [S1063-7761(97)00507-6]

## 1. INTRODUCTION

A systematic description of the properties of relativistic quantum particles is possible only within quantum field theory. The relativistic effects are large when the energies of the particles are comparable to their rest energies. Particles (real or virtual) can be created at such energies; therefore, consideration of only a given original system of particles is incorrect in the general case.[1,2]

However, in some problems particle production can be disregarded, and familiar wave equations can be used for one particle (the single-particle approximation) or several given particles.[1] The calculations are based on relativistic generalizations of the Schrödinger equation for free particles: the Dirac equation for particles with spin 1/2, and the Klein–Gordon equation for particles with spin 0.

These equations contain a particle momentum operator. The general approach to describing particles in an external electromagnetic field is to replace that operator by a generalized momentum, which includes the field potentials $\mathbf{A}$ and $A_0$.[1] For example, the relativistic corrections to atomic energy levels, which determine their exact structure, are found in this manner.

We note that such an approach to problems of this type (for example, atomic spectra) has some logical loose ends and cannot be regarded as a systematic theory.

This paper proposes a method for obtaining the relativistic wave equation for a system of particles for problems in which the creation of real or virtual particles can be neglected. The method is based on the results of the quantum theory of scattering and the ideas which were used to describe deep impurity levels in doped insulators.[3] In this case only the free wave functions of all the types of particles in the system, the interaction potentials between the particles, and their statistics are important.

We shall not be concerned with spin effects here. In the final analysis, for the hydrogen atom they lead to relativistic corrections of order $\alpha^4$, where $\alpha$ is the fine structure constant. The purpose of our work is to obtain a relativistic equation for the bound states of a system of interacting particles, in which a complete basis of orthonormal vacuum wave functions is taken into account for each type of particle in the system. In the case of a pairwise interaction this equation should have definite permutational symmetry with respect to the parameters of any particle.

Below we derive and investigate an equation for the bound states of two interacting charged particles (a proton and an electron). We disregard the electronic vacuum polarization, which, as we know, distorts the Coulomb field in a region of the order of the Compton wavelength of the electron. The interaction potential $V(\mathbf{r})$ between the two particles can be represented in the form

$$V(\mathbf{r})=\begin{cases} -\dfrac{e^2}{r}, & r>\rho_p \\ V_1(\mathbf{r}) & r\leqslant\rho_p \end{cases}. \tag{1}$$

Here $\rho_p$ is a quantity of the order of the radius of the proton: $\simeq 0.8-0.9$ fm.[4] The form of $V_1(\mathbf{r})$ is unknown, but it is physically reasonable that the wave function of the electron should be influenced by the corpuscular properties of the proton in the region of space $r\leqslant\rho_p$. If the proton is treated as a rigid solid particle, the electron wave function for any bound state should vanish at $r\leqslant\rho_p$, since this region is occupied by the proton with its finite dimensions.

Of course, the radius of this region is very small in comparison with the characteristic length scales of the problem, viz., the Bohr radius and the Compton wavelength of the electron. We shall not take into account the perturbation of the wave functions due to the potential $V_1$ in the region

 1063-7761/97/070027-07$10.00  

$r \leqslant \rho_p$. In the differential equations for the systems discussed below, we seek a solution outside this small region, and then we shall go to the limit $r \to 0$ in that equation after allowing all the operators to act on the wave function. Then, the action of the operator $\varepsilon(\hat{\mathbf{p}})$ on the wave function is completely determined and unique.

We show that this equation contains two types of solutions. The first corresponds to Bohr bound states. It predicts that the energy of translational motion of the system as a whole on the structure of the atomic levels in the laboratory frame. This effect comes from the finite value of the ratio $m/M$ between the masses of the particles, and leads to removal of the degeneracy of the levels with respect to the orbital angular momentum $l$ and partial removal of the degeneracy with respect to its projection. We calculate the shifts of the $ns$ and $sp$ atomic levels at low values of the energy of translational motion of the center of inertia of the system $E^{(2)}(g) \ll mc^2 m/2M$.

The fact that a bound state (other than a Bohr state) appears in the proposed approach to the problem with an exponential wave function damping radius equal to the Compton wavelength of the electron and a binding energy $E_b = M + m - \sqrt{|M^2 - m^2|}$ is an unexpected result. The equation of state has a significantly non-Hamiltonian form, so that it does not make sense to separate the contributions from the potential and kinetic energies to the binding energy of a particle. For this state of the system the vacuum polarization will significantly distort the interaction energy, and the results must be regarded only as a prediction of the possibility of such a state. A complete description should include the state of the perturbed electronic vacuum.

The system of units in which $c = \hbar = 1$ is used below.

## 2. METHOD FOR DERIVING THE RELATIVISTIC EQUATION FOR BOUND STATES OF A SYSTEM OF INTERACTING PARTICLES

To obtain the relativistic wave equation for the bound states of a system of particles we use the expansion of the complete wave function of the discrete spectrum in a complete basis of orthonormal wave functions of the free states for each type of particle in the system. The basis for a particle of mass $m$ consists of free states with particle energies $\varepsilon = \pm\varepsilon(\mathbf{p})$, where $\varepsilon(\mathbf{p}) = (m^2 + \mathbf{p}^2)^{1/2}$. The form of wave functions of free particles is dictated by the uniformity of space and time, i.e., symmetry under any parallel displacement of the four-coordinate system. Then the free state of a particle with definite four-momentum $p^\nu = (\varepsilon, \mathbf{p})$ should be represented by the plane wave $\exp(-ipx)$. When the continuity equation ($\partial_\mu j^\mu = 0$, where $j^\mu$ is the current four-vector) is used, the normalized wave functions for the free-particle states have the form

$$\psi^{\pm}_{\mathbf{p},\sigma} = \sqrt{\frac{m}{\varepsilon(\mathbf{p})}}\, \exp(i(\mathbf{p}\cdot\mathbf{r} \pm \varepsilon(\mathbf{p})t))u_{p,\sigma}, \tag{2}$$

where $u_{p,\sigma}$ is an orthonormal spin function.

Henceforth we neglect the spin interaction. Then the spin part in Eq. (2) can be omitted.

Let $\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n)$ be the wave function of a bound state of a system of particles with rest masses $m_i$ and radius vectors $\mathbf{r}_i$. Then, according to the quantum theory of scattering, this wave function should satisfy the Lippman–Schwinger equation[5,6]

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_n; E) = \int \ldots \int \prod_{i=1}^{n} d\mathbf{r}_i' G^0(\mathbf{r}_1 - \mathbf{r}_1', \ldots, \mathbf{r}_n - \mathbf{r}_n'; E) \times \sum_{i>j} V(\mathbf{r}_i' - \mathbf{r}_j')\psi(\mathbf{r}_1', \mathbf{r}_2', \ldots, \mathbf{r}_n'; E), \tag{3}$$

where $V(\mathbf{r}_i - \mathbf{r}_j)$ is the interaction between particles $i$ and $j$, $E$ is the total energy of the system, and $G^0$ is the causal Green's function of the noninteracting system. To determine it we introduce the field operator of a system of noninteracting particles:

$$\mathbf{\Psi}_s = \prod_{i=1}^{n} \psi^{(i)}, \tag{4}$$

where

$$\psi^{(i)} = \sum_{\mathbf{p}} (a^{(+),i}_{\mathbf{p}} \psi^{+}_{\mathbf{p}}(\mathbf{r}_i, t) + a^{(-),i}_{\mathbf{p}} \psi^{-}_{\mathbf{p}}(\mathbf{r}_i, t)), \tag{5}$$

where $a^{(-),i}_{\mathbf{p}}$ and $a^{(+),i}_{\mathbf{p}}$ are the annihilation operators of the $i$th particle in the bands of free states with energies $\varepsilon = -\varepsilon(\mathbf{p})$ and $\varepsilon = +\varepsilon(\mathbf{p})$, respectively. The lower band is assumed to be filled. The commutation rules for these operators are assumed to be the same as for fermions.

The Green's function is defined as

$$G^0(\{\mathbf{r}_i\}, t; \{\mathbf{r}_i'\}, t') = \langle T(\Psi_s(\{\mathbf{r}_i\}, t)\Psi_s^{+}(\{\mathbf{r}_i'\}, t'))\rangle, \tag{6}$$

where $\langle \ldots \rangle$ denotes averaging over the vacuum ground state. Expression (6) is an $n$-particle two-time Green's function, which is used in the quantum theory of scattering.[6]

Substituting (2), (4), and (5) into (6) and calculating the Fourier component with respect to time, we obtain

$$G^0(\mathbf{r}_1 - \mathbf{r}_1', \ldots, \mathbf{r}_n - \mathbf{r}_n'; E) = \sum_{\mathbf{p}_1, \ldots, \mathbf{p}_n} \prod_{i=1}^{n} \frac{m_i}{\varepsilon_i(\mathbf{p}_i)} \times \left( \sum_{\substack{\{j_l = 1,2\} \\ l = 1, \ldots, n}} (E + (-1)^{j_1}\varepsilon_1(\mathbf{p}_1) + (-1)^{j_2}\varepsilon_2(\mathbf{p}_2) + \ldots + (-1)^{j_n}\varepsilon_n(\mathbf{p}_n))^{-1} \right) \times \exp\left( i \sum_{k=1}^{n} \mathbf{p}_k(\mathbf{r}_k - \mathbf{r}_k') \right). \tag{7}$$

With consideration of (7), in Eq. (3) it is convenient to transform to coordinates in which the radius vector $\mathbf{R}$ of the center of mass of the system can be separated out:

$$\mathbf{R} = \frac{\Sigma_i \mathbf{r}_i \varepsilon_i(\mathbf{p}_i)}{\Sigma_i \varepsilon_i(\mathbf{p}_i)}. \tag{8}$$

The choice of the remaining coordinates $\mathbf{x}_j$ ($j=1,\ldots,n-1$, $n\geqslant 2$) is dictated by convenience or the symmetry of the problem. The wave function of the system can then be written in the form

$$\psi(\mathbf{x}_1,\ldots,\mathbf{x}_{n-1};\mathbf{g})\exp(i\mathbf{g}\cdot\mathbf{R}),$$

where $\mathbf{g}$ is the momentum of the system as a whole in the laboratory frame. It specifies the relationship between the $\mathbf{p}_k$ in Eq. (7): $\mathbf{g}=\sum_{k=1}^{n}\mathbf{p}_k$. As a result, $(n-1)$-fold integration over the momenta $\mathbf{p}$ and coordinates $\mathbf{x}$ of the system remains in the Lippman–Schwinger equation for a system with $n\geqslant 2$.

Equation (3) can also be obtained directly from the integral equation of quantum scattering theory, which we write in the form

$$\psi(\{\mathbf{r}_i\};E)=\frac{1}{E-H_0+i\epsilon}\sum_{i>j}V(\mathbf{r}_i'-\mathbf{r}_j')\psi(\{\mathbf{r}_i'\};E),$$

where $H_0$ is the unperturbed Hamiltonian of the system, which, as was assumed in Ref. 6, is equal to the sum of the unperturbed Hamiltonians for each particle in the system. As in Eq. (3), the particular solution of the homogeneous equation is omitted, since $\psi$ refers to bound states of the system.[5,6]

Using the expansion of $\psi(\{\mathbf{r}_i\};E)$ in the complete basis of functions of the system $\mathbf{\Phi_P}$ consisting of the wave functions (2) for each particle

$$\psi=\sum_{\mathbf{P}}\mathbf{\Phi_P}(\mathbf{\Phi_P},\psi),$$

where $\mathbf{P}=\{\mathbf{p}_i\}$, and going to the limit $\epsilon\to+0$, we obtain Eqs. (3) and (7). Generally speaking, the form of the unperturbed Hamiltonian $H_0$ for the $i$th particle is not important to us, just the fact that $H_0\psi^{\pm}_{\mathbf{p},\sigma}=\pm\varepsilon(\mathbf{p})\psi^{\pm}_{\mathbf{p},\sigma}$.

### 3. ONE PARTICLE IN A COULOMB FIELD

From (7) for the Green's function we obtain

$$G^0(\mathbf{r}-\mathbf{r}';E)=\sum_{\mathbf{p}}G^0(\mathbf{p};E)\exp(i\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')),$$

where

$$G^0(\mathbf{p};E)=\frac{m}{\varepsilon(\mathbf{p})}\left(\frac{1}{E-\varepsilon(\mathbf{p})}+\frac{1}{E+\varepsilon(\mathbf{p})}\right)=\frac{2mE}{\varepsilon(\mathbf{p})}\left(\frac{1}{E^2-\varepsilon^2(\mathbf{p})}\right). \tag{9}$$

Taking Eq. (9) into account, we can easily show that the integral equation for the bound states

$$\psi(\mathbf{r};E)=\int d\mathbf{r}_1 G^0(\mathbf{r}-\mathbf{r}';E)V(\mathbf{r}_1)\psi(\mathbf{r}_1;E)$$

corresponds to the differential equation

$$\varepsilon(\hat{\mathbf{p}})(E^2-\varepsilon^2(\hat{\mathbf{p}}))\psi(\mathbf{r};E)=2mEV(r)\psi(\mathbf{r};E), \tag{10}$$

where $\hat{\mathbf{p}}=-i\nabla_{\mathbf{r}}$.

First and foremost, let us discuss the features distinguishing Greens' function (9) and Eq. (10) from the more familiar expressions. If we were to change the sign between the two pole functions in expression (9) (strictly speaking, it is determined by the commutation relations of the particle creation and annihilation operators), we would obtain the exact form of the Green's function for the Klein–Gordon equation:

$$G^0(p;\mathbf{E})=\frac{m}{\varepsilon(\mathbf{p})}\left(\frac{1}{E-\varepsilon(\mathbf{p})}-\frac{1}{E+\varepsilon(\mathbf{p})}\right)=\frac{2m}{E^2-\varepsilon^2(\mathbf{p})}.$$

The corresponding integral equation for the bound states would then be the Klein–Gordon equation in an external scalar potential:

$$(-\Delta_{\mathbf{r}}+2mV(r))\psi(\mathbf{r};E)=(E^2-m^2)\psi(\mathbf{r};E).$$

When we are dealing with shallow bound states $\sqrt{|E^2-m^2|}\ll m$, such a choice of the sign, of course, does not lead to fundamentally incorrect results: Bohr states with small relativistic corrections are obtained. However, if there are deep bound states with binding energies of the order of $m$, such a choice of sign would result in a complete loss of the information about them.

In principle, the same situation obtains in the Dirac equation. In momentum space the Green's function for the Dirac equation has the form[1]

$$G^0(\mathbf{p};E)=\frac{\gamma p+m}{2\varepsilon(\mathbf{p})}\left(\frac{1}{E-\varepsilon(\mathbf{p})+i0}-\frac{1}{E+\varepsilon(\mathbf{p})-i0}\right)=\frac{\gamma p+m}{E^2-\varepsilon^2(\mathbf{p})+i0}.$$

Here $\gamma$ denotes the Dirac matrices in the standard representation, and the four-vector $p=(E,\mathbf{p})$.

It is noteworthy that in contrast to (9), here the amplitudes of the simple poles for the two allowed regions of the free-particle spectra $E=\pm\varepsilon(\mathbf{p})$ have different signs. In this sense the structure of the Green's function for the Dirac equation is similar to a boson propagator and, in particular, to the Green's function for the Klein–Gordon equation. This analogy can also be traced in the expression for the Green's function of phonons in a solid, which has the form

$$D^0(\mathbf{k};\omega)=\frac{\omega_0(\mathbf{k})}{2}\left(\frac{1}{\omega-\omega_0(\mathbf{k})+i0}-\frac{1}{\omega+\omega_0(\mathbf{k})-i0}\right).$$

Here $\omega_0(\mathbf{k})$ is the dispersion law for phonons. Clearly, the structure of the Green's function for the Dirac and Klein–Gordon equations is similar to that for phonons, with the exception of the amplitudes of the simple poles. However, in the nonrelativistic limit, in which $\mathbf{p}\to 0$ and $\varepsilon(\mathbf{p})\to m$, we have

$$\gamma p+m=2m\begin{pmatrix}1&0\\0&0\end{pmatrix},$$

and the fermion Green's function for the Dirac equation transforms into the boson Green's function for the Klein–Gordon equation, i.e., here we are dealing only with Bohr states.

Note that the expression (9) that we have derived here corresponds completely to the familiar Lehmann spectral representation for the fermion Green's function[7] with real and positive-definite pole amplitudes. Of course, Bohr states are also realized in Eq. (10), which, as can easily be shown, reduces to the Schrödinger equation when $r \neq 0$, if the binding energy is small compared to the particle mass. Below we obtain a solution of this equation for wave functions with $s$ symmetry.

We seek a solution of Eq. (10) in the form $\psi = \exp(-r/\xi)$. Introducing the notation $\lambda = 1/m$ for the Compton wavelength of a particle, we obtain

$$(E^2-\lambda^{-2}+\xi^{-2})\left((1-\lambda^2\xi^{-2})^{1/2}+\frac{\lambda^2}{r\xi}(1-\lambda^2\xi^{-2})^{-1/2}\right) - \frac{2}{r\xi}(1-\lambda^2\xi^{-2})^{1/2} = 2E\left(-\frac{e^2}{r}\right). \tag{11}$$

Combining terms containing the same powers $r^j$ $(j=0,1)$, from (11) we find two algebraic equations in the unknowns $E$ and $\xi$:

$$(E^2-\lambda^{-2}+\xi^{-2})(1-\lambda^2\xi^{-2})^{1/2}=0, \tag{12}$$

$$(E^2-\lambda^{-2}+\xi^{-2})(1-\lambda^2\xi^{-2})^{-1/2}-2\lambda^{-2}(1-\lambda^2\xi^{-2})^{1/2} + 2e^2E\xi\lambda^{-2}=0. \tag{13}$$

Solution of (12) yields two values for the exponential wave function damping radius:

$$\xi_1=(\lambda^{-2}-E_1^2)^{-1/2}, \quad \xi_2=\lambda.$$

Substituting $\xi_1$ into (13), we obtain the energy of the Bohr $1s_{1/2}$ ground state for a fermion,

$$E_1 = m\sqrt{1-\alpha^2},$$

where $\alpha$ is the fine structure constant. Here $\xi_1$ is the Bohr radius.

Solving (13) for $E$, we easily obtain

$$\lim_{\xi\to\lambda}\frac{E^2(\xi)}{(1-\lambda^2\xi^{-2})^{1/2}}=0,$$

and when $\xi_2=\lambda$, the binding energy of the particle is

$$E_2=0.$$

Thus, this deep level is located in the middle of the energy gap of width $2m$ in the spectrum of free states of the particle.

We note here that Eq. (10) has a significantly non-Hamiltonian form. We have shown that this equation has two solutions with wave functions of $s$ symmetry. One of the solutions, $E_1$, corresponds to the Bohr $1s_{1/2}$ ground state for the electron. For such solutions and low binding energy compared to the particle mass, Eq. (10) can be reduced to Hamiltonian form (the Schrödinger equation). Then we can separate the contributions of the potential and kinetic energies to the binding energy of the particle. In the case of $E_2$, such a separation makes no sense.

In a certain sense the analogous problem of the bound states of Fermi particles in many-body systems also exists in the nonrelativistic energy range. In this context we can mention the problem of the deep levels in semiconductors, whose solution was the subject of a paper by Keldysh,[3] as well as the description of hydrogenic atoms in a low-temperature plasma, which was presented, for example, in papers by Kilimann *et al.* (see Ref. 8 and the literature cited therein). In contrast to the present work, in those studies the corresponding medium was a ''vacuum.'' The wave equation obtained in Ref. 3 for a deep impurity level in an insulator likewise does not have a Hamiltonian form. Only by making a number of simplifications can it be reduced to a form similar to the relativistic Dirac equation, which contains several additional terms, such as an ''attracting'' center, regardless of the sign of the carrier.

## 4. FINITE-MASS NUCLEUS

### 4.1. Equation for bound states

According to Eq. (7), we can represent the Green's function of a system of two particles in the form

$$G^0(\mathbf{r}_1-\mathbf{r}_1',\mathbf{r}_2-\mathbf{r}_2';E)=\sum_{\mathbf{p},\mathbf{q}}\frac{mM}{\varepsilon(\mathbf{p})\Sigma(\mathbf{q})}\left(\frac{1}{E-\Sigma(\mathbf{q})-\varepsilon(\mathbf{p})} + \frac{1}{E-\Sigma(\mathbf{q})+\varepsilon(\mathbf{p})} + \frac{1}{E+\Sigma(\mathbf{q})-\varepsilon(\mathbf{p})} + \frac{1}{E+\Sigma(\mathbf{q})+\varepsilon(\mathbf{p})}\right) \times\exp(i(\mathbf{p}\cdot(\mathbf{r}_1-\mathbf{r}_1')+\mathbf{q}\cdot(\mathbf{r}_2-\mathbf{r}_2'))), \tag{14}$$

where $M$ is the mass of the nucleus and $\Sigma(\mathbf{q})=(M^2+\mathbf{q}^2)^{1/2}$.

Here it is convenient to separate out the coordinates of the center of mass of the system:

$$\mathbf{r}=\mathbf{r}_1-\mathbf{r}_2, \quad \mathbf{R}=\frac{\mathbf{r}_1\varepsilon(\mathbf{p})+\mathbf{r}_2\Sigma(\mathbf{q})}{\varepsilon(\mathbf{p})+\Sigma(\mathbf{q})}.$$

We seek the wave function of the system in the form

$$\psi(\mathbf{r},\mathbf{R};\mathbf{g})=\chi(\mathbf{r};\mathbf{g})\exp(i\mathbf{g}\cdot\mathbf{R}),$$

Here $\mathbf{g}$ is the momentum of the system as a whole in the laboratory frame.

Then for the wave function $\chi(\mathbf{r};\mathbf{g})$ we obtain

$$\chi(\mathbf{r};\mathbf{g})=\int d\mathbf{r}_1\sum_{\mathbf{q}}\frac{Mm}{\Sigma(\mathbf{q})\varepsilon(\mathbf{q}-\mathbf{g})} \times\left(\sum\frac{1}{E\pm\Sigma(\mathbf{q})\pm\varepsilon(\mathbf{q}-\mathbf{g})}\right) \times\exp\left(i\left(\mathbf{q}-\mathbf{g}\,\frac{\Sigma(\mathbf{q})}{\Sigma(\mathbf{q})+\varepsilon(\mathbf{q}-\mathbf{g})}\right)\cdot(\mathbf{r}-\mathbf{r}_1)\right) \times V(\mathbf{r}_1)\chi(\mathbf{r}_1;\mathbf{g}). \tag{15}$$

To obtain the differential equation for the bound states of the system corresponding to integral equation (15), we must find the solution $\mathbf{q}(\mathbf{f},\mathbf{g})$ of the equation

$$\mathbf{f}=\mathbf{q}-\mathbf{g}\,\frac{\Sigma(\mathbf{q})}{\Sigma(\mathbf{q})+\varepsilon(\mathbf{q}-\mathbf{g})}. \tag{16}$$

If a solution of (16) in the form $\mathbf{q}(\mathbf{f},\mathbf{g})$ is found, the differential equation for the bound states of the system can easily be obtained from (15):

$$\left(\frac{d\mathbf{q}(\hat{\mathbf{f}},\mathbf{g})}{d\mathbf{f}}\right)^{-1}\Sigma(\mathbf{q})\varepsilon(\mathbf{q}-\mathbf{g})\times\left(\sum\frac{1}{E\pm\Sigma(\mathbf{q})\pm\varepsilon(\mathbf{q}-\mathbf{g})}\right)^{-1}\chi(\mathbf{r};\mathbf{g})=MmV(\mathbf{r})\chi(\mathbf{r};\mathbf{g}). \tag{17}$$

Here $\mathbf{q}=\mathbf{q}(\hat{\mathbf{f}},\mathbf{g})$ and $\hat{\mathbf{f}}=-i\nabla_{\mathbf{r}}$.

#### 4.2. Two types of states in the center-of-mass system

The radius vector $\mathbf{R}$ describes the motion of the system as a whole. Using this degree of freedom, we can obtain the effective mass of a system of two particles in any bound state. This will be done below, but we shall first ascertain the influence of the mass of the nucleus on the spectrum of bound states. For this purpose we consider a system of two bound particles at rest. Then $\mathbf{g}=0$ and $\mathbf{q}=\mathbf{f}$. Taking into account Eqs. (16) and (17), we obtain

$$\varepsilon(\hat{\mathbf{f}})\Sigma(\hat{\mathbf{f}})(E^4-2E^2(\varepsilon^2(\hat{\mathbf{f}})+\Sigma^2(\hat{\mathbf{f}}))+(\varepsilon^2(\hat{\mathbf{f}})-\Sigma^2(\hat{\mathbf{f}}))^2)\chi(\mathbf{r};E)=4mME(E^2-\varepsilon^2(\hat{\mathbf{f}})-\Sigma^2(\hat{\mathbf{f}}))V(\mathbf{r})\chi(\mathbf{r};E). \tag{18}$$

We find the solutions of (18) corresponding to a wave function with $s$ symmetry: $\psi=\exp(-r/\xi)$. Substituting this function into (18) and combining terms containing the same powers $r^j$ $(j=0,1)$, after some simple but cumbersome calculations, we find two algebraic equations for $E$ and $\xi$:

$$Z=E^4-2E^2(\lambda_M^{-2}+\lambda_m^{-2}-2\xi^{-2})+(\lambda_M^{-2}-\lambda_m^{-2})^2=0, \tag{19}$$

$$Z\xi^{-1}(1-\lambda_m^2\xi^{-2})^{1/2}(1-\lambda_M^2\xi^{-2})^{1/2}\times\left(\frac{\lambda_m^2}{1-\lambda_m^2\xi^{-2}}+\frac{\lambda_M^2}{1-\lambda_M^2\xi^{-2}}\right)-2^3E^2\xi^{-1}(1-\lambda_m^2\xi^{-2})^{1/2}\times(1-\lambda_M^2\xi^{-2})^{1/2}=-4e^2E(E^2-\lambda_M^{-2}-\lambda_m^{-2}+2\xi^{-2}). \tag{20}$$

Here $\lambda_m=1/m$ and $\lambda_M=1/M$ are the Compton wavelengths of the particles.

Equations (19) and (20) have two solutions. The first corresponds to the $s$ Bohr ground state, which is characterized by the radius

$$\xi_1=a_B\sqrt{\gamma(\gamma+\gamma^{-1}+2\sqrt{1-\alpha^2})} \tag{21}$$

and energy

$$E_1=M\sqrt{1-\frac{\gamma\alpha^2}{(\gamma+\gamma^{-1}+2\sqrt{1-\alpha^2}}}+m\sqrt{1-\frac{\gamma^{-1}\alpha^2}{(\gamma+\gamma^{-1}+2\sqrt{1-\alpha^2}}}, \tag{22}$$

where $a_B$ is the Bohr radius, and $\gamma=m/M$.

For the second solution of (19) and (20) we obtain

$$\xi_2=\max(\lambda_m,\lambda_M) \tag{23}$$

and

$$E_2=\sqrt{|M^2-m^2|}. \tag{24}$$

The binding energy of the system in the deep bound state can be determined from the last equation:

$$E_b=M+m-\sqrt{|M^2-m^2|}. \tag{25}$$

When $M\gg m$, from (25) we have

$$E_b=m\left(1+\frac{\gamma}{2}\right), \tag{26}$$

which in the limit $M\to\infty$ is consistent with the result in Sec. 3.

#### 4.3. Influence of the momentum of the system as a whole on the Bohr energy levels in the nonrelativistic case

Let the system as a whole move with a certain velocity $\ll c$. Let $E^{(2)}(g)=g^2/2(M+m)$ be the energy of translational motion of the system.

When $g=0$, the integral over $\mathbf{q}$ for the Bohr states is taken for characteristic momentum values $\simeq\alpha m$. Therefore, in (15) $\Sigma(\hat{\mathbf{q}})$ and $\varepsilon(\hat{\mathbf{q}})$ can be expanded into series in $\mathbf{q}$. The same thing can be done for $g\neq 0$, if

$$E^{(2)}(g)\ll m\frac{m}{2M}\simeq 140\ \text{eV}.$$

Thus, for the electronic subsystem the nonrelativistic case corresponds to rather low energies of translational motion of the system as a whole.

Here we confine our attention to energy contributions up to $\alpha^4$. Therefore, the solution of (16) can be written in the form

$$\mathbf{q}=\mathbf{f}+\frac{M}{M+m}\mathbf{g}+\frac{\hbar^2\mathbf{g}}{m(M+m)c^2}\left(-\frac{M-m}{M}f^2+\frac{m}{M+m}\mathbf{g}\cdot\mathbf{f}\right). \tag{27}$$

Using this solution, we can reduce Eq. (17) to an equation of the Schrödinger type. For this purpose, the ''resonance'' term $(E_g-(\Sigma(\mathbf{q})-\varepsilon(\mathbf{q}-\mathbf{g}))^{-1}$ in the sum in the parentheses in (17) is expanded up to fourth-order terms in $\mathbf{f}$ and $\mathbf{g}$. It is easily seen that in the three remaining terms of the sum, we must keep only the zeroth order of the expansion. This leads to the contribution $1/2\mu_2$, where $\mu_2=mM(M+m)/((M+m)^2+Mm)$. In the first three factors on the left-hand side of (17), the expansion must be restricted to second order in $\mathbf{f}$ and $\mathbf{g}$. As a result we obtain

$$\left(E(g)-E^{(3)}(g)-\frac{\hat{\mathbf{f}}^2}{2\mu}-\frac{\hat{\mathbf{f}}^4}{8\mu_1^3}-\hat{W}(\mathbf{g},\mathbf{f})\right)\chi(\mathbf{r};\mathbf{g})$$

$$= \left( 1 - \frac{3\hat{\mathbf{f}}^2}{4\mu_3^2} + \frac{E(g) - E^{(3)}(g)}{2\mu_2} \right) V(\mathbf{r})\chi(\mathbf{r};\mathbf{g}). \qquad (28)$$

Here $E^{(3)}(g)$ is the sum of the first three terms in the expansion of $\sqrt{(M+m)^2+g^2}$ with respect to $g^2$; the symbols $\mu = mM/(M+m)$, $\mu_1^3 = M^3m^3/(M^3+m^3)$, and $\mu_3^2 = m^2M^2/(M^2+Mm+m^2)$ are used, and the perturbation operator appearing when the ''resonance'' term in Eq. (17) is expanded, which is important for the following, is

$$\hat{W}(\mathbf{g},\mathbf{f}) = \frac{g^2}{2(M+m)} \frac{\Delta_{\mathbf{r}} - 2(\mathbf{n}_g \nabla_{\mathbf{r}})^2}{2Mm}, \qquad (29)$$

where $\mathbf{n}_g = \mathbf{g}/g$.

It is clear from (28) that the solutions correspond to Bohr states with allowance for relativistic corrections $\simeq \alpha^4$. As was noted above, for these corrections the corresponding matrix elements should be calculated as $\lim_{\delta\to 0}\int_\delta d\mathbf{r}$. When $M \to \infty$ and $g = 0$, for the Bohr ground state we obtain from (28) the first three terms in the expansion of $m\sqrt{1-\alpha^2}$ in $\alpha^2$.

The operator (29) represents the influence of the translational motion of the system as a whole on the structure of the atomic levels in the laboratory frame. This is a very unusual result, and it is a consequence of the finite value of the ratio $m/M$.

It is easy to show that for the $ns$ states, Eq. (28) leads to the level shifts

$$\Delta E_{ns} = -\frac{\alpha^2}{6n^2} \frac{m}{M} E^{(2)}(g). \qquad (30)$$

For the $2p$ states we have

$$\Delta E_{2p} = +\frac{\alpha^2}{40} \frac{m}{M} E^{(2)}(g) \qquad (31)$$

for the state with momentum projection $m = 0$, and

$$\Delta E_{2p} = -\frac{3\alpha^2}{40} \frac{m}{M} E^{(2)}(g) \qquad (32)$$

for states with momentum projection $m = \pm 1$.

Thus, removal of the degeneracy with respect to the orbital angular momentum $l$ and partial removal of the degeneracy with respect to its projection $m$ in terms of $|m|$ are predicted when the particle moves as a whole. Let us evaluate the shift in the nonrelativistic case. When $E^{(2)}(g) = 50$ eV, for the $1s$ state we obtain

$$\Delta E_{1s} \simeq 2 \times 10^{-3} \text{ cm}^{-1},$$

which is approximately 17 times less than the Lamb shift for the $1s$ state.

It is noteworthy that the influence of the translational motion of the system as a whole on the Bohr levels leads to a correction to the mass of the system, which depends on its state. For $ns$ states, from (30) we obtain

$$\Delta M_{ns} = \frac{\alpha^2}{6n^2} m.$$

We were unable to solve (17) for the deep bound state $E_2$ when $\mathbf{g} \neq 0$. Since the electron in this state should be regarded as a relativistic particle, its contribution to the effective mass of the system should significantly exceed the rest mass of the electron.

To conclude this section we point out the following situation. In this paper we derived the equations for bound states (10) and (17) on the basis of the Lippman–Schwinger equation in form (3). A similar result can be obtained if we start the derivation with the Dyson equation for the two-particle Green's function with the corresponding self-energy part. Such a procedure for obtaining the Bethe–Salpeter equation[9] for dispersive media was used in Refs. 8 and 10. In Ref. 10 the wave equation for the bound states of two charged nonrelativistic particles in a medium with a dielectric constant that includes frequency and spatial dispersion was obtained by a simple replacement of the vacuum photon propagator by the corresponding propagator in a dispersive medium in the Bethe–Salpeter equation (see, for example, Ref. 7). It is easy to show that this equation also contains the influence of the translational motion of the system as a whole on the spectrum of bound states in the laboratory frame. This corresponds qualitatively to the results of the treatment presented here for an essentially two-band model of the vacuum.

## 5. CONCLUSIONS

The solution corresponding to a bound state in a proton + electron system with an exponential wave function damping radius equal to the Compton wavelength of the electron is an unexpected result. We have taken an approach in which the creation of real or virtual particles was neglected and only a given initial system of particles was taken into account. However, for this state of the system, the vacuum polarization will significantly distort the interaction energy, and the results should be regarded only as a prediction of the possibility of such a state. Its complete description must include the state of the perturbed electronic vacuum.

However, we believe that consideration of the vacuum polarization will not qualitatively alter the result. This is supported, in particular, by the aforementioned theory of the bound states of charged particles in dispersive media. Although only the shielding of the Coulomb potential due to temporal and spatial dispersion was taken into account initially,[10] a subsequent more systematic treatment of the bound states of Fermi particles in many-body systems[8] shows that renormalization of the self-energy part does not have an appreciable influence on the spectrum of bound states, not only at a qualitative level, but also at a quantitative level in some cases.

We presume that vacuum polarization should provide for the stability of atomic systems and, in particular, the hydrogen atom H in the presence of the unoccupied deep level $E_2$. Then the matrix element of the radiative transition $E_{1s} \to E_2$ with emission of a photon has the form

$$\langle vac^* | \langle E_2 | \hat{O}(\mathbf{r}) | E_{1s} \rangle | vac^0 \rangle,$$

where $\hat{O}(\mathbf{r})$ is the transition operator acting on the electronic wave function, and $|vac^*\rangle$ and $|vac^0\rangle$ are the vacuum states for the $E_2$ and $E_{1s}$ states of the system, respectively. This

transition matrix element vanishes if the vacuum states are orthogonal, i.e., if $\langle vac^*|vac^0\rangle = 0$. We attribute the stability of atomic systems to just this situation.

Let us now discuss the $E_2$ state, in which we also include vacuum polarization. The only existing reality that might correspond to this situation is the neutron, which is known to be an unstable particle. In the free state it decays according to the reaction

$$n \to p + e + \tilde{\nu}_e$$

with a lifetime $\simeq 15$ min.

In the present approach, this decay can be attributed to the ionization of the $E_2$ state by high-energy particles in the Earth's inner radiation belt with an energy greater than the binding energy of the system. Then decay of the system should be accompanied by the appearance of a scattered particle, a free proton, and a knocked-out free electron, and by relaxation of the excited electronic vacuum state $|vac^*\rangle$. Since it follows from our treatment that neutron decay is regulated by the flux of the original near-terrestrial high-energy particles, the theory could be verified by the variation of the neutron lifetime in response to variation in the energy distribution and flux of particles with energies greater than the binding energy of the system.

We note that if it were possible to create a perturbed vacuum state $|vac^e\rangle$ at the space-time location of the nucleus of a hydrogen atom for which $\langle vac^*|vac^e\rangle \neq 0$, the following reaction with the release of energy should take place, in principle:

$$\mathbf{H} + |vac^e\rangle \to |E_2\rangle + h\nu + |vac^s\rangle,$$

where $\mathbf{H}$ is a hydrogen atom, $h\nu$ is a $\gamma$ ray with energy of order $m$, and $|vac^s\rangle$ is the excess excitation of the electronic vacuum.


[1] V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Quantum Electrodynamics*, Pergamon Press, Oxford (1982).
[2] N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, 2nd Am. ed., Wiley, New York (1980).
[3] L. V. Keldysh, Zh. Éksp. Teor. Fiz. **45**, 364 (1963) [Sov. Phys. JETP **18**, 253 (1964)].
[4] H. Frauenfelder and E. M. Henley, *Subatomic Physics*, Prentice-Hall, Englewood Cliffs, N.J. (1974).
[5] J. M. Ziman, *Elements of Advanced Quantum Theory*, Cambridge University Press, New York (1969).
[6] T.-Y. Wu and T. Ohmura, *Quantum Theory of Scattering*, Prentice-Hall, London (1962).
[7] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Physics*, Prentice-Hall, Englewood Cliffs, N.J. (1963).
[8] K. Kilimann, D. Kremp, and G. Ropke, Teor. Mat. Fiz. **55**, 448 (1983).
[9] E. E. Salpeter and H. A. Bethe, Phys. Rev. **84**, 1232 (1951).
[10] É. A. Manykin, M. I. Ozhovan, and P. P. Poluéktov, Teor. Mat. Fiz. **49**, 289 (1981).


Translated by P. Shelnitz

# Correlation echoes in the stochastic excitation of inhomogeneously broadened two-level systems

S. A. Baruzdin

*St. Petersburg State Electrical Engineering University, 197376 St. Petersburg, Russia*
(Submitted 30 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 63–77 (July 1997)

The third-order cross-correlations between a free induction signal of an inhomogeneously broadened two-level system and white Gaussian noise exciting this system are studied. The temporal properties of the third-order cross-correlation functions are found to correspond to the characteristics of ordinary two- and three-pulse spin and light (photon) echoes excited by determinate radio pulses. The nonlinear properties of correlation echoes are studied as functions of the noise pulse parameters. It is established that the correlation echo amplitude is determined not only by the noise pulse parameters but also by the position on the time axis of the noise counts that form the given type of echo. Finally, the behavior of the spin and light correlation echoes in the appropriate ranges is discussed.  [S1063-7761(97)00607-0]

## 1. INTRODUCTION

The use of stochastic excitation in NMR spectroscopy was first proposed by Ernst[1] (see also Ref. 2) to lower the excitation power and to ensure that the excitation covers a wide band. The method became known as the stochastic resonance method. Its use also allows various multidimensional spectra to be obtained, while lowering the excitation power facilitates solving the problem of the receiver's ''dead time.'' In addition to NMR, stochastic excitation is used in EPR studies and in optics.[3,4]

Realizing this method requires extensive data processing, however. For instance, to measure the pulse characteristic of a linear system, $h(t)$, related to the frequency transfer coefficient through a Fourier transformation, the system is excited not by a delta-function pulse but by white noise $x(t)$ with a spectral power density $N_0$ and a correlation function

$$\langle x(t_1)x(t_2)\rangle = N_0\delta(t_1-t_2).$$

What is determined here is the cross-correlation function between the output process of the system,

$$y(t)=\int_{-\infty}^{\infty} x(\xi)h(t-\xi)d\xi,$$

and the input process delayed by time $\sigma$:

$$C_1(\sigma)=\langle y(t)x(t-\sigma)\rangle = N_0 h(\sigma). \tag{1}$$

Thus, short coherent pulses are replaced in the experiment by band white noise, which allows the excitation power to be several orders of magnitude lower. Here, however, instead of forming $h(t)$ directly in real time, according to (1) we must, first, statistically average the products $y(t)x(t-\sigma)$ and, second, measure the function $h(t)$ at different points, each time changing the delay time $\sigma$. Simultaneous measurement of several points of the pulse characteristic in parallel data processing is also possible.

In addition to studies of the characteristics of linear systems, white Gaussian noise used as excitation provides a convenient model for consistent studies of properties of nonlinear systems.[1–5] In this case instead of the cross-correlation function (1) one studies higher-order cross-correlation functions, which can be related to the pulse characteristics of the respective order.

Theoretical studies of the correlation properties of the responses of two-level spin systems described by Bloch equations and excited by white Gaussian noise were carried out by Ernst[1], Knight, and Keiser.[6] Here the excitation was continuous, and the spin system was in a stable stationary state.

Pulsed stochastic excitation is used alongside continuous excitation. In Ref. 7 the statistical properties of the free induction signal of an inhomogeneously broadened spin system excited by finite samples of white Gaussian noise were investigated. It was found, among other things, that in this case the free induction signal is a nonstationary Gaussian process with zero mathematical expectation, which carries no information about the system under investigation. Note that in the optical range, optical free induction serves as the free induction signal.[8]

Nevertheless, stochastic excitation can form responses whose mathematical expectation is nonzero. For instance, in experiments on excitation of stimulated light echo, the first and second delta-function coherent excitation pulses were replaced by longer noise pulses.[4,9] These pulses were formed from the incoherent radiation of a pulsed laser whose beam was split into two beams time-delayed with respect to each other. Thus, the noise counts of the first and second pulses separated by the delay time were coherent, which ensured their coherent addition in the stimulated light echo, and this became known as an ''echo with internal coherence,'' which means coherence of two noise pulses within each excitation period. By measuring the mean value of the amplitude of this echo as a function of the delay time between excitation pulses it is possible to extract information about the times of longitudinal ($T_1$) and transverse ($T_2$) relaxations of the system. Stimulated spin echo can be excited in a similar way. Its nonlinear properties were studied in Ref. 10.

In the experiment described in Ref. 3, Paff and Blumich

studied the stochastic analog of stimulated spin echo excited by quasicontinuous white Gaussian noise $x(t)$. The transmitter and receiver were alternately gated after isolating the response by separating it in time from the excitation process. A correlator was used to calculate the third-order cross-correlation function between the response $y(t)$ of the nonlinear spin system and the excitation:

$$C_3(\sigma_1>\sigma_2>\sigma_3)=\int_{-\infty}^{\infty} y(t)x(t-\sigma_1)x(t-\sigma_2)x(t-\sigma_3)dt, \tag{2}$$

where $\sigma_i$ is the delay time in the correlator.

The spin system was in a stable state, in which both $x(t)$ and $y(t)$ were interpreted as stationary processes, which actually allows the cross-correlation functions (2) to be calculated from the time average of the product in the integrand.[1)] For this reason $C_3(\sigma_1,\sigma_2,\sigma_3)$ is independent of the time $t$.

The function $C_3(\sigma_1,\sigma_2,\sigma_3)$ was successively measured by the correlator point by point. Here the delay times were chosen by the condition that

$$\sigma_1-\sigma_2=t_1=\text{const}, \quad \sigma_2-\sigma_3=t_2=\text{const},$$

and the delay time $\sigma_3=t_3$ was varied. The maximum of $C_3(\sigma_1,\sigma_2,\sigma_3)$ was found to be at $\sigma_3=\sigma_1-\sigma_2$.

This result can be compared with the time at which a stimulated echo excited by three ordinary delta-function pulses appears. If the distance between the first two excitation pulses is $t_1=\sigma_1-\sigma_2$ and the distance between the second and third pulses is $t_2=\sigma_2-\sigma_3$, a stimulated echo is formed after a time $t_3=\sigma_3=\sigma_1-\sigma_2$ following the third pulse has elapsed. Thus, the maximum in $C_3(\sigma_1,\sigma_2,\sigma_3)$ corresponds to the time when stimulated echo is formed, and the third-order cross-correlation function (2) resembles a response to three delta-function pulses, i.e., is the stochastic analog of stimulated echo. At the same time, the use of stochastic excitation made it possible to lower the excitation power in this experiment by a factor of 1000 and easily solved the dead-time problem of the receiver.

The measured cross-correlation functions were distorted by system noise caused by the finite time of integration in the correlator in the averaging process. Here the signal-to-noise ratio increases in proportion to the square root of the integration time.

The possibility and advisability of extending these experiments to the optical and EPR ranges are also noted.

The present work studies the nonlinear properties of correlation echoes, i.e., responses induced in inhomogeneously broadened two-level spin and optical systems by finite samples of a band white Gaussian noise of length $\tau$. These echoes can be isolated from the free induction signal in calculating third-order cross-correlation functions.

## 2. THE EQUATION OF MOTION OF THE STATE VECTOR OF THE SYSTEM

If the length $\tau$ of a noise pulse is assumed to be much smaller than $T_1$ and $T_2$, the effect of relaxation processes can be ignored. We will also assume that the period within which the excitation pulses act on the system is much longer than $T_1$, so that by the time the next excitation pulse has arrived the system has returned to its state of thermodynamic equilibrium.

Let us first examine an inhomogeneously broadened two-level spin system. Bearing in mind the above, we can write the equation of motion of the magnetization vector **M** in a magnetic field with induction **B** as follows:[2]

$$\frac{d\mathbf{M}}{dt}=\gamma\mathbf{M}\times\mathbf{B}, \tag{3}$$

where $\gamma$ is the gyromagnetic ratio.

Suppose that the spin system placed in a nonuniform magnetic field with induction $B_z\mathbf{e}_z$ is subjected, on the interval $|t|\leqslant\tau/2$, to a circularly polarized magnetic field

$$\mathbf{B}_1(t)=B(t)\cos[\omega_0 t+\varphi(t)]\mathbf{e}_x-B(t)\sin[\omega_0 t+\varphi(t)]\mathbf{e}_y,$$

where $B(t)$ and $\varphi(t)$ are functions describing the variations of the amplitude and phase of the magnetic induction, and $\omega_0$ is the vibration frequency coinciding with the central frequency of the inhomogeneously broadened absorption line.

In the system of coordinates rotating with a frequency $\omega_0$ about the longitudinal axis the equation of motion of the magnetization vector is[7]

$$\frac{dm_+}{dt}=\left(\frac{dm_-}{dt}\right)^*=i\Omega m_+-i\widetilde{R}(t)m_z,$$

$$\frac{dm_z}{dt}=-\frac{i}{2}R^*(t)m_+ +\frac{i}{2}\widetilde{R}(t)m_-, \tag{4}$$

where $m_+=(m_-)^*$ are the complex-valued transverse components of the magnetization vector, $m_z$ is the longitudinal component,

$$\widetilde{R}(t)=\gamma\widetilde{B}(t)=\gamma B(t)\cos\varphi(t)-i\gamma B(t)\sin\varphi(t),$$

$$\Omega=\gamma B_z-\omega_0=\omega-\omega_0.$$

In the corresponding two-level systems of the optical range the magnetic dipole transitions are replaced by electric dipole transitions. The equation of motion (3) now corresponds to the equation of motion of a pseudoelectric dipole **P** in a pseudoelectric field **E** (see Refs. 8, 11 and 12):

$$\frac{d\mathbf{P}}{dt}=\gamma_e\mathbf{P}\times\mathbf{E}, \tag{5}$$

where $\gamma_e$ is the gyroelectric ratio.

The analogy between Eqs. (3) and (5) are similar can be explained by employing a common vector model used in analyzing free induction signals and spin and light echoes in inhomogeneously broadened two-level systems.[8,13,14]

Thus, Eq. (5) can be written in a rotating coordinate system in the form (4) if in the latter we replace the components $m_\pm$ and $m_z$ with the corresponding components $p_\pm$ and $p_z$ of the pseudopolarization vector and if $\widetilde{R}(t)$ is interpreted as the complex-valued envelope of the exciting process and is linked to the pseudoelectric field $\widetilde{E}(t)$ by the relationship $\widetilde{R}(t)=\gamma_e\widetilde{E}(t)$ (see Ref. 15).

Let us introduce a new state vector of the system, **Y**, whose components are related to those of the vectors introduced earlier in the following manner:

$$y_{\pm}(t)=[p_{\pm}(t,\Omega)\ \text{ or }\ m_{\pm}(t,\Omega)]\exp[\pm i\Omega(t-t_0)],$$

$$y_z(t)=p_z(t,\Omega)\ \text{ or }\ m_z(t,\Omega), \tag{6}$$

where $t_0=-\tau/2$. Introduction of these components is equivalent to going over to a system of coordinates rotating with a frequency $\omega=\omega_0+\Omega$ about the longitudinal axis.

In terms of the new variables the equation of motion of the state vector of the system, (4), can be written in matrix form:

$$\frac{d\mathbf{Y}}{dt}=\mathbf{Q}(t)\mathbf{Y}, \tag{7}$$

$$\mathbf{Y}=\begin{pmatrix} y_+ \\ y_- \\ y_z \end{pmatrix},$$

$$\mathbf{Q}(t)=\begin{pmatrix} 0 & 0 & -\tilde{s}(t) \\ 0 & 0 & -s^*(t) \\ s^*(t)/2 & \tilde{s}(t)/2 & 0 \end{pmatrix}, \tag{8}$$

$$\tilde{s}(t)=i\tilde{R}(t)\exp[-i\Omega(t+\tau/2)]. \tag{9}$$

Prior to excitation the system was in an equilibrium state, in which the components of the vector $\mathbf{Y}$ are determined by the initial conditions

$$y_{\pm}\left(-\frac{\tau}{2}\right)=0,\quad y_z\left(-\frac{\tau}{2}\right)=Y_0.$$

Here $Y_0$ is the equilibrium value of the magnetization ($M_0$) or pseudopolarization ($P_0$).

For further analysis it is convenient to introduce two matrices:

$$\mathbf{Q}_1=\begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 1/2 & 0 \end{pmatrix},\quad \mathbf{Q}_2=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 1/2 & 0 & 0 \end{pmatrix}. \tag{10}$$

As a result, the equation of motion (7) of the state vector of the system can be written as follows:[16]

$$\frac{d\mathbf{Y}}{dt}=\tilde{s}(t)\mathbf{Q}_1\mathbf{Y}+s^*(t)\mathbf{Q}_2\mathbf{Y}. \tag{11}$$

Equation (11) can be integrated in two symmetric forms:

$$\mathbf{Y}(t)=\exp\left[\mathbf{Q}_1\int_{t_0}^{t}\tilde{s}(t')dt'\right]\mathbf{Y}(t_0)+\int_{t_0}^{t}\exp\left[\mathbf{Q}_1\int_{\xi}^{t}\tilde{s}(t')dt'\right]s^*(\xi)\mathbf{Q}_2\mathbf{Y}(\xi)d\xi, \tag{12}$$

$$\mathbf{Y}(t)=\exp\left[\mathbf{Q}_2\int_{t_0}^{t}s^*(t')dt'\right]\mathbf{Y}(t_0)+\int_{t_0}^{t}\exp\left[\mathbf{Q}_2\int_{\xi}^{t}s^*(t')dt'\right]\tilde{s}(\xi)\mathbf{Q}_1\mathbf{Y}(\xi)d\xi. \tag{13}$$

The validity of Eqs. (12) and (13) can be verified by differentiation, and the result is Eq. (11).

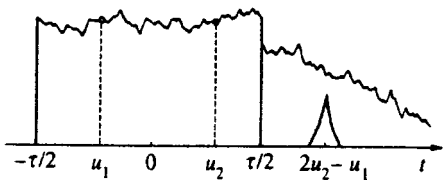



FIG. 1. Time diagram of excitation of two-pulse correlation echo from two noise counts at times $u_1$ and $u_2$.

### 3. SPIN AND LIGHT CORRELATION ECHOES

To analyze the nonlinear properties of the two-pulse correlation echo formed by the noise counts $u_1$ and $u_2$, we must determine the third-order cross-correlation function between the free induction signal and the excitation. We start by defining the corresponding cross-correlation function for a single spectral component of the nonstationary response of an inhomogeneously broadened system,

$$\mathbf{C}_e(t)=\langle\mathbf{\Phi}(t)\rangle=\langle\tilde{s}(u_1)[s^*(u_2)]^2\mathbf{Y}(t)\rangle,$$

$$2u_2-u_1>\frac{\tau}{2},\quad t\geqslant\frac{\tau}{2}. \tag{14}$$

The time diagram of the envelopes of the excitation pulse, of the response of the system in the form of a free induction signal, and of the two-pulse correlation echo separated from the free induction signal is depicted in Fig. 1.

We assume that the system is excited by band white noise with a Gaussian distribution and a zero average, with the bandwidth of the noise being much larger than the width of the excited inhomogeneous line. In this case the correlation functions of the complex-valued amplitude of the noise can be approximated as follows:

$$\langle\tilde{R}(t)R^*(u)\rangle=N_0\delta(t-u),$$

$$\langle\tilde{R}(t)\tilde{R}(u)\rangle=0,\quad |t,u|<\frac{\tau}{2}, \tag{15}$$

where $N_0$ is the spectral power density of the complex-valued Gaussian process $\tilde{R}(t)$.

To calculate the cross-correlation function (14) we multiply Eq. (12) by $\tilde{s}(u_1)[s^*(u_2)]^2$ and average, assuming that $t>t_0=u_2$. The result is

$$\mathbf{C}_e(t)=\mathbf{D}_1+\mathbf{D}_2,$$

$$\mathbf{D}_1=\left\langle\exp\left[\mathbf{Q}_1\int_{u_2}^{t}\tilde{s}(t')dt'\right]\mathbf{\Phi}(u_2)\right\rangle,$$

$$\mathbf{D}_2=\left\langle\int_{u_2}^{t}\exp\left[\mathbf{Q}_1\int_{\xi}^{t}\tilde{s}(t')dt'\right]s^*(\xi)\mathbf{Q}_2\mathbf{\Phi}(\xi)d\xi\right\rangle. \tag{16}$$

In averaging we allow for the fact that the response of a physically realizable system, $\mathbf{Y}(t)$, is independent of the excitation $\tilde{s}(\xi)$ and the functions of this process for $\xi>t$ in view of the causality principle, and the fact that the moments of a Gaussian process are factorable.[17] The vector $\mathbf{Y}$, which enters into the expressions for $\mathbf{D}_1$ and $\mathbf{D}_2$, can be determined by solving Eq. (7) by the method of successive approximations:[7]

$$\mathbf{Y}(t)=\left[\mathbf{I}+\int_{-\tau/2}^{t} d\xi_1 \mathbf{Q}(\xi_1) + \int_{-\tau/2}^{t} d\xi_1 \int_{-\tau/2}^{\xi_1} d\xi_2 \mathbf{Q}(\xi_1)\mathbf{Q}(\xi_2)+\cdots\right]\mathbf{Y}\left(-\frac{\tau}{2}\right), \tag{17}$$

where $\mathbf{I}$ is the identity matrix. According to Eqs. (17) and (8), the components of $\mathbf{Y}$ are given by a series whose terms can be expressed in terms of the products of the exciting process, which makes factorization of $\mathbf{D}_1$ and $\mathbf{D}_2$ possible. Here we can write

$$\mathbf{D}_1=\left\langle \exp\left[\mathbf{Q}_1\int_{u_2}^{t}\tilde{s}(t')dt'\right]\right\rangle \mathbf{C}_e(u_2) + \left\langle [s^*(u_2)]^2 \exp\left[\mathbf{Q}_1\int_{u_2}^{t}\tilde{s}(t')t'\right]\right\rangle \langle \tilde{s}(u_1)\mathbf{Y}(u_2)\rangle. \tag{18}$$

The cofactors containing matrix exponentials in (18) can be averaged by expanding them in series and allowing for the fact that $\tilde{s}(t)$ is a complex-valued Gaussian process defined by (9) and (15):

$$\begin{aligned}\left\langle [s^*(u_2)]^2 \exp\left[\mathbf{Q}_1\int_{u_2}^{t}\tilde{s}(t')dt'\right]\right\rangle &= \left\langle [s^*(u_2)]^2\left\{\mathbf{I} + \mathbf{Q}_1\int_{u_2}^{t}\tilde{s}(t')dt' + \frac{1}{2!}\left[\mathbf{Q}_1\int_{u_2}^{t}\tilde{s}(t')dt'\right]^2+\cdots\right\}\right\rangle \\ &= \frac{1}{2!}\mathbf{Q}_1^2\int_{u_2}^{t}dt'\int_{u_2}^{t}dt''\langle \tilde{s}(t')s(t'')[s^*(u_2)]^2\rangle \\ &= \frac{N_0^2\mathbf{Q}_1^2}{2}\int_{u_2}^{t}dt'\int_{u_2}^{t}dt''2\delta(t'-u_2)\,\delta(t''-u_2) \\ &= \frac{N_0^2\mathbf{Q}_1^2}{4}.\end{aligned} \tag{19}$$

If in (19) we put $s^*(u_2)=1$, we find that

$$\left\langle \exp\left[\mathbf{Q}_1\int_{u_2}^{t}\tilde{s}(t')dt'\right]\right\rangle=\mathbf{I}. \tag{20}$$

Plugging (19) and (20) into (18) yields

$$\mathbf{D}_1=\mathbf{C}_e(u_2)+\frac{N_0^2\mathbf{Q}_1^2}{4}\langle \tilde{s}(u_1)\mathbf{Y}(u_2)\rangle. \tag{21}$$

The expression describing $\mathbf{D}_2$ can be transformed in the following manner:

$$\mathbf{D}_2=\int_{u_2}^{t}\left\langle \exp\left[\mathbf{Q}_1\int_{\xi}^{t}\tilde{s}(t')dt'\right]\right\rangle \mathbf{Q}_2\langle s^*(\xi)\mathbf{\Phi}(\xi)\rangle d\xi + \int_{u_2}^{t}\left\langle \exp\left[\mathbf{Q}_1\int_{\xi}^{t}\tilde{s}(t')dt'\right]s^*(\xi)\right\rangle \mathbf{Q}_2\mathbf{C}_e(\xi)d\xi. \tag{22}$$

Averaging the cofactors containing matrix exponentials in (22) yields

$$\mathbf{D}_2=\mathbf{Q}_2\int_{u_2}^{t}\mathbf{C}_e(\xi)d\xi+\frac{N_0\mathbf{Q}_1\mathbf{Q}_2}{2}\int_{u_2}^{t}\mathbf{C}_e(\xi)d\xi. \tag{23}$$

Now we determine the cross-correlation function (14) taking Eq. (13) as the basis:

$$\begin{aligned}&\mathbf{C}_e(t)=\mathbf{D}_3+\mathbf{D}_4,\\ &\mathbf{D}_3=\left\langle \exp\left[\mathbf{Q}_2\int_{u_2}^{t}s^*(t')dt'\right]\mathbf{\Phi}(u_2)\right\rangle=\mathbf{C}_e(u_2),\\ &\mathbf{D}_4=\int_{u_2}^{t}\left\langle \exp\left[\mathbf{Q}_2\int_{\xi}^{t}s^*(t')dt'\right]\tilde{s}(\xi)\mathbf{Q}_1\mathbf{\Phi}(\xi)\right\rangle d\xi \\ &\quad=\mathbf{Q}_1\int_{u_2}^{t}\langle \tilde{s}(\xi)\mathbf{\Phi}(\xi)\rangle d\xi+\frac{N_0\mathbf{Q}_1\mathbf{Q}_2}{2}\int_{u_2}^{t}\mathbf{C}_e(\xi)d\xi.\end{aligned} \tag{24}$$

If we integrate Eq. (11) and then multiply the result by $\tilde{s}(u_1)[s^*(u_1)]^2$ and average, we find that the same cross-correlation function can be written as

$$\mathbf{C}_e(t)=\mathbf{C}_e(u_2)+\mathbf{Q}_1\int_{u_2}^{t}\langle \tilde{s}(\xi)\mathbf{\Phi}(\xi)\rangle d\xi + \mathbf{Q}_2\int_{u_2}^{t}\langle \tilde{s}(\xi)\mathbf{\Phi}(\xi)\rangle d\xi. \tag{25}$$

Adding (16) to (24) and subtracting (25) from the sum, we get

$$\mathbf{C}_e(t)=\frac{N_0^2\mathbf{Q}_1^2}{4}\langle \tilde{s}(u_1)\mathbf{Y}(u_2)\rangle+\mathbf{C}_e(u_2) + N_0\mathbf{L}\int_{u_2}^{t}\mathbf{C}_e(\xi)d\xi, \tag{26}$$

$$\mathbf{L}=\frac{\mathbf{Q}_1\mathbf{Q}_2+\mathbf{Q}_2\mathbf{Q}_1}{2}. \tag{27}$$

Differentiation of (26) yields the equation

$$\frac{d\mathbf{C}_e}{dt}=N_0\mathbf{L}\mathbf{C}_e(t),$$

whose solution is

$$\mathbf{C}_e(t)=\exp[N_0\mathbf{L}(t-u_2)]\mathbf{C}_e(u_2),\quad t\geqslant u_2. \tag{28}$$

To determine the correlation moment $\mathbf{C}_e(u_2)$ in (28), we set $t_0=u_1<t<u_2$ in Eq. (13), multiply the equation by $\tilde{s}(u_1)[s^*(t)]^n$ ($n=1$ or 2), and average:

$$\begin{aligned}&\langle \tilde{s}(u_1)[s^*(t)]^n\mathbf{Y}(t)\rangle \\ &=\left\langle \tilde{s}(u_1)[s^*(t)]^n\exp\left[\mathbf{Q}_2\int_{u_1}^{t}s^*(t')dt'\right]\mathbf{Y}(u_1)\right\rangle \\ &+\left\langle \tilde{s}(u_1)[s^*(t)\right]^n\int_{u_1}^{t}\exp\left[\mathbf{Q}_2\int_{\xi}^{t}s^*(t')dt'\right] \\ &\times\tilde{s}(\xi)\mathbf{Q}_1\mathbf{Y}(\xi)d\xi.\end{aligned} \tag{29}$$

At $n=2$ the first term on the right-hand side of Eq. (29) vanishes and the second can be written as

$$\mathbf{C}_e(t)=2\int_{u_1}^{t}\langle s^*(t)\tilde{s}(\xi)\rangle \times\left\langle \exp\left[\mathbf{Q}_2\int_{\xi}^{t}s^*(t')dt'\right]\tilde{s}(u_1)s^*(t)\mathbf{Q}_1\mathbf{Y}(\xi)\right\rangle d\xi.$$

If we allow for Eqs. (9) and (15), we get

$$\mathbf{C}_e(t)=N_0\mathbf{Q}_1\langle\tilde{s}(u_1)s^*(t)\mathbf{Y}(t)\rangle. \tag{30}$$

To find $\langle\tilde{s}(u_1)s^*(t)\mathbf{Y}(t)\rangle$, we set $n=1$ in (29). Then the first term on the right-hand side of Eq. (29) vanishes and

$$\langle\tilde{s}(u_1)s^*(t)\mathbf{Y}(t)\rangle=\frac{N_0\mathbf{Q}_1}{2}\langle\tilde{s}(u_1)\mathbf{Y}(t)\rangle,\quad t\geqslant u_1. \tag{31}$$

On the basis of (30) and (31) we conclude that at $t=u_2$

$$\mathbf{C}_e(u_2)=\frac{N_0^2\mathbf{Q}_1^2}{2}\langle\tilde{s}(u_1)\mathbf{Y}(u_2)\rangle. \tag{32}$$

To find $\langle\tilde{s}(u_1)\mathbf{Y}(u_2)\rangle$ in (32), we use the correlation moment (28),

$$\mathbf{C}_e(t)=\langle\tilde{s}(u_1)[s^*(u_2)]^2\mathbf{Y}(t)\rangle=\exp[N_0\mathbf{L}(t-u_2)] \times\langle\tilde{s}(u_1)[s^*(u_2)]^2\mathbf{Y}(u_2)\rangle, \tag{33}$$

in which we set $s^*(u_2)=1$, after which we replace $u_2$ by $u_1$. Then at $t=u_2$ we have

$$\langle\tilde{s}(u_1)\mathbf{Y}(u_2)\rangle=\exp[N_0\mathbf{L}(u_2-u_1)]\langle\tilde{s}(u_1)\mathbf{Y}(u_1)\rangle. \tag{34}$$

To find $\langle\tilde{s}(u_1)\mathbf{Y}(u_1)\rangle$ in (34), we use an equation that follows from (12):

$$\langle\tilde{s}(t)\mathbf{Y}(t)\rangle=\left\langle\tilde{s}(t)\exp\left[\mathbf{Q}_1\int_{-\tau/2}^{t}\tilde{s}(t')dt'\right]\right\rangle\mathbf{Y}\left(-\frac{\tau}{2}\right) +\int_{-\tau/2}^{t}\left\langle\tilde{s}(t)\exp\left[\mathbf{Q}_1\int_{\xi}^{t}\tilde{s}(t')dt'\right] \times s^*(t)\mathbf{Q}_2\mathbf{Y}(\xi)\right\rangle d\xi,\quad t\leqslant u_1. \tag{35}$$

The first term on the right-hand side is zero, and averaging the second term yields

$$\langle\tilde{s}(t)\mathbf{Y}(t)\rangle=N_0\int_{-\tau/2}^{t}\delta(t-\xi) \times\exp[-i\Omega(t-\xi)]\mathbf{Q}_2\langle\mathbf{Y}(\xi)\rangle d\xi=\frac{N_0\mathbf{Q}_2}{2}\langle\mathbf{Y}(t)\rangle. \tag{36}$$

If in (31), which is valid for $t\geqslant u_1$, we set $\tilde{s}(u_1)=1$, then

$$\langle s^*(t)\mathbf{Y}(t)\rangle=\frac{N_0\mathbf{Q}_1}{2}\langle\mathbf{Y}(t)\rangle.$$

In view of the symmetry of symmetry of (12) and (13) we can also write

$$\langle\tilde{s}(t)\mathbf{Y}(t)\rangle=\frac{N_0\mathbf{Q}_2}{2}\langle\mathbf{Y}(t)\rangle,\quad t\geqslant u_1.$$

Then at $t=u_1$, with allowance for the second half of this result described by Eq. (36), we obtain

$$\langle\tilde{s}(u_1)\mathbf{Y}(u_1)\rangle=N_0\mathbf{Q}_2\langle\mathbf{Y}(u_1)\rangle. \tag{37}$$

Finally, to determine $\langle\mathbf{Y}(u_1)\rangle$ we use (33), where we set $\tilde{s}(u_1)[s^*(u_2)]^2=1$ and replace $u_2$ by $-\tau/2$:

$$\langle\mathbf{Y}(t)\rangle=\exp\left[N_0\mathbf{L}\left(t+\frac{\tau}{2}\right)\right]\left\langle\mathbf{Y}\left(-\frac{\tau}{2}\right)\right\rangle. \tag{38}$$

Since initially the state of the system, defined by the initial conditions, is determinate, at $t=u_1$ Eq. (38) has the form

$$\langle\mathbf{Y}(u_1)\rangle=\exp\left[N_0\mathbf{L}\left(u_1+\frac{\tau}{2}\right)\right]\mathbf{Y}\left(-\frac{\tau}{2}\right). \tag{39}$$

The cross-correlation function (14) at the end of the excitation pulse can be found from Eqs. (28), (32), (34), (37), and (39) if in (28) we set $t=\tau/2$. This yields

$$\mathbf{C}_e\left(\frac{\tau}{2}\right)=\exp\left[N_0\mathbf{L}\left(\frac{\tau}{2}-u_2\right)\right]\frac{N_0^2\mathbf{Q}_1^2}{2}\exp[N_0\mathbf{L}(u_2-u_1)]N_0\mathbf{Q}_2\exp\left[N_0\mathbf{L}\left(u_1+\frac{\tau}{2}\right)\right]\mathbf{Y}\left(-\frac{\tau}{2}\right). \tag{40}$$

After we have plugged the initial conditions and the matrices (10) and (27) into (40), we find that the components $y_-$ and $y_z$ of the state vector are not correlated with the excitation, and the correlation of the component $y_+$ is described by the following expression:

$$\left\langle\tilde{s}(u_1)[s^*(u_2)]^2y_+\left(\frac{\tau}{2}\right)\right\rangle=\frac{N_0^3Y_0}{4}\exp\left(-\frac{N_0\alpha_e}{4}\right), \tag{41}$$

$$\alpha_e=\frac{3\tau}{2}+u_1. \tag{42}$$

When the excitation pulse is terminated at $t>\tau/2$, i.e., when we have $\tilde{s}(t)=0$, the state vector of the system does not change (this follows from Eqs. (7) and (8)), $\mathbf{Y}(t)=\mathbf{Y}(\tau/2)$, and

$$\langle\tilde{s}(u_1)[s^*(u_2)]^2y_+(t)\rangle=\left\langle\tilde{s}(u_1)[s^*(u_2)]^2y_+\left(\frac{\tau}{2}\right)\right\rangle, \tag{43}$$

with $t\geqslant\tau/2$.

To sum the contributions of all the spectral components of an inhomogeneously broadened system, we must go back to the initial system of coordinates rotating with a frequency $\omega_0$ via Eqs. (6), (9), and (43).

Then, in the case of magnetic resonance, the cross correlation between the spectral component $m_+(t,\Omega)$ of the transverse component of the magnetization vector, which has a frequency detuning $\Omega=\omega-\omega_0$ with respect to the central frequency of the line, and the excitation can be written as follows:

$$\langle\tilde{R}(u_1)[R^*(u_2)]^2m_+(t,\Omega)\rangle=i\exp[i\Omega(t+u_1-2u_2)] \times\left\langle\tilde{s}(u_1)[s^*(u_2)]^2y_+\left(\frac{\tau}{2}\right)\right\rangle. \tag{44}$$

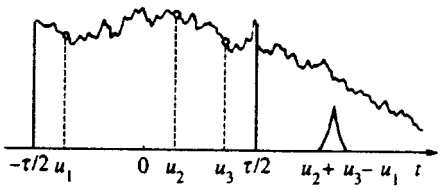



FIG. 2. Time diagram of the excitation of stimulated correlation echo from three noise counts at times $u_1$, $u_2$, and $u_3$.

To find the third-order cross-correlation function between the free induction signal and the excitation, we must integrate (44) with a weight function determined by the shape of the low-frequency equivalent of the inhomogeneously broadened absorption line, $g(\Omega)$:

$$\langle \widetilde{R}(u_1)[R^*(u_2)]^2 m_+(t)\rangle = \int_{-\infty}^{\infty} g(\Omega)\langle \widetilde{R}(u_1) \times [\widetilde{R}(u_2)]^2 m_+(t,\Omega)\rangle d\Omega. \tag{45}$$

Plugging (41)–(44) into Eq. (45) finally yields

$$\langle \widetilde{R}(u_1)[R^*(u_2)]^2 m_+(t)\rangle = i\pi \frac{M_0 N_0^3}{2} \exp\left(-\frac{N_0\alpha_e}{4}\right) G(t+u_1-2u_2),$$

$$G(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(\Omega)\exp(i\Omega t)d\Omega. \tag{46}$$

Reasoning along similar lines, we can find the third-order cross-correlation function

$$\langle \widetilde{R}(u_1)R^*(u_2)R^*(u_3)m_+(t)\rangle \tag{47}$$

for $t>\tau/2>u_3>u_2>u_1>-\tau/2$ and $u_2+u_3-u_1>\tau/2$. This function corresponds to stimulated correlation echo. The time diagram of the envelopes of the excitation pulse, the free induction signal, and the stimulated correlation echo separated from the free induction signal and formed by the noise counts $u_1$, $u_2$, and $u_3$ is depicted in Fig. 2.

Using the above method, we can show that the cross-correlation function (47) is given by the following expression:

$$\langle \widetilde{R}(u_1)R^*(u_2)R^*(u_3)m_+(t)\rangle = i\pi M_0 N_0^3 \exp\left(-\frac{N_0\alpha_s}{4}\right) G(t+u_1-u_2-u_3), \tag{48}$$

$$\alpha_s = \frac{3\tau}{2} + u_1 - u_2 + u_3. \tag{49}$$

Equations (44)–(49) are also valid for the corresponding inhomogeneously broadened systems of the optical range if $m_+$ and $M_0$ are replaced by $p_+$ and $P_0$, respectively.

#### 4. DISCUSSION

The mathematical expectations of the complex-valued envelopes of the two-pulse and stimulated correlation echoes are given by the expressions (46) and (48). Here the time dependence, determined by the functions $G(t+u_1-2u_2)$ and $G(t+u_1-u_2-u_3)$, fully corresponds to ordinary two- and three-pulse echoes, i.e., the responses generated by determinate delta-function radio pulses at times $u_1$ and $u_2$ or $u_1$, $u_2$, and $u_3$, respectively (see Figs. 1 and 2). The same is true of the shape of the echo-responses and the times of their formation.

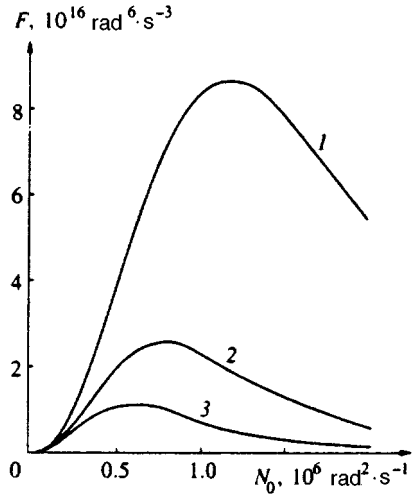



FIG. 3. Dependence of the normalized amplitude of correlation echo on the spectral power density of the noise, $N_0$. For two-pulse echo $\alpha=\alpha_e=3\tau/2+u_1$, and for stimulated echo $\alpha=\alpha_s=3\tau/2+u_1-u_2+u_3$. The values of the parameters $\alpha$ are 10 $\mu$s (curve *1*), 15 $\mu$s (curve *2*), and 20 $\mu$s (curve *3*).

As noted in the Introduction, Paff and Blumich[3] studied the stochastic analog of stimulated spin echo. Comparison of the cross-correlation functions (46) and (48) with the results of their experiment shows that the maximum of the function $G(t+u_1-u_2-u_3)$ coincides with that of the function $C_3(\sigma_1,\sigma_2,\sigma_3)$ in (2) at $u_i=t-\sigma_i$ $(i=1,2,3)$. Their shapes coincide to within system noise caused by the finite averaging time.

However, due to the nonstationary nature of the free induction signal, all averaging in the present paper was done over the ensemble and not over time (as opposed to the situation in Ref. 3). As a result, the cross-correlation functions have, in addition to $G(t+u_1-2u_2)$ and $G(t+u_1-u_2-u_3)$, exponential factors that depend on the position of the noise counts $u_i$ on the time axis, in accordance with Eqs. (42) and (49).

In the present paper we also find the dependence of the correlation echo amplitudes on the parameters of the noise pulse. According to (46) and (48), this dependence is nonlinear and is given by the function

$$F(N_0) = N_0^3 \exp\left(-\frac{N_0\alpha}{4}\right), \tag{50}$$

with $\alpha=\alpha_e$ for two-pulse echo and $\alpha=\alpha_s$ for stimulated echo. The diagrams of these functions for different values of $\alpha$ are depicted in Fig. 3.

In the limit $N_0\alpha \ll 4$ the dependence is cubic. The rate of growth of the function diminishes as $N_0$ increases. The maxi-

mum of the normalized correlation echo amplitude determined by the function $F(N_0)$ is $[12/\alpha e]^3$ and is reached at $N_0\alpha = 12$.

The exponential dependence in (50) is a reflection of the effect of noise in the intervals between noise counts forming the echo-response. We analyze this effect by using the example of a two-pulse echo formed on the noise counts $u_1$ and $u_2$. These moments divide the noise pulse into three intervals. Here at time $u_1$ the longitudinal component $y_z$ of the state vector is transformed into the $y_-$ component, and at time $u_2$ the $y_-$ component is transformed into $y_+$. To put it schematically,

$$y_z \rightarrow y_- \rightarrow y_+ .$$

The matrix components in (40) describe the damping of the components of the state vector **Y** caused by the noise in the three intervals. Equation (40) implies that in the first interval the longitudinal component $y_z$ decays according to the law $\exp[-N_0(u_1+\tau/2)/2]$, while in the second and third intervals the components $y_-$ and $y_+$ decay with a rate that is two times smaller: $\exp[-N_0(u_2-u_1)/4]$ and $\exp[-N_0(\tau/2-u_2)/4]$, respectively. The overall damping is determined by the expressions (41) and (42). Here it occurs that according to (41) and (42) the amplitude of the two-pulse correlation echo depends both on the length $\tau$ of the noise pulse and on the time $u_1$. Since $|u_1| < \tau/2$ holds, for a fixed pulse length $\tau$ the values of $\alpha_e$ may differ by a factor of two, depending on the position of $u_1$. For $N_0\alpha_e \ll 4$ the position of $u_1$ has no effect on the amplitude of the echo. In the opposite case the echo amplitude depends on the position of $u_1$, and this dependence is quite strong. For instance, at $t = 10\ \mu$s the value of $\alpha_e$ can change from $10^{-5}$ s as $u_1 \rightarrow -\tau/2$ to $2\times 10^{-5}$ s as $u_1 \rightarrow \tau/2$ (see Fig. 3). Here at $N_0 = 10^6$ rad$^2$ s$^{-1}$ the correlation echo amplitude changes more than by a factor of ten.

The parameters of the noise pulse given in the above example and the $F$ vs $N_0$ curves in Fig. 3 correspond to the noise parameters used in exciting nuclear spin echo and are determined from the condition that

$$\tau = 10\ \mu\text{s} \ll T_2 .$$

In the optical range the time $T_2$ amounts to several tens of nanoseconds.[8] Then for a 10-ns noise pulse the value of $\alpha$ in Fig. 3 is a thousand times smaller and $N_0$ becomes a thousand times larger. The function $F(N_0)$ grows by a factor of $10^9$.

Note that in most light-echo excitation experiments the excitation of inhomogeneously broadened lines, whose width amounts to $10^9$–$10^{11}$ s$^{-1}$, is partial.[8,11]

In these conditions coherent light pulses roughly 10-ns long[8] have a spectral width $\sim 10^8$ s$^{-1}$. At the same time, the spectral width of a noise pulse of length $\tau$ is practically independent of the pulse length. Thus, with an appropriate choice of the spectral width of band white noise a 10-ns pulse can ensure uniform excitation of the entire inhomogeneously broadened line. For this reason the function $G(t)$ may not coincide with the envelope of the light echo formed in the event of partial excitation of the inhomogeneously broadened line.

Another distinctive feature of the optical range that sets it apart form the range used in NMR experiments is the condition of spatial synchronism.[8] Here the free induction signal propagates in the same direction as the exciting pulse in the form of a plane wave with a vector **k**. The two-pulse and stimulated echoes in the light induction signal are formed by pairs and triples of noise pulse counts.

As is known,[8] two-pulse and stimulated light echoes propagate in the directions

$$\mathbf{k}_e = -\mathbf{k}_1 + 2\mathbf{k}_2 , \tag{51}$$

$$\mathbf{k}_s = -\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 , \tag{52}$$

respectively, where $\mathbf{k}_i$ is the wave vector of the $i$th excitation light pulse. Since in the case under consideration all noise counts belong to the same excitation noise pulse propagating in the direction **k**, we have

$$\mathbf{k}_1 = \mathbf{k}_2 = \mathbf{k}_3 = \mathbf{k}. \tag{53}$$

Equations (51)–(53) imply that $\mathbf{k}_e = \mathbf{k}_s = \mathbf{k}$, and all echoes propagate in the direction in which the light induction signal propagates.

In conclusion let us briefly discuss the results related to stimulated correlated echo. In the formation of this echo the components of the state vector are transformed at times $u_1$, $u_2$, and $u_3$ according to the scheme $y_z \rightarrow y_- \rightarrow y_z \rightarrow y_+$. The longitudinal component $y_z$ decays twice as rapidly as the transverse components. As a result the overall damping is determined by the exponential factor in (48). Here, according to (49), the stimulated echo amplitude depends on the position of all three noise counts $u_1$, $u_2$, and $u_3$ on the time axis. As in the case of two-pulse echo, for a noise pulse of fixed length $\tau$ the values of $\alpha_s$ can differ by a factor of two, depending on the position of $u_1$, $u_2$, and $u_3$.

The present work was made possible by a grant from the Ministry of General and Vocational Education of the Russian Federation (1995 Grant Competition in the Field of Electronics and Radio Engineering, Project 5-32, Grant GR/TOR-17).

[1)] In Eq. (2) of Ref. 3 the product of stationary processes is averaged over time. Formally the procedure can be written as $\lim_{T\rightarrow\infty} (1/2T) \int_{-T}^{T} f(t)dt$, a feature mentioned in Ref. 20 of the cited paper. Note that in ergodic processes ensemble and time averages coincide.

[1] R. R. Ernst, J. Mater. Res. **3**, 10 (1970).
[2] R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of NMR in One and Two Dimensions*, Clarendon Press, Oxford (1987), Chap. 4.
[3] J. Paff and B. Blumich, Phys. Rev. A **43**, 3640 (1991).
[4] R. Beach and S. R. Hartmann, Phys. Rev. Lett. **53**, 663 (1984).
[5] V. I. Tikhonov, *Nonlinear Transformations of Stochastic Processes, Pergamon, Oxford, 1965.*
[6] W. R. Knight and R. Keiser, J. Mater. Res. **48**, 293 (1982).
[7] S. A. Baruzdin, Fiz. Tverd. Tela (St. Petersburg) **37**, 3497 (1995) [Phys. Solid State **37**, 1922 (1995)].
[8] É. A. Manykin and V. V. Samartsev, *Optical Echo Spectroscopy* [in Russian], Nauka, Moscow (1984), Chaps. 1–4.
[9] R. Beach, D. DeBeer, and S. R. Hartmann, Phys. Rev. A **32**, 3467 (1985).
[10] S. A. Baruzdin, Radiotekh. Elektron. **41**, 959 (199).
[11] N. N. Akhmediev, B. S. Borisov, V. A. Zuĭkov, V. V. Samartsev, R. G. Usmanov, and B. Sh. Khamidullin, Izv. Akad. Nauk SSSR, Ser. Fiz. **50**, 1488 (1986).

[12] A. Korpel and M. Chatterjee, Proc. IEEE **69**, 22 (1981) [Mir Publishers, Moscow].
[13] V. S. Kuz'min, I. Z. Rutkovskiĭ, A. P. Saĭko, A. D. Tarasevich, and G. G. Fedoruk, Zh. Eksp. Teor. Fiz. **97**, 880 (1990) [Sov. Phys. JETP **70**, 493 (1990)].
[14] V. S. Kuz'min, I. Z. Rutkovskiĭ, A. P. Saĭko, A. D. Tarasevich, and G. G. Fedoruk, Izv. Akad. Nauk SSSR, Ser. Fiz. **53**, 2358 (1989).
[15] V. B. Ustinov, M. M. Kovalevskiĭ, and S. A. Baruzdin, Izv. Akad. Nauk SSSR, Ser. Fiz. **50**, 1495 (1986).
[16] S. A. Baruzdin, in *Digest of 1st Intern. Conf. on Differential Equations and Their Applications* [in Russian], St. Petersburg Univ. Press, St. Petersburg (1996), p. 25.
[17] V. I. Tikhonov, *Statistical Radio Engineering* [in Russian], Radio i Svyaz', Moscow (1982), p. 185.

Translated by Eugene Yankovsky

# Production of heavy atomic clusters upon interaction of laser radiation with matter

Yu. P. Gangrskiĭ, S. G. Zemlyanoĭ, B. N. Markov, G. V. Myshinskiĭ, and V. O. Nesterenko

*Joint Institute of Nuclear Research, 141980 Dubna, Moscow Region, Russia*

I. V. Vorykhalov, I. N. Izosimov, A. A. Rimskiĭ-Korsakov, and V. V. Smirnov

*V. G. Khlopin Radium Institute, 194021 St. Petersburg, Russia*
(Submitted 5 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 78–88 (July 1997)

The results of an investigation of the formation of Pb, Th, and U clusters over a broad range of numbers of atoms (from a few atoms to macroparticles) upon interaction of high-power pulsed laser radiation with matter are presented. Clusters of fissionable elements are studied for the first time. A setup for determining the yield of clusters and the number of atoms in them, which is based on the use of several different methods (laser resonance fluorescence, time-of-flight measurements, and counting the number of tracks of fission fragments from the cluster nuclei), is described. The dependence of the yield of clusters with various numbers of atoms on the conditions for their formation is discussed. © *1997 American Institute of Physics.* [S1063-7761(97)00707-5]

## 1. INTRODUCTION

The investigation of atomic clusters is presently one of the most promising and rapidly developing research areas in physics.[1–6] A unique possibility for investigating the evolution of systems from an atom to a solid has appeared owing specifically to atomic clusters. The information obtained in these investigations is of great importance for developing our understanding of atoms, nuclei, solids, and crystals.

Several directions can be identified in experiments with clusters. One of the most promising is the study of large clusters, which combine the properties of both micro- and macrosystems. The investigation of how such clusters form in various processes, and of shell structure in them, is of great interest. It is expected that the transition from micro- to macrosystems in these clusters can occur when the number of atoms in them is of the order of several tens of thousands. This number naturally depends on the kind of atoms, the temperature, and several other characteristics of the clusters. However, these questions have been investigated to a comparatively small extent. There is an especially small amount of data on the properties of clusters of the heaviest elements (with atomic numbers $Z_i > 82$). Actually, the data cover elements up to lead, whose clusters were studied, for example, in Refs. 1 and 7–11. There is essentially no information on clusters of actinides and transuranian elements. This is apparently due primarily to the difficulties of working with radioactive elements.

This paper presents the results of research on clusters of heavy elements (Pb, Th, and U) over a broad range of the numbers of atoms in them (up to millions). Atomic clusters of fissionable elements are considered for the first time. The clusters were obtained via the interaction of high-power pulsed laser radiation with the metals cited. The experimental setup and measurement of the dependence of the yield of different clusters on the number of atoms in them at various laser powers, and therefore different cluster formation temperatures, are described.

The interaction of high-power laser radiation with matter has been studied in great detail.[12,13] At pulsed laser power densities below $10^9$ W/cm$^2$, three processes dominate: 1) heating without alteration of the phase state, 2) melting and vaporization, and 3) ionization and plasma formation. These processes are accompanied by the ejection of atoms, molecules, singly- and multiply-charged ions, clusters, and macroscopic particles of the material from the surface of the irradiated sample. The relationship of the intensities of all these components is determined both by the characteristics of the laser radiation and by the surface properties of the irradiated sample. This leads to substantially different conditions for the occurrence of cluster formation. Three principal cluster formation paths can be identified; a) condensation from vaporized atoms at the initial instant of fragmentation of the material, b) disintegration of hot macroscopic pellets, and c) intense vaporization of atoms and dimers from the latter. It can be expected that because of the variety of conditions in these processes, the clusters formed will be characterized by a large set of masses and temperatures.

## 2. EXPERIMENTAL SETUP

The setup used in our experiments to investigate clusters over a broad range of masses (from a few atoms or ions to macroscopic particles) incorporates three different methods for detecting them. The detection of single atoms or singly-charged ions is based on recording their resonance fluorescence, which is excited by a tunable laser. In clusters of up to several hundred atoms, time-of-flight measurements were employed. Finally, the heaviest clusters (with more than a thousand atoms) were identified by counting the number of nuclear fission fragments in the cluster.

A block diagram of the setup is presented in Fig. 1. The output of a pulsed laser was focused on a sample in a

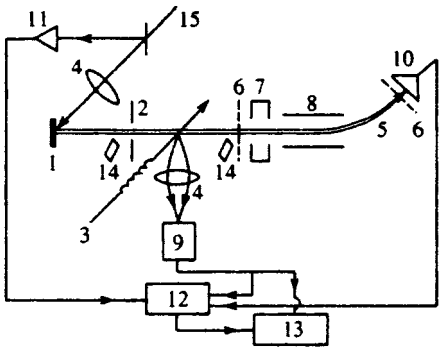



FIG. 1. Experimental setup: *1* — sample; *2* — collimator; *3* — continuous dye laser beam; *4* — optical lenses; *5* — beam of atoms and clusters ; *6* — accelerating grids; *7* — electrostatic lens; *8* — deflecting plates; *9* — photomultiplier; *10* — ion detector; *11* — detector of pulsed laser radiation; *12* — time analyzer; *13* — computer; *14* — dielectric detectors; *15* — pulsed laser beam.

vacuum chamber, which was evacuated by a turbomolecular pump to a residual pressure of $5\times10^{-6}$ Torr. The laser radiation was focused by a lens with a diameter of 20 mm and a focal length of 45 mm. The dimensions of the laser spot on the surface of the sample were varied by moving the lens using a vacuum bellows drive.

Pulsed lasers of two types, viz., a copper vapor laser (LGI-201) and an yttrium–aluminum garnet laser (LTIPCh-7), were used to obtain the clusters. The parameters of these lasers (wavelength $\lambda$, frequency $\omega$, pulse width $\Delta t$, pulsed power density $q$, and mean power $\overline{P}$) are presented in Table I. These laser parameters clearly make it possible to vary the incident power density at the sample over a broad range (from $10^6$ to $10^9$ W/cm$^2$), and thereby to study the formation of clusters over a broad range of temperatures.

The flux of atoms, ions, and clusters formed upon interaction of pulsed laser radiation with the sample surface was shaped by a system of collimators into a beam with a specified angular divergence. This beam crossed the laser beam employed for resonant excitation of the atoms or ions at a right angle. The source of the latter radiation was a dye laser pumped by a continuous-wave argon ion laser. The laser frequency was automatically swept over a predetermined range with a width up to 30 GHz. When a stabilization system was employed, the half-width of the laser line was at most 20 MHz, which is significantly less than the Doppler shift due to the energy and angular spread of the beam of atoms or ions. The spontaneous optical fluorescence emitted by the excited atoms or ions was focused by a system of lenses onto the cathode of a photomultiplier operating in the single-photon counting mode.[14]

TABLE I. Parameters of the lasers used to obtain clusters.

| Laser | $\lambda$, nm | $\omega$, Hz | $\Delta t$, ns | $E_p$, mJ | $\overline{P}$, W | $q$, W/cm$^2$ |
|---|---|---|---|---|---|---|
| LGI-201 | 510+570 | $10^4$ | 20 | 0.5 | 5 | $10^6-10^7$ |
| LGIPCh-7 | 1064 | 25 | 10 | 50 | 0.12 | $10^7-3\times10^9$ |
| | 532 | 25 | 10 | 20 | 0.6 | $10^7-2\times10^9$ |

The spectra of the resonant frequencies of the atoms or ions measured in this manner enable one to determine their intensity and mean velocity (from the Doppler line broadening). These data are a source of information on the conditions for cluster formation when high-power laser radiation interacts with matter (specifically, on the plasma temperature and density at the surface of the sample).

The characteristics of the clusters (mass and velocity spectra) were determined by measuring the time of flight of charged clusters over a given distance (from the surface of the sample to the detector).[15] The charged clusters entered a region with an accelerating electric field after passing through the shaping collimator. This field was produced by a high voltage (up to 2.8 kV) applied to a grid (the length of the accelerating gap was 30 mm). The beam of clusters with some given energy obtained in this manner was focused by an electrostatic lens and entered a drift gap 400 mm long. A cluster detector (secondary electron multiplier) was located at the end of this gap. Deflecting plates, within the gap were used to separate the charged and neutral clusters so that the latter would not strike the detector.

Pulses from this detector were the ''stop'' signals for the time analyzer, and pulses from the photodiode onto which a portion of the pulsed laser radiation was directed served as the ''start'' signals. The latter signals corresponded to the time of escape of a cluster from the sample surface to sufficient accuracy. The measured time of flight $T$ is related to the cluster mass $M$ (in mass units) by

$$M=T^2\left[\frac{7.25\times10^{-7}L_1}{E_0+U}+\frac{1.44\times10^{-6}(L_2\sqrt{E_0+U}-L_1\sqrt{E_0})}{U}\right]^{-2}, \tag{1}$$

where $U$ is the accelerating potential [V], $E_0$ is the initial energy of the clusters [eV], and $L_1$ and $L_2$ are, respectively, the lengths of the accelerating and drift gaps [mm].

The velocity distribution of neutral atoms could also be measured in the same manner. In this case, a pulse from the photomultiplier recording resonantly scattered fluorescence served as the ''stop'' signal. This velocity distribution also served as a source of information regarding the temperature at the sample surface.

As the mass of the singly-charged clusters increases, their kinetic energy remains unchanged (it is determined by the accelerating voltage), but their velocity decreases. For heavy clusters containing more than 10 000 atoms, the velocity approaches the thermal value ($10^4-10^5$ cm/s), and the signal from the ion detector decreases to the noise level. Therefore, another method was used to detect such heavy clusters and to determine their mass. This method is applicable only to clusters consisting of atoms of fissionable elements (for example, Th or U), and is based on counting the fission fragments produced by an intense flux of neutrons or $\gamma$ rays. The number of fission fragments from a cluster is given by

$$N=2A\int_0^{E_0}\sigma(E)Y(E,t)dE, \tag{2}$$

where $A$ is the number of atoms in the cluster, and $\sigma(E)$ and $Y(E,t)$ are the fission cross section and the total flux of bombarding particles with energy $E$ during time $t$.

The fission fragments were detected by a dielectric detector, a Mylar film 20 $\mu$m thick. After chemical treatment of the detector (etching in a 20% NaOH solution for 15 min), the fission fragment tracks became visible under an optical microscope (their lengths amounted to $\simeq 10$ $\mu$m).[16] Counting the number of tracks that emerge from one point and are due to the splitting of cluster atoms yields the number of atoms of the element undergoing fission in the cluster. The relation between these quantities is given by

$$A = N\left[2\varepsilon \int_0^{E_0} \sigma(E) Y(E,t)\, dE\right]^{-1}, \tag{3}$$

where $\varepsilon$ is the fragment detection efficiency of the dielectric detector. In the experiments described here, in which the clusters were deposited on the detector surface, $\varepsilon = 0.5$, i.e., the observed number of tracks corresponded to the number of nuclei undergoing fission in a cluster.

It is clear from (3) that the coefficient relating the number of tracks to the number of atoms in a cluster is determined by the fission cross section of the nuclei sample and the total flux of bombarding particles. For example, for a sample of $^{235}$U and irradiation of the detector by hot neutrons, the fission cross section $\sigma = 5\times 10^{-22}$ cm$^2$, and the intensity of the neutrons in a thermal column of a nuclear reactor can reach $10^{15}$ neutrons/s$\cdot$cm$^2$. Under these conditions, after the detector has been irradiated for several hours ($\simeq 2\times 10^4$ s)

$$\int_0^{E_0} \sigma(E) I(E)\, dE\, dt \simeq 10^{-2}, \tag{4}$$

i.e., each observed track corresponds to about a hundred cluster atoms. Heavier clusters containing many thousands or millions of atoms are detected more conveniently using smaller neutron fluences or irradiation by $\gamma$ rays, for which the fission cross section is appreciably smaller ($\sigma \simeq 0.1-0.5$ barns). Thus, this method makes it possible to detect and identify clusters that are inaccessible to the time-of-flight method.

Measurement is most convenient when some dozens of tracks (up to a hundred) are observed from each cluster. When the number of tracks is large, they overlap, and, consequently, there are counting errors. Since the number of nuclear fission events in a cluster is a statistical quantity, the error counting, and hence the number of atoms in a cluster, is given by

$$\frac{\Delta N}{N} = \frac{a}{\sqrt{KN}}, \tag{5}$$

where $K$ is the number of clusters observed. When $N$ and $K$ ranged from 10 to 100, the counting accuracy for cluster atoms, which characterizes the mass resolution of the method, was 3–10%. Although such resolution fails to provide a detailed picture of the mass distribution in some cases, this method is highly sensitive and can be used successfully for small yields of clusters.

TABLE II. Characteristics of the elements investigated.

| | Na | Pb | Th | U |
|---|---|---|---|---|
| Density, g/cm$^3$ | 0.371 | 11.34 | 11.72 | 19.04 |
| Boiling point, K | 1159 | 978 | 5063 | 4303 |
| Heat of sublimation, kJ/mole | 90.1 | 178 | 540 | 494 |
| Resistivity, $10^{-6}\Omega\cdot$cm | 4.28 | 19.2 | 13 | 21 |

Dielectric detectors of fission fragments are essentially insensitive to other types of radiation, viz., plasma fluorescence, atoms, and ions. They can therefore be placed in direct proximity to the surface of an irradiated sample, and high measurement sensitivity can thereby be achieved. The disintegration of hot clusters as they move away from the sample can be evaluated by positioning detectors at different distances (10–400 mm) from the sample surface and comparing the mass distribution of the clusters for the various detectors.

## 3. EXPERIMENTAL RESULTS

Experiments were performed on the setup described above to determine the mass spectrum of particles (from a few atoms and ions to macroparticles) formed when laser radiation acts on samples of lead, thorium, and uranium ($^{235}$U and $^{238}$U) over a broad range of power densities. Table II presents some characteristics of these elements, as well as of sodium, which is widely used in cluster research,[17] for comparison. These characteristics permit evaluation of the thermodynamic and electrical properties of the elements investigated. Comparison with sodium reveals that they require a considerably higher atomization temperature and energy, while their comparatively low electrical conductivity indicates a smaller concentration of free electrons. In each of our experiments, we determined the mass of vaporized material from the sample by weighing the sample before and after the experiment, and by measuring the size of the crater produced on the sample surface (the two methods led to similar results).

Figure 2 presents the measured time-of-flight spectrum of $^{238}$U atoms for laser radiation with a wavelength of 1064 nm and two different power densities (the area of the optical radiation spot on the sample surface was $\simeq 1$ mm$^2$). Knowing the length of the flight path (from the sample to the laser beam) and the mass of the atoms, we can easily calculate the

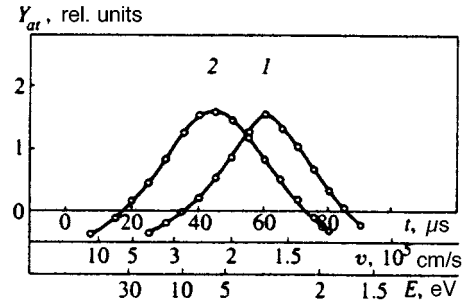



FIG. 2. Time-of-flight spectra of uranium atoms for various power densities: *1* — $q = 3\times 10^7$ W/cm$^2$; *2* — $q = 10^8$ W/cm$^2$.

TABLE III. Velocities and yields of atoms upon laser irradiation of Pb and U samples.

| λ, nm | $q$, W/cm$^2$ | $d$, mm | $v$, $10^5$ cm/s | $Y_{at}$, atoms/pulse | |
|---|---|---|---|---|---|
| | | | | Pb | U |
| 510+570 | $3.0\times10^7$ | 0.3 | $1.5\times10^5$ | $3.5\times10^{12}$ | $5.0\times10^{10}$ |
| 1064 | $1.1\times10^8$ | 1.4 | $2.0\times10^5$ | $4.0\times10^{13}$ | $1.5\times10^{12}$ |
| 1064 | $2.5\times10^8$ | 1.0 | $2.2\times10^5$ | $1.5\times10^{14}$ | $1.0\times10^{13}$ |
| 1064 | $5.0\times10^8$ | 0.7 | $2.5\times10^5$ | $6.0\times10^{14}$ | $4.0\times10^{13}$ |
| 1064 | $3.0\times10^9$ | 0.3 | $2.8\times10^5$ | $2.5\times10^{14}$ | $1.5\times10^{14}$ |

velocity and energy spectra from these spectra. The corresponding values are presented in Fig. 2 on parallel horizontal axes. The resulting spectra correspond to Maxwellian velocity and energy distributions of the atoms vaporized from the sample. They permit evaluation of the temperature of the sample surface in the laser radiation spot. The temperature was evaluated with consideration of the increase in the velocity of the atoms due to the gas-dynamic expansion of the plasma.[13,18] Table III presents the mean velocities and yields of atoms from Pb and U samples for various power densities and diameters of the laser spot.

Figure 3 presents the yield of single $^{238}$U atoms and the mass of vaporized material that condenses into clusters and macroparticles as a function of laser power density, measured under the same conditions (their mass spectrum is considered below). The temperature of the sample surface is presented on a parallel horizontal axis. It is clear that optimum power densities (or temperatures) — at which their yields are maximum — are observed for clusters and single atoms (this temperature is lower for clusters). The measurements also show that the yield of ions over the entire experimental range of laser intensity is less than 1% of the total amount of vaporized material. We therefore conclude that a considerable fraction of the material is vaporized in the form of clusters, neutral atoms, and macroparticles.

Figure 4 presents the mass distribution measured by the time-of-flight method of particles vaporized from Pb and U samples by laser radiation at 510 and 570 nm with power density $3\times10^7$ W/cm$^2$. We see that singly-charged clusters

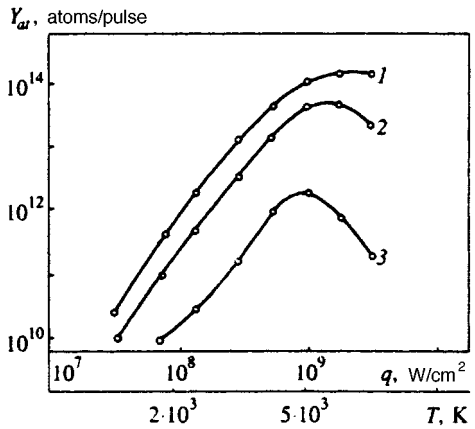



FIG. 3. Dependence of the total amount of the material (*1*), the number of neutral atoms (*2*), and the number of uranium clusters with $N>1000$ (*3*) per laser pulse on the power density of the laser radiation.

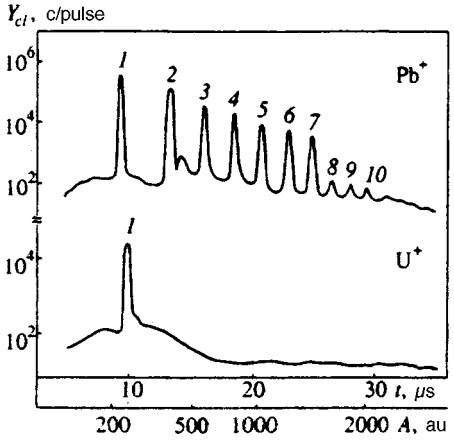



FIG. 4. Mass distribution of lead and uranium clusters measured by the time-of-flight method for $q=3\times10^7$ W/cm$^2$. Peaks are labeled by the number of atoms in a cluster.

containing up to 10 atoms are observed in the case of Pb, while only single ions are observed in the case of U. A similar mass distribution with a single peak of single ions is obtained for the Th sample. The fraction of ionized clusters containing from 2 to 100 atoms is less than 1% for both elements. Similar results were obtained at other power densities (the lowest power density was $10^7$ W/cm$^2$, and the corresponding temperature was $\simeq$2000 K) and for laser radiation at a different wavelength (1064 nm).

The mass distributions of the clusters containing more than 1000 atoms were measured using the track method for counting fission fragments after the irradiation of dielectric detectors in a neutron flux. Such clusters (or macroparticles) were observed in the case of the Th and U samples. For the highest neutron fluence and a U sample, one track corresponded to 1100 atoms at the detector surface. The single tracks could belong either to individual U atoms or atoms grouped in a cluster with $A<1000$. A comparison of the number of tracks with the number of atoms determined from the resonance fluorescence measurements showed that at least 95% of the single tracks belong to individual atoms and that at most 5% of the clusters have $A<1000$. At the same time, instances with more than two tracks correspond to clusters (or macroparticles) containing more than 1000 atoms.

An example of one of the distributions of such systems with respect to the number of atoms in them is presented in Fig. 5. To minimize the statistical error in the number of clusters, they are combined into groups containing $A-3A$ atoms, where $A$ is the minimum number of atoms in a cluster belonging to a given group. In this manner, the error in the yield of clusters in each group is at most 5%. It is clear from Fig. 5 that, as when there are few atoms in the clusters, their yield decreases with increasing $A$. This is typical of the entire range of laser intensity.

Table IV presents the yield of various components (atoms, ions, small ($A=2-100$), and large ($A>1000$) clusters) for various laser intensities. We see that for Pb, the fraction of small clusters amounts to $\simeq10^{-2}$ of the number of single atoms, while the fraction for U and Th is less than $10^{-3}$. The fraction of large clusters with $N>100$ atoms is

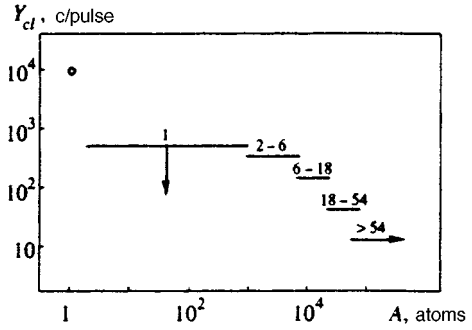



FIG. 5. Mass distribution of uranium clusters determined from the number of fission fragment tracks. $q=3\times10^7$ W/cm$^2$. The lines are labeled with the number of fission fragment tracks observed. The point corresponds to the number of single atoms.

even smaller. An appreciable yield of these clusters was observed owing to the higher sensitivity of the measurement method, the absence of a background, and summation over a broad range of numbers of atoms in a cluster.

## 4. DISCUSSION

These experiments exhibited significant differences in the yield of clusters of different elements produced by high-power laser radiation. Under identical irradiation conditions, an appreciable yield of ionized Pb clusters containing 2–10 atoms was observed, but similar clusters of Th and U were not detected (their yield was at least two orders of magnitude lower than the yield for Pb).

The temperature of the plasma produced when laser radiation interacts with the sample surface can be estimated from the measured velocity spectra of the atoms of the elements investigated. At laser power densities ranging from $3\times10^7$ to $3\times10^9$ W/cm$^2$, this temperature is $2\times10^3-6\times10^3$ K. These values are appreciably greater than in typical experimental setups, for sample, in systems employing supersonic vapor jets through a narrow nozzle.[19] Such a high plasma temperature is clearly not optimal for cluster formation. However, at a lower temperature (which can be obtained by lowering the laser power), the yield of vaporized atoms drops sharply, especially from such refractory elements as Th and U, which naturally reduces the yield of clusters.

The probability of producing small clusters with $A=2-100$ from macroscopic pellets, which can be observed by measuring fission fragments of Th and U nuclei in them, is also very low. Their temperature is probably too low for disintegration or intense vaporization of atoms from them, which could otherwise have led to the production of small clusters. This hypothesis is confirmed by measurements of the yields and mass distributions of pellets with $A>10^5$ at various distances from the sample. It was found that the yield of such pellets decreases as the square of the distance from the sample, and that the mass distributions determined from the number of fission fragment tracks are essentially identical. This can mean that the disintegration of the pellets or the vaporization of a large number of atoms from them, which could lead to the formation of small clusters, is unlikely. Estimates of the time needed for that many atoms ($>1000$) to vaporize, which are similar to those described in Ref. 20, demonstrated that it is large ($>10^{-3}$ s) — much greater than the time of flight of the pellets to the fragment detectors.

Large clusters with $A>1000$ might form with high probability from macroscopic pellets. A typical feature is their oscillatory dependence of the surface energy — and therefore the yield — on the number of atoms.[21] However, the method for counting fission fragments used in the present work does not make it possible to identify such oscillations. Therefore, the production of large Th and U clusters from macroscopic pellets requires further study.

Small clusters can form when vaporized atoms condense. This probably takes place in lead, in which we observed the production of ionized clusters containing 2–10 atoms. However, only individual ionized atoms were observed for U and Th. In all likelihood, such a difference in cluster formation probabilities is associated with significantly different electronic states of the resulting systems. This is probably reflected in their thermophysical properties (Table II). While the conditions for producing small clusters for lead, which has low specific heat, were quite acceptable, more optimal conditions are apparently needed for Th and U. They could probably be achieved by using a cluster cooling system,[1] which would permit more accurate control of their formation conditions. We propose doing this in our forthcoming investigations.

Consistent previous investigations,[7–10] the mass distribution of ionized Pb clusters exhibits a sharp decrease in yield at $A>7$. This can basically be accounted for by the icosahedral packing of atoms in lead clusters ,[8–10] which is responsible for the features at $A=7$, 10, 13, and 19.

## 5. CONCLUSIONS

These experiments showed that the present setup, which incorporates three different methods, offer great promise for

TABLE IV. Yield of various components upon irradiation of Pb and U samples by laser radiation at $\lambda=1064$ nm.

| | $Y_{at}$, atoms/pulse | | $Y_{ion}$, atoms/pulse | | $Y_{cl}$, clusters/pulse ($A=2$–$100$) | | $Y_{cl}$, clusters/pulse ($A>1000$) |
|---|---|---|---|---|---|---|---|
| $q$, W/cm$^2$ | Pb | U | Pb | U | Pb | U | U |
| $1.1\times10^8$ | $1.5\times10^{13}$ | $5.3\times10^{11}$ | $1.8\times10^{11}$ | $1.0\times10^9$ | $2.5\times10^{11}$ | $<10^9$ | $6.5\times10^{10}$ |
| $2.5\times10^8$ | $5.5\times10^{13}$ | $2.5\times10^{12}$ | $6.5\times10^{11}$ | $2.3\times10^{10}$ | $6.0\times10^{11}$ | $<10^9$ | $1.5\times10^{11}$ |
| $5.0\times10^8$ | $2.2\times10^{14}$ | $1.2\times10^{13}$ | $2.2\times10^{12}$ | $1.2\times10^{11}$ | – | – | $1.8\times10^{12}$ |
| $3.0\times10^9$ | $3.5\times10^{14}$ | $1.8\times10^{13}$ | $4.5\times10^{13}$ | $2.5\times10^{12}$ | – | – | $2.7\times10^{11}$ |

research on the formation of clusters of heavy elements over a broad range of masses. It has been shown that the yield of clusters formed when high-power laser radiation interacts with matter depends on the temperature of the plasma formed. At the same time, the distribution of the clusters according to the number of atoms in them is significantly less critical with respect to these conditions. The probability of cluster formation under such conditions varies significantly for various elements (it is markedly greater for Pb than for Th and U). The atomic systems of the fissionable elements U and Th have been investigated for the first time. A new method for measuring the number of atoms in them based on counting fission fragments is proposed.

In conclusion, we thank Yu. Ts. Oganesyan and Yu. É. Penionzhkevich for their sustained interest in this work, V. V. Kresin for useful discussions, and V. P. Perelygin and his coworkers for their assistance in processing and scanning the dielectric detectors of the fission fragments.

[1] W. A. de Heer, W. D. Knight, M. Y. Chou *et al.*, Solid State Phys. **40**, 93 (1987).
[2] V. V. Kreisin, Phys. Rep. **220**, 1 (1992).
[3] V. O. Nesterenko, Fiz. Elem. Chastits At. Yadra **23**, 1665 (1992).
[4] W. A. de Heer, Rev. Mod. Phys. **65**, 611 (1993).
[5] M. Brack, Rev. Mod. Phys. **65**, 677 (1993).
[6] C. Brechignac and J. P. Connerade, J. Phys. B: At. Mol. Opt. Phys. **27**, 3795 (1994).
[7] J. Muhlbach, K. Satteler, P. Pfau, and E. Recknagel, Phys. Rev. Lett. **87**, 415 (1992).
[8] K. Satteler, Surf. Sci. **156**, 292 (1985).
[9] H. S. Lim, C. K. Ong, and F. Ercolessi, Surf. Sci. **270**, 1109 (1992).
[10] H. S. Lim, C. K. Ong, and F. Ercolessi, Z. Phys. D **26**, 45 (1993).
[11] H. R. Siekmann, E. Holubkrappe, B. Wrenger *et al.*, Z. Phys. B **90**, 201 (1993).
[12] J. F. Ready, *Effects of High-Power Laser Radiation*, Academic Press, New York (1971).
[13] Yu. A. Bykovskiĭ, S. M. Sil'nov, E. A. Sotnichenko, and B. A. Shestakov, Zh. Éksp. Teor. Fiz. **93**, 500 (1987) [Sov. Phys. JETP **66**, 285 (1987)].
[14] Yu. P. Gangrskiĭ, K. P. Marinova, B. N. Markov *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **49**, 2261 (1985).
[15] I. N. Izosimov, I. V. Vorykhalov, and A. A. Rimskiĭ-Korsakov, Izv. Akad. Nauk SSSR, Ser. Fiz. **59**, 21 (1995).
[16] R. L. Fleischer, P. B. Price, and R. M. Walker, *Nuclear Tracks in Solids: Principles and Applications*, University of California Press, Berkeley, Calif. (1975).
[17] *Physical Constants. A Handbook* [in Russian], I. S. Grigor'ev and E. Z. Meĭlikhov (eds.), Énergoizdat, Moscow (1991).
[18] I. A. Zel'dovich and Yu. P. Raizer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena*, Academic Press, New York (1966–1967).
[19] S. Bjornholm, J. Borggreen, O. Echt *et al.*, Phys. Rev. Lett. **65**, 1627 (1990).
[20] B. M. Smirnov, Zh. Éksp. Teor. Fiz. **110**, 47 (1996) [JETP **83**, 24 (1996)].
[21] B. M. Smirnov, Zh. Éksp. Teor. Fiz. **108**, 1810 (1995) [JETP **81**, 992 (1995)].

Translated by P. Shelnitz

# High-order optical-harmonic generation in films of nonmetallic crystals

V. A. Kovarskiĭ

*Institute of Applied Physics, Moldavian Academy of Sciences, MD-2028 Kishinev, Republic of Moldova*
(Submitted 29 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 89–96 (July 1997)

High-order optical-harmonic generation in nonmetallic films interacting with pulses of laser light is examined. The wave functions of the current carriers in a crystal in an external electromagnetic field are chosen in the form of Volkov–Keldysh solutions. An explicit expression for the intensity of the $s$th harmonic, which depends on the crystal parameters, is derived. A plateau and a cutoff effect, similar to those in the case of harmonic generation on an isolated atom, have been detected. Finally, numerical estimates are made for GaAs films excited by pulses of radiation from a carbon dioxide laser. © *1997 American Institute of Physics.*
[S1063-7761(97)00807-X]

## 1. INTRODUCTION

The high-order harmonic generation (HHG) effect, in which in atomic gases the intensity of the harmonics slowly decreases with increasing harmonic number, has been the object of numerous studies, both experimental and theoretical (see, e.g., Refs. 1–4). The main interest here lies in the possibility of a sizable anti-Stokes transformation of the incident laser radiation into hard UV radiation or even soft x-ray radiation.

The number density of the emitting atoms plays an important role in the intensity of high-order harmonics. Recently there have been reports on observations of high-order harmonics in the reflection of intense monochromatic radiation from a solid surface.[5] The main parameter in the theory of the HHG effect is the ponderomotive energy

$$\varepsilon_F^{(0)}=\frac{e_0^2F^2}{4m_0\omega^2},$$

where $e_0$ and $m_0$ is the electron charge and mass, and $F$ and $\omega$ are the amplitude and frequency of the exciting laser radiation. The maximum harmonic is characterized by a number $N_{\max}$ corresponding to the end of the plateau on the curve representing the dependence of the logarithm of intensity on the harmonic's number, which constitutes a cutoff effect:

$$N_{\max}=\frac{1}{\hbar\omega}(I+(2-3)\varepsilon_F^{(0)}), \tag{1}$$

where $I$ is the ionization potential of the atom.

In the present paper the HHG problem is analyzed for nonmetallic crystal films. Usually in crystals, the quantity $\varepsilon_F$, in which the free electron mass $m_0$ is replaced by the effective current-carrier mass $\mu$, acts as the ponderomotive potential. Since $\mu$ can be smaller than $m_0$ by a factor of several tens, the shorter the pulses of laser excitation, the larger the parameter $\varepsilon_F$.

Obviously, to observe the harmonics the films must be thin. As we will shortly see, for GaAs crystals and exciting radiation from a carbon dioxide laser, it is possible to observe the 95th harmonic in films with thickness 300–1000 Å and a peak field strength $F\simeq 4\times 10^6$ V cm$^{-1}$.

The emission of photoelectrons due to absorption of harmonics, with a distribution close to the intensity distribution of the harmonics, is also predicted. The method of calculating the intensity of harmonics is based on the original ideas of calculating the multiphoton photoelectric effect in nonmetallic crystals proposed in the landmark paper of Keldysh[6] and in the paper of the present author and Sedletsky on the HHG effect in atoms with a hydrogen-like ground state.[7]

## 2. GENERAL EXPRESSION FOR THE INTENSITY OF HARMONIC GENERATION IN BAND-TO-BAND TRANSITIONS

Consider a crystal with band gap $\Delta_{cv}$ in the Kane model. Suppose that a direct optical transition at $\mathbf{k}=0$ is allowed, and that a linearly polarized wave of frequency $\omega$ and amplitude $\mathbf{F}\|X$, where $X$ is a chosen axis of the crystal, is incident on the crystal. We assume that

$$e_0F|X_{cv}|<\Delta_{cv},$$

where $X_{cv}$ is the band-to-band transition matrix element, and $\Delta_{cv}$ is the band gap. The wave function of the system in the coordinate representation depends on the electron and hole radius vectors, $\mathbf{r}_e$ and $\mathbf{r}_h$ $(\mathbf{r}_e,\mathbf{r}_h\equiv\mathbf{r})$.

If we allow for the interaction of the external electromagnetic field with the electron and hole, the appropriate Schrödinger equation has the form

$$i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t}=[H_0(\mathbf{r})+H^d(\mathbf{r},t)+H^{nd}(\mathbf{r},t)]\Psi(\mathbf{r},t). \tag{2}$$

Here $H_0(\mathbf{r})$ is the crystal Hamiltonian without an external field, and $H^d(\mathbf{r},t)$ and $H^{nd}(\mathbf{r},t)$ are the diagonal (in indices $c$ and $v$) and off-diagonal (mixing the bands $c$ and $v$) parts of the interaction with the external electromagnetic field.

We assume that in the ground state (a completely filled valence band and a vacant conduction band), the electron and hole wave function can be chosen in the form $\psi_0(\mathbf{r})=\delta(\mathbf{r}_e-\mathbf{r}_h)$. In what follows, we assume that interaction with the electromagnetic wave is taken into account by $H^d$ exactly, while $H^{nd}$ is allowed for only in the lowest perturbation order. The adopted model simplifies the calculation of the average dipole moment $\overline{X}(t)$ considerably:

$$\overline{X(t)}=\int \Psi^*(\mathbf{r},t)X\Psi(\mathbf{r},t)dr. \tag{3}$$

The observable quantity is the intensity $|X_\Omega|^2$ of a harmonic at frequency $\Omega$:

$$X_\Omega=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-i\Omega t}\overline{X(t)}dt. \tag{4}$$

We use the Furry representation ($f$ ).[8,9] We select $\Psi(t)$ in the form

$$\Psi(t)=u(t)\psi_f(t). \tag{5}$$

Here

$$i\hbar\frac{\partial u}{\partial t}=[H_0+H^d(t)]u, \tag{6}$$

$$u(t)=T\ \exp\left\{-\frac{i}{\hbar}\int_{-\infty}^{t}(H_0+H^d(t_1))dt_1\right\}, \tag{7}$$

$$i\hbar\frac{\partial\psi_f}{\partial t}=[H^{nd}(t)]_f\psi_f(t), \tag{8}$$

$$[H^{nd}(t)]_f=u^{-1}(t)H^{nd}(t)u(t), \tag{9}$$

$$\psi_f(t)=S(t,-\infty)\psi_f(-\infty). \tag{10}$$

where

$$S(t,-\infty)=T\ \exp\left\{-\frac{i}{\hbar}\int_{-\infty}^{t}[H^{nd}(t_1)]_f\,dt_1\right\}, \tag{11}$$

$$\psi_f(-\infty)=\psi_0(\mathbf{r}). \tag{12}$$

We introduce the Green's function

$$G(\mathbf{r},t;\mathbf{r}',t')=-\frac{i}{\hbar}\theta(t-t')\delta(\mathbf{r}-\mathbf{r}')u(t)u^{-1}(t'). \tag{13}$$

Clearly, the expression for $\overline{X(t)}$ becomes

$$\overline{X(t)}=2\ \mathrm{Im}\left\{\int_{-\infty}^{\infty}dt'\langle\psi_0|H^{nd}(t)\hat{G}(\mathbf{r},t;\mathbf{r}'t')H^{nd}(t')|\psi_0\rangle\right\},$$

$$\hat{G}(\mathbf{r},t;\mathbf{r}',t')=-\frac{i}{\hbar}\ \theta(t-t')\sum_{\mathbf{k}}\ \varphi_{\mathbf{k}(t)}(\mathbf{r},t)\varphi^{+}_{\mathbf{k}(t)}(\mathbf{r}',t'),$$

$$\frac{1}{\mu}=\frac{1}{m_e}+\frac{1}{m_c}, \tag{14}$$

where

$$\varphi_{\mathbf{k}(t)}(\mathbf{r},t)=\psi^{c}_{\mathbf{k}(t)}(\mathbf{r}_e,t)\psi^{v}_{\mathbf{k}(t)}(\mathbf{r}_h,t),$$

with $\psi^{c}_{\mathbf{k}(t)}(\mathbf{r},t)$ and $\psi^{v}_{\mathbf{k}(t)}(\mathbf{r},t)$ solutions of the Volkov type for an electron and hole in the external electromagnetic field,[6] and

$$\mathbf{k}(t)=\mathbf{k}+\frac{e\mathbf{F}}{\hbar\omega}\sin\ \omega t,$$

$$\psi^{c,v}_{\mathbf{k}(t)}(\mathbf{r},t)=\exp[i(\mathbf{k}(t)\mathbf{r})]U^{(c,v)}_{\mathbf{k}(t)}(\mathbf{r})\exp\left\{-\frac{i}{\hbar}\int_0^t\varepsilon^{c,v}_{\mathbf{k}(t_1)}dt_1\right\}, \tag{15}$$

with $U^{c,v}_{\mathbf{k}(t)}(\mathbf{r})$ the periodic parts of the Bloch function. As a result we find that

$$\overline{X(t)}=2\ \mathrm{Im}\Bigg[-\frac{i}{\hbar}\frac{eF}{\omega\mu}\int dt'\sum_{\mathbf{k}}\ \theta(t-t')\times\langle v\mathbf{k}(t)|X|c\mathbf{k}(t)\rangle\left\langle c\mathbf{k}(t')\right|-i\hbar\frac{\partial}{\partial x}\left|v\mathbf{k}(t')\right\rangle\times\sin\ \omega t'\ \exp\left\{-\frac{i}{\hbar}\int_t^{t'}\varepsilon_{\mathbf{k}}(t_1)dt_1\right\}\Bigg], \tag{16}$$

where $\varepsilon_{\mathbf{k}(t)}$ describes the dispersion law, which in the Kane model is

$$\varepsilon_{\mathbf{k}(t)}=\varepsilon^{c}_{\mathbf{k}(t)}-\varepsilon^{v}_{\mathbf{k}(t)}=\sqrt{\Delta_{cv}^2+\frac{\hbar^2\mathbf{k}^2(t)}{\mu}\Delta_{cv}}. \tag{17}$$

The band-to-band matrix elements of position and momentum in the external field are calculated, as usual, over the unit cell volume. Since the transverse momentum $\mathbf{p}_\perp=\hbar\mathbf{k}_\perp$ is not perturbed by the electromagnetic field, the main contribution to the sum over $\mathbf{k}$ in (16) is provided by the term with $p_\perp^2=0$ (this follows directly from (17), since $p_\perp^2$ enters into (17) together with $\Delta_{cv}$). As we will shortly see, the dependence of the probability of harmonic generation on the band gap $\Delta_{cv}$ leads to a sharp (exponential) decrease in the rate of the process as $\Delta_{cv}$ grows. Hence from now on we assume $p_\perp^2=0$. The position matrix element for a band-to-band transition ($p_\perp^2=0$) has the following form in the Kane model:[10]

$$\langle v\mathbf{k}(t)|X|c\mathbf{k}(t)\rangle\equiv X_{12}(t)=\frac{i\hbar}{2}\ \sqrt{\frac{\Delta_{cv}}{\mu}}\times\frac{1}{\Delta_{cv}+(\hbar^2/\mu)(k_x+(eF/\hbar\omega)\sin\ \omega t)^2}. \tag{18}$$

Accordingly, the momentum matrix element is

$$p_{12}=\mu\frac{\Delta_{cv}}{\hbar}X_{12}. \tag{19}$$

Thus, we must calculate

$$\overline{X(t)}=2\ \mathrm{Im}\Bigg(-\frac{i}{\hbar}\frac{eFX_{12}(t)}{\omega\mu}\times\int_{-\infty}^{\infty}dt'\sum_{k_x}\ \theta(t-t')p_{12}(t')\sin\ \omega t'\times\exp\left\{-\frac{i}{\hbar}\int_t^{t'}\varepsilon_{k_x(t_1)}dt_1\right\}\Bigg). \tag{20}$$

Here an exact calculation is difficult, so we adopt an approximate approach based on taking the value of the preexponential factor $p_{12}(t')$ at the saddle point $t'_*$ calculated at $k_x=0$ outside the sign of integration over $t'$. The fact that the factor is calculated at $k_x=0$ means that the term with $k_x=0$ provides the main contribution to the sum over $k_x$. This statement is corroborated by the calculations below. The function

$$f(t')=\pm it'\omega-\frac{i}{\hbar}\int_t^{t'}\varepsilon_{k_x(t_1)}dt_1 \tag{21}$$

has stationary points, so that

$$i\omega-\frac{i}{\hbar}\varepsilon_{k_x(t)}=0. \tag{22}$$

At $k_x=0$ we have

$$t'_*=\frac{i}{\omega}\sinh^{-1}\gamma,\quad \gamma=\frac{\omega}{\omega_{\rm tun}},$$

$$\omega_{\rm tun}=e_0F\sqrt{\frac{\Delta_{cv}/\mu}{\Delta_{cv}^2-\hbar^2\omega^2}}, \tag{23}$$

where the tunneling frequency $\omega_{\rm tun}$ is determined by the tunneling transition from the valence band to the conduction band in the field of the external electromagnetic wave. In the adopted adiabatic limit $\hbar\omega\ll\Delta_{cv}$, the expression for $\omega_{\rm tun}$ can be written as

$$\omega_{\rm tun}=\omega_{\rm tun}^{(0)}\left(1+\frac{1}{2}\left(\frac{\hbar\omega}{\Delta_{cv}}\right)^2\right),\quad \omega_{\rm tun}^{(0)}=\frac{e_0F}{\sqrt{\mu\Delta_{cv}}}. \tag{24}$$

In (20) it is convenient to replace the summation over $k_x$ by integration over $k_x$:

$$\sum_{k_x}\cdots=\frac{1}{2\pi}\int\cdots dk_x.$$

The integral over $k_x$ contains a simple pole. Indeed, the matrix element $p_{12}(t'_*)$ has the poles

$$k_x^{(1)}=ik_0\frac{\kappa^2}{2\sqrt{2}},\quad k_x^{(2)}=-ik_0\sqrt{2},$$

$$\frac{\hbar^2k_0^2}{2\mu}=\Delta_{cv},\quad \kappa^2=\left(\frac{\hbar\omega}{\Delta_{cv}}\right)^2. \tag{25}$$

The integral over $k_x$ is calculated in the Appendix. Thus, the integral over $k_x$ can be calculated directly, and in calculating $X_\Omega$ we can replace $k_x$ by $k_x^{(1)}$ in the integrals with respect to $t$ and $t'$. Since $|k_x^{(1)}/k_0|$ is much smaller than unity, we can, to make matters simple, set it equal to zero. This was the condition for taking the pre-exponential factor at the point $t'_*$ calculated at $k_x=0$ outside the integral. Further calculations of the integral over $t'$ are done by simplifying the expression containing a square root in the exponential:

$$j=\frac{i}{\hbar}\int_{t'}^{t}dt_1\sqrt{\Delta_{cv}^2+\Delta_{cv}\frac{e^2F^2}{\omega^2\mu}\sin^2\omega t_1}$$

$$\simeq i\left\{(q+\rho)\omega(t-t')-\frac{\rho}{2}(\sin 2\omega t-\sin 2\omega t')\right\},$$

$$q=\frac{\Delta_{cv}}{\hbar\omega},\quad \rho=\frac{\varepsilon_F}{\hbar\omega},\quad \varepsilon_F=\frac{e_0^2F^2}{4\mu\omega^2}. \tag{26}$$

Note that there can be no objection to expanding the root in (26) after all singularities are taken into account. This corresponds to the way in which the conduction and valence bands are treated in the model of parabolic dispersion laws.

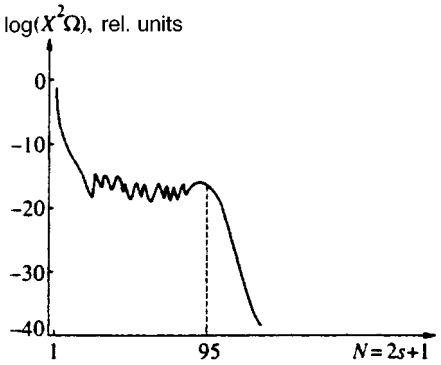



FIG. 1. Harmonic intensity (in relative units) as a function of the harmonic's number ($q=15$ and $\rho=40$).

Note that the expression inside the braces in (16) is the classical action, and finding $X_\Omega$ leads to the energy conservation law[4]

$$\hbar\Omega=\varepsilon_{\mathbf{k}(t)}. \tag{27}$$

This condition can be used in the matrix element $X_{12}(t)$.

Let us employ the expansion

$$e^{i\rho\sin z}=\sum_{n=-\infty}^{+\infty}J_n(\rho)e^{inz}, \tag{28}$$

where $J_n(\rho)$ is the Bessel function of a real argument.

Now we can easily integrate with respect to $t$ and $t'$. We find

$$X_\Omega=X_0\sqrt{\rho}\left(\sum_{m=-\infty}^{\infty}\left[J_m\left(\frac{\rho}{2}\right)J_{m-s-1}\left(\frac{\rho}{2}\right)+J_m\left(\frac{\rho}{2}\right)J_{m-s}\left(\frac{\rho}{2}\right)\right]\frac{1}{m-s+a}\right),$$

$$X_0=\frac{\pi}{8}\frac{\Delta_{cv}^{5/2}\sqrt{\hbar\omega}}{\hbar^3\Omega^2\omega^2}\sqrt{\frac{\hbar}{\mu\omega}},$$

$$\Omega=(2s+1)\omega,\quad a=\frac{1}{2}(q+\rho-1). \tag{29}$$

Next we use the well−known formula[11]

$$\sum_m J_m(x)J_{m+s}(x)\frac{1}{m+a}=\frac{\pi}{\sin\pi a}J_{s-a}(x)J_a(x), \tag{30}$$

where $a$ is an integer.

Thus we find that

$$X_\Omega=\frac{X_0\pi\sqrt{\rho}}{\sin a\pi}\left[J_{a+1}\left(\frac{\rho}{2}\right)J_{s-a}\left(\frac{\rho}{2}\right)-J_a\left(\frac{\rho}{2}\right)J_{s-a}\left(\frac{\rho}{2}\right)\right]. \tag{31}$$

This formula implies that for $\rho\ll1$ the intensity of the $(2s+1)$st harmonic is

$$|X_{2s+1}|^2\sim\rho^{2s+1}\sim(F^2)^{2s+1},$$

which corresponds to the result obtained by applying perturbation-theory techniques. Here the harmonics for which $\rho/2\simeq s-a$ are the most interesting ones. The curve

representing the dependence of the harmonic's intensity on the harmonic's number has a plateau (Fig. 1), which extends to the highest harmonic,

$$N_{\max}=2s_{\max}+1=\frac{1}{\hbar\omega}(\Delta_{cv}+2\varepsilon_F),$$

corresponding to the cutoff effect.

### 3. DISCUSSION

A specific feature of solids is that they can absorb the generated harmonics. The harmonics can be observed in reflected light, as was done by von der Linde *et al.*[5] But if we are dealing with thin films, the spectrum of the generated harmonics can be observed by transillumination in the ranges where $\hbar\Omega \gg \Delta_{cv}$. For instance, for GaAs we have $\Delta_{cv}=1.5$ eV and $\mu \simeq 0.06m_0$. The crystal is transparent to the radiation of a carbon dioxide laser with $\hbar\omega \simeq 0.1$ eV. The calculation done in Ref. 5 is valid for fields $F \ll F_{\rm cr}$, with

$$F_{\rm cr}=\frac{\Delta_{cv}}{e_0|X_{12}|}\simeq 3\times 10^7 \text{ V/cm}. \tag{32}$$

Figure 1 depicts the dependence of the harmonic intensity for generation of harmonics in a field $F=4\times 10^6$ V/cm ($\rho=40$ and $q=15$) on the harmonic's number. We see that the highest harmonic has the number $N_{\max}$, which means that $\hbar\Omega=9.5$ eV. The absorption coefficient at this wavelength in GaAs is $10^5$–$10^6$ cm$^{-1}$ (see Ref. 12).

Thus, to observe this harmonic we need a film with a thickness 300–1000 Å. A convenient way of recording the harmonics is to study the photoemissive effect they produce. For GaAs the photoemission threshold is $\eta=5.5$ eV (see Ref. 13), which corresponds to a threshold harmonic with a number $N_{\rm th}=55$. The absorption coefficient in the interval between the 55th and 95th harmonics, which is capable of causing a photoemissive effect in the crystal, varies relatively weakly. As a result, the energy distribution of the photoelectrons generally follows the harmonic intensity distribution (see Fig. 1).

Note that a direct multiphoton photoelectric effect in the crystal determines the energy of an emitted electron in terms of the free electron mass $m_0$. Here, if we employ the Volkov solutions for a free electron, the corresponding cutoff effect for the energy of the emitted electron is clearly given by

$$\varepsilon=\eta+2\varepsilon_F^{(0)}, \quad \varepsilon_F^{(0)}=\frac{e_0^2F^2}{4m_0\omega^2}. \tag{33}$$

Since the ponderomotive potentials $\varepsilon_F^{(0)}$ and $\varepsilon_F$ differ by a factor of $\mu/m_0$, so will the values of the energies of the emitted electron for these two competing mechanisms of free photoelectron formation.

Naturally, a fraction of the harmonics that land in the band gap will freely leave the crystal. In our example of a GaAs crystal, for a photon energy $\hbar\omega=0.05$ eV and a field strength $1.4\times 10^6$ V cm$^{-1}$ the crystal will emit harmonics in the plateau region with numbers ranging from the 20th to the 30th.

I am grateful to E. Yu. Kanarovskiĭ for his useful remarks.

### APPENDIX

Consider the integral

$$I=\int_{-\infty}^{\infty}dz\,\frac{F(z)}{(z-i\alpha_1)(z+i\alpha_2)},$$

$$F(z)=\theta(t-t')f(z)\times\exp\left\{iB\int_{t'}^{t}\sqrt{z^2+zA(t_1)+C(t_1)}\,dt_1\right\},$$

where $F(z)$ has no poles and decreases as a power function at infinity. We examine the pole in the upper half-plane at $z=i\alpha_1$. The residue at this point requires that the Jordan lemma be valid. We close the integration contour in the upper half-plane and examine the behavior of the integrand $F(z)/(z+i\alpha_2)$ at $z=|z|e^{i\varphi}$:

$$|z|\to\infty, \quad \varphi\in[0,\omega].$$

Clearly, $F(z)$ with $|z|\to\infty$ is

$$\theta(t-t')\exp\{i(t-t')|z|e^{i\varphi}\}\to 0.$$

(One can easily verify that on the rays $\varphi=\pi/4$ and $3\pi/4$ the integrand in the integral with respect to $t_1$ is also independent of $t$ as $|z|\to\infty$.) The integrals for the complex conjugate function, where the contour is closed in the lower half-plane, can be calculated in a similar manner.


[1] A. L'Huillier, L. A. Lompre, G. Mainfray, and C. Manus, in *Proc. 5th Intern. Conf. on Multiphoton Processes*, Paris (1990).
[2] J. L. Krause, K. J. Schafer, and K. C. Kulander, Phys. Rev. Lett. **68**, 3535 (1992).
[3] P. B. Corkum, Phys. Rev. Lett. **71**, 1994 (1993).
[4] M. Lewinstein, Ph. Balcou., M. Yu. Ivanov, A. L'Huillier, and P. B. Corkum, Phys. Rev. A **49**, 2117 (1994).
[5] D. von der Linde, T. Engers, A. Jenke, P. Agostini, G. Grillon, E. Nibbering, A. Mysyrowicz, and A. Antonetti, Phys. Rev. A **52**, R25 (1995).
[6] L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1945 (1964) [Sov. Phys. JETP **20**, 1307 (1965)].
[7] V. A. Kovarsky and O. A. Sedletsky, in *Proc. 7th Intern. Conf. on Multiphoton Processes*, Garmisch-Partenkirchen, Germany (1996).
[8] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 2nd ed., Pergamon Press, Oxford (1982).
[9] A. Kovarskiĭ, N. F. Perel'man, and I. Sh. Averbukh, *Multiquantum Processes* [in Russian], Énergoatomizdat, Moscow (1985).
[10] Yu. A. Bychkov and A. M. Dykhne, Zh. Éksp. Teor. Fiz. **58**, 1734 (1970) [Sov. Phys. JETP **31**, 928 (1970)].
[11] V. A. Kovarskiĭ and N. F. Perel'man, Izv. Akad. Nauk Mold. SSR, Ser. Fiz. Tekh. Matem. No. 1, 37 (1973).
[12] B. Seraphin and H. Bennett, in *Semiconductors and Semimetals*, R. K. Willardson and A. C. Beer (eds.), Vol. 3: *Optical Properties of III–V Compounds*, Academic Press, New York (1967).
[13] *Tables of Physical Quantities* [in Russian], E. K. Kikoin (ed.), Atomizdat, Moscow (1976).


Translated by Eugene Yankovsky

# Polarization characteristics of the ''forbidden'' second optical harmonic of femtosecond laser pulses in a bacteriorhodopsin solution

A. V. Balakin, N. I. Koroteev, A. V. Pakulev, and A. P. Shkurinov

*International Laser Center and Physics Department, M. V. Lomonosov Moscow State University, 119899 Moscow, Russia*

D. Boucher, P. Masselin, and E. Fertein

*Laboratoire de Physico Chimie de l'Atmosphère, Université du Littoral, 59379 Dunkerque, Cedex 1, France*
(Submitted 9 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 97–114 (July 1997)

The generation of the second harmonic of femtosecond laser pulses in a bacteriorhodopsin solution has been experimentally studied for various wavelengths and polarization states of radiation at the fundamental frequency. The polarization properties of the effect are analyzed under various experimental conditions. The nature and properties of the signal are treated as the manifestation of a superposition of nonlinear optical effects of various orders (the second and the fourth). The second-order effects can have both an electric-dipole and a magnetic-dipole or electric-quadrupole character. In analyzing fourth-order processes, besides the direct electric-dipole contribution, the possibility of the participation of cascade processes at second- and third-order nonlinearities is also allowed. © *1997 American Institute of Physics.* [S1063-7761(97)00907-4]

## 1. INTRODUCTION

Nonlinear second-order optical processes such as the generation of the second optical harmonic and the generation of radiation at sum and difference frequencies make it possible to obtain unique information concerning the internal structure of crystals, thin films, surfaces, and interfaces of both centrosymmetric and noncentrosymmetric materials, as well as concerning laser-induced processes on the surface.[1,2]

From the standpoint of a phenomenological description of second-order nonlinear optical processes, one important factor is the symmetry of the medium. Thus, inside an isotropic centrosymmetric (i.e., invariant under inversion) medium, even-order nonlinear processes are forbidden in the electric-dipole approximation.[3] This forbiddenness can be removed at the interface between two media, where the symmetry breaks down,[4] while inside centrosymmetric materials, generation of the second optical harmonic and generation of radiation at the sum and difference frequencies can appear only because of ''nonlocal''—quadrupole and magnetic-dipole—interactions.[1,2,5,6] In homogeneous isotropic media in which there is no macroscopic inversion center (i.e., in noncentrosymmetric media, including isotropic media such as optically active liquids), nonlinear processes generated by even-order dipolar optical susceptibilities become possible.[7] Noncentrosymmetric properties are possessed by many molecular systems in which intrinsic (internal) asymmetry exists,[8] including solutions of natural biological macromolecules with helical structure.[9]

Substantial interest in the study of the asymmetry properties of complex molecules by nonlinear-optics methods has resulted in recent years in a number of interesting experimental[10,11] and theoretical[12,13] papers on second-harmonic generation from the free surfaces of solutions containing noncentrosymmetric molecules and native optically active liquids. It has been shown that second-harmonic generation in this case can be sensitive to the noncentrosymmetry of the medium being studied, because of the appearance of specific components of the nonlinear second-order susceptibility tensor $\chi^{(2)}$ that are absent in a centrosymmetric medium.

Even-order nonlinear optical processes in an isotropic medium that were sensitive to its noncentrosymmetry (gyrotropy, chirality) were observed for the first time in nonracemic solutions of *d* and *l* arabinose, and were associated with electric-dipole nonlinear optical susceptibilities of second order (generation of radiation at a sum frequency)[14] and fourth order (generation of the second harmonic in noncollinear beams).[15]

Our previous papers[16,17] described for the first time the generation of the second harmonic of femtosecond laser radiation in a finely dispersed suspension of purple membranes (with fragment sizes less than 50 nm) containing bacteriorhodopsin. It was shown from an analysis of the dependences of the intensity of the second-harmonic signal on the intensity of the fundamental-frequency pulses that the recorded signal is caused by bulk nonlinear processes of both second and fourth order and their interference.

Second-harmonic generation in an aqueous suspension of large fragments of purple membranes having a characteristic size of up to 1 $\mu$m was reported in Refs. 18 and 19, in which the recorded process was explained by hyper-Rayleigh scattering resulting from high nonlinear second-order polarizability of each fragment of the membrane as a whole, and also in Ref. 20, in which, on the contrary, it was noted that the recorded process is not associated with hyper-Rayleigh scattering. The nature of the emergence of the second-harmonic signal in a bacteriorhodopsin suspension has been the subject of intense discussion in the literature.[21,22] However, the experimental and theoretical studies have yet to

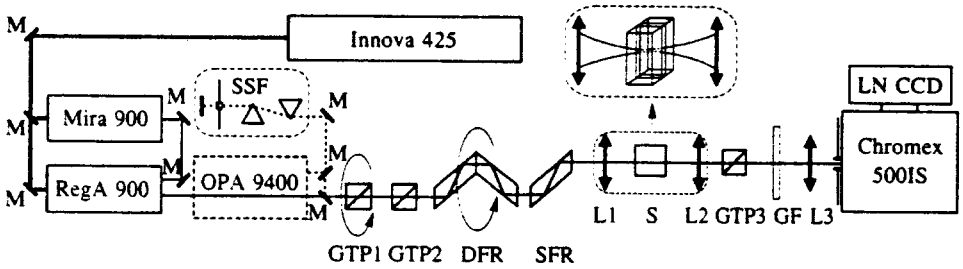



FIG. 1. Layout of experimental apparatus. GTP are Glan prisms, DFR is a double Fresnel rhomb, SFR is a single Fresnel rhomb, SSF is a spatial spectral filter, M are mirrors, L are lenses, GF is a glass color filter, S is the sample, and LN CCD is a CCD camera.

produce a definitive picture of the phenomenon, which has come to be called in the literature the generation of the forbidden second harmonic.

The bacteriorhodopsin molecules are grouped in the membrane into unit cells with three molecules apiece (a trimer), while the individual cells are arranged in a hexagonal two-dimensional lattice.[23] Since the size of a single trimer does not exceed 6 nm, each membrane fragment in the references indicated above contained about 50 000 bacteriorhodopsin molecules. It can be assumed that the second-harmonic radiation from molecules localized in an individual membrane was spatially coherent in this case, while the radiation generated by different membranes added incoherently. Each membrane fragment can be regarded as a single scattering center whose hyperpolarizability is proportional to the mean number of bacteriorhodopsin molecules found in it. Coherent generation of the second harmonic of monochromatic radiation within a homogeneous isotropic medium is strictly forbidden in the electric-dipole approximation even when the medium is macroscopically noncentrosymmetric.[7] With respect to the generation of the second harmonic of circularly polarized beams, a ''double forbiddenness'' is active in such media—in addition to the first, there is a forbiddenness associated with the impossibility of conserving the momentum of a pulse of optical radiation when the second and higher harmonics are generated.[24,25]

This paper, which is a continuation of Refs. 16 and 17, discusses the polarization dependence of the generation of the second harmonic of femtosecond pulses in a finely dispersed suspension of fragments of purple membranes. The use of femtosecond pulsed laser radiation made it possible to increase the peak radiation power while keeping the energy of the light pulses at a level low enough not to damage the test object. The characteristics of the processes under investigation differ radically from the basic characteristics of hyper-Rayleigh scattering and can be interpreted only by taking into account the fourth-order electric-dipole nonlinear-optical susceptibility of the medium and second-order nonlocal processes.

## 2. EXPERIMENTAL APPARATUS AND SAMPLE PREPARATION

A diagram of the experimental apparatus is shown in Fig. 1. Laser radiation of femtosecond width was formed and amplified by means of a Mira 900 solid-state titanium–sapphire laser and a RegA 9000 regenerative amplifier, pumped with all the lines of an Innova 425 $Ar^+$ laser. Depending on the experimental conditions, we used light pulses either obtained directly from the output of the RegA 9000 or converted by means of an OPA 9400 optical parametric amplifier (shown in Fig. 1 with a dashed outline). In the latter case, the femtosecond continuum obtained by focusing the amplified pulses of the master oscillator into a sapphire plate was used as trigger radiation. All the laser equipment is produced by Coherent, Inc. When the parametric amplifier was used, the experimental system also included a two-prism spatial spectral filter SSF (shown in the figure with a dashed outline), which suppressed the residues of the continuum radiation. The filter uses a two-pass scheme with zero dispersion, and the input and output radiation were decoupled in the vertical plane.

The polarization and energy characteristics of the femtosecond radiation were controlled by a common scheme in both cases. A combination of two Glan–Taylor prisms (GTP1 and GTP2) was used to isolate vertical plane polarization and to smooth the variation of the radiation energy. The light beam remained plane polarized as it passed through a double Fresnel rhomb (DFR), while the angle $\varphi$ between the polarization orientation direction of the light wave and the vertical plane containing the beam axis was varied by rotating the rhomb around an axis coinciding with the propagation direction of the ray. A single Fresnel rhomb (SFR) was used to vary the degree of ellipticity of the radiation. Rhomb SFR was oriented in such a way that the complex vector of the electric field of the transmitted light wave varied according to the law

$$\mathbf{E}(\varphi)=A(\mathbf{e}_x \cos \varphi+\mathbf{e}_y i \sin \varphi), \tag{1}$$

where $A$ is the field amplitude, and $\mathbf{e}_x$ and $\mathbf{e}_y$ are, respectively, vertical and horizontal unit vectors forming a plane perpendicular to the propagation axis of the beam.

Thus, according to Eq. (1), the radiation had linear vertical polarization when $\varphi=0°$, which changed to right-circular polarization for $\varphi=45°$, next to linear horizontal polarization for $\varphi=90°$, then to left circular polarization for $\varphi=135°$, and again to linear vertical polarization for $\varphi=180°$. The intermediate values of angle $\varphi$ corresponded to different states of the degree of ellipticity and orientation of the polarization ellipse.

At the output of SFR, the laser radiation had the following parameters: the wavelength was 820 nm, the spectral width was 8 nm, the pulsewidth at half-height was about 250 fs, the energy per pulse was adjusted over the range 0–1.2 $\mu$J, and the pulse repetition rate was 200 kHz. When the optical parametric amplifier was included in the setup, we could obtain at the output of the SFR light pulses tunable

over the spectral range 470–700 nm, having a width of about 150 fs at half-height, a maximum energy of 45 nJ, and a repetition rate of 200 kHz.

Focusing lens L1, with a focal length of 17 cm, provided a beam convergence angle of about 2° inside the medium of interest in the first case and about 1° in the second. Assuming a Gaussian beam and allowing for the initial divergence, the beam-waist diameters at the focus of the lens were 20 $\mu$m and 30 $\mu$m, respectively. Fused-quartz spectroscopic cell S, in which the test solution was placed, had internal dimensions of $10\times 2.5$ mm, and focusing took place along the long side of the cell. Using quartz lens L2, the radiation to be recorded at the second-harmonic frequency was collimated and directed into the analyzer—a Glan–Taylor prism (GTP3). Prism GTP3 was oriented in such a way that the recorded second-harmonic signal was minimal when the orientation of the polarization vector of the exciting radiation was horizontal. The use of the analyzer and the choice of its orientation made it possible to eliminate the influence of the polarization dependence of the reflectance of the spectrograph gratings.

The second-harmonic radiation was next isolated with the help of suitable glass filters (GF) and was focused by lens L3, with focal length 5 cm, onto the entrance slit of a Chromex 500IS spectrograph. An SZS 21 filter was used when the sample was excited by radiation with a wavelength of 820 nm, and a UFS 1 filter was used for excitation by 590-nm radiation. The spectrum of the second-harmonic signal was recorded by means of a Princeton Instruments Inc. CCD camera (LN CCD) cooled with liquid nitrogen.

A specially prepared finely dispersed suspension of fragments of the purple membranes of *Halobacterium Halobium*, containing bacteriorhodopsin, was used for the experiments. To reduce the effects associated with structural ordering of the placement of the bacteriorhodopsin in the membranes and to maximize the isotropic distribution of the molecules in the sample, the membrane fragments, prepared by the standard method,[23] were fragmented by means of additional ultrasound processing with multistep precipitation and filtering.[16] The resulting extract was prepared in a phosphate buffer (PBS), which kept the acidity of the solution at $pH=7.35$.

The quality of the suspension was monitored via microscopic measurements of the characteristics of a monolayer film obtained by precipitating the prepared sample onto a quartz substrate, with drying at constant humidity. The characteristic size of the membrane fragments was estimated by means of an atomic-force microscope[26]: for the sample used in the experiments, it was at most 50 nm (the accuracy with which the size was measured was determined by the resolution of the microscope, and equalled 50 nm). Immediately after preparation, the suspension contained not only pulverized but also relatively large particles (more than 50 nm across), which precipitated to the bottom of the cell after a sufficiently long delay as the concentration stabilized.

The precipitation dynamics of the large fragments was studied by recording the absorption spectra of the sample with a spectrophotometer every 6 min, beginning at the time of preparation. Figure 2 shows the spectra recorded immediately after the suspension was placed in the cell (solid curve) and 1.5 h later, after natural precipitation of the large particles (dashed curve). The presence of relatively large membrane fragments in the suspension increased the optical density of the sample uniformly over the entire spectrum, because it increased the concentration of bacteriorhodopsin molecules. The kinetics of optical absorption had an exponential character, with a characteristic time of 1–1.5 hr to reach the steady state.

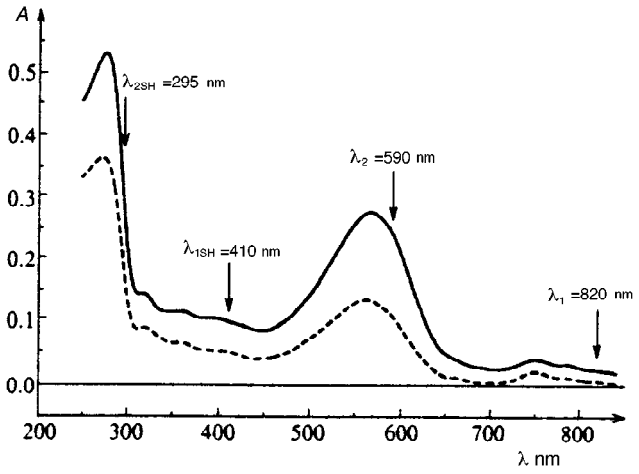



FIG. 2. Absorption spectra of an unprecipitated sample (solid curve) and a completely precipitated sample (dashed curve) containing bacteriorhodopsin. *A* is the optical density, and λ is the wavelength in nanometers.

A naturally precipitated sample with properties close to those of a colloidal solution was used in subsequent experiments. The absorption spectrum of the sample was similar to the spectrum shown in Fig. 2 as a dashed curve. The maximum absorptance in the visible region ($\lambda=568$ nm) was at most 10% when the cell was 1 cm long, which corresponded to a concentration of bacteriorhodopsin molecules of about $10^{-6}$ M.

### 3. EXPERIMENTAL RESULTS

For this paper, we experimentally studied the polarization dependence of the second-harmonic signal intensity in a suspension of purple membranes, both off-resonance at the fundamental frequency ($\lambda_\omega=820$ nm), and with resonant excitation ($\lambda_\omega=590$ nm). The wavelength of the fundamental and the second harmonic for both cases is marked in Fig. 2 by vertical arrows. The absorption of the test sample can be considered negligible at $\lambda_\omega=820$ nm, but it is nonzero at the second-harmonic wavelength ($\lambda_{2\omega}=410$ nm). Absorption is strong at $\lambda_\omega=590$ nm and the corresponding second-harmonic wavelength $\lambda_{2\omega}=295$ nm, which can easily be seen from Fig. 2.

Figure 3 shows how the energy of the pulses of second-harmonic radiation generated in a bacteriorhodopsin suspension, normalized to the maximum value, depends on the polarization state of the pump radiation upon entering the test medium for nonresonant irradiation of the sample ($\lambda_\omega=820$ nm). The different curves correspond to different experimental conditions.

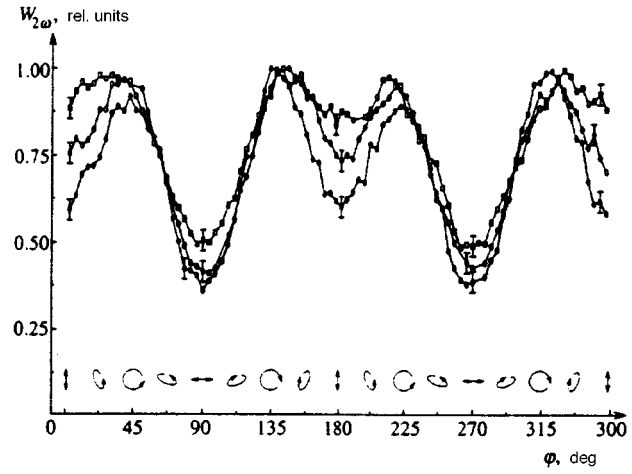



FIG. 3. Energy of second-harmonic pulses in a bacteriorhodopsin solution vs. the polarization state of radiation at the fundamental frequency ($\lambda_\omega = 820$ nm) in the absence of absorption for various energies $W_\omega$ and widths $\tau_p$ of the exciting pulses: *1*—$W_\omega = 1.2$ $\mu$J, $\tau_p = 250$ fs; *2*—$W_\omega = 0.17$ $\mu$J, $\tau_p = 250$ fs; *3*—$W_\omega = 0.17$ $\mu$J, $\tau_p = 0.9$ ps. Each experimental curve is normalized to its maximum value. Angle $\varphi$ (in degrees) defines the polarization state of the input radiation according to Eq. (1).

Curve *1* was obtained using input radiation pulses with an energy of 1.2 $\mu$J/pulse. The energy has characteristic maxima for circular polarization of the initial radiation and minima when the polarization is linear. An appreciable difference of the signal intensity of the second harmonic can be seen for right ($\varphi = 45°$ and 225°) and left ($\varphi = 135°$ and 315°) circular polarization. This difference can probably be associated with macroscopic noncentrosymmetry (chirality) of the test medium, which causes the nonlinear susceptibility to be different for right- and left-circularly polarized initial radiation. Since the concentration of the solution was kept at a low level, no appreciable manifestations of linear optical rotation or circular dichroism were observed in the sample at the fundamental frequency.

Another typical feature of the polarization dependence shown here is that the signal is different for the two mutually perpendicular linear polarizations of the fundamental radiation. It should be pointed out that the difference of the second-harmonic signal from zero when the input polarization is orthogonal to the output cannot be explained by simple scattering depolarization, since the latter, as shown by special measurements that we carried out, was negligible.

In order to check whether the second-harmonic generation process is a consequence of surface effects at the interface between the test sample and the cell material, the polarization dependence measured under the same conditions was studied at 45° incidence of the exciting radiation at the input face of the cell. Surface generation of the second harmonic is forbidden for normal incidence of the input radiation at the interface between two media. Since the input beam had finite divergence, it stands to reason that an interface could contribute to the second-harmonic signal even at normal incidence. In this case, a significant increase in second-harmonic intensity should be observed as the angle of incidence increases, reaching a maximum value at 45°.

In our experiments, the maximum recorded signal did not change with oblique incidence, while the dependence of the conversion efficiency on the input polarization state was essentially identical (to within the experimental errors) to the dependence of *I* in Fig. 3. Moreover, we took special measures in all of our experiments to position the beam waist of the fundamental radiation at the center of the cell and not at the input or output window. Effects associated with nonlinear susceptibility of the interface between two media thus did not substantially influence the experimental results.

Curve *2* in Fig. 3 reflects the dependence obtained under the same conditions as for *1* but with a lower input-pulse energy (0.17 $\mu$J). It can be seen from Fig. 3 that the character of the dependence changes as the energy decreases. It should be pointed out that there are differences between the second-harmonic signal values at $\varphi = 0°$, 180°, and 360° (where the plane-polarized radiation at the input is parallel to the direction of maximum transmission of the analyzer) and $\varphi = 90°$ and 270° (with orthogonal orientation), as well as at $\varphi = 45°$ and 135° (right- and left-circular polarization of the input radiation, respectively). Decreasing the peak intensity of the pump radiation (without changing the pulsewidth and the beam size) reduces the difference between the maxima and consequently reduces the sensitivity of the experiment to the chirality of the medium.

As the peak intensity of the exciting radiation is further reduced (curve *3*), the dips in the neighborhood of $\varphi = 0°$, 180°, and 360° disappear almost completely. The differences in the second-harmonic signal amplitudes for right- and left-circularly polarized input radiation also disappear. The results shown by the given curve are obtained by means of a pulse with incompletely compensated dispersion of the group velocity (linear chirp) of the exciting radiation, having a pulsewidth of about 900 fs and the same spectral width and energy as the pulse for curve *2*. The polarization dependence represented by curve *3* is close to the analogous dependence that we obtained under the same conditions, but using nanosecond exciting radiation whose peak power was at least an order of magnitude lower.[16]

When the sample undergoes resonant excitation at a wavelength of $\lambda_\omega = 590$ nm, the dependence of the second-harmonic intensity on the polarization state of the initial radiation has an essentially different character (Fig. 4). The second-harmonic signal in this dependence is also normalized to the maximum value, which was a factor of 3–4 lower in absolute value than the maximum signal in the nonresonant case (Fig. 3, curve *1*). However, it must be emphasized that the energy of the pulses at 590 nm was at most 45 nJ; i.e., the peak intensity of the radiation in the test sample, taking into account the other parameters (shorter pulsewidth and less divergence) was two orders of magnitude less than in the nonresonant case. An increase in the efficiency of the second-harmonic generation process was thus observed in going to resonant excitation conditions. The increase in modulation depth (to 70%) shows that the observed amplification is selective with respect to the polarization state of the pump radiation upon entering the medium; it is greatest for a circularly polarized wave and least for a plane-polarized wave. Moreover, the difference in the conversion efficiency disappears for input radiation plane-polarized in the vertical

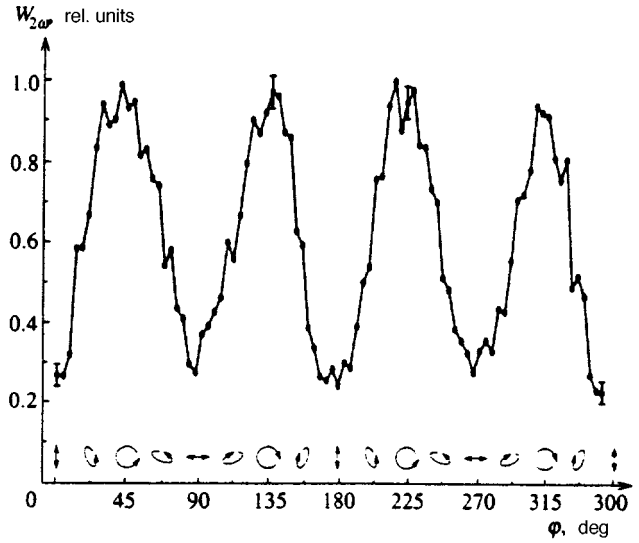



FIG. 4. Energy of second-harmonic pulses in a bacteriorhodopsin solution vs. the polarization state of the input radiation in the presence of absorption at the fundamental frequency ($\lambda_\omega = 590$ nm, $W_\omega = 45$ nJ, $\tau_p = 150$ fs). The curve is normalized to the maximum value.

($\varphi = 0°$, 180°) and horizontal ($\varphi = 90°$, 270°) planes, as well as for right-circular ($\varphi = 45°$, 225°) and left-circular ($\varphi = 135°$, 315°) polarization.

## 4. GENERAL APPROACH TO ANALYZING THE PROCESS OF GENERATING THE SUM FREQUENCY AND THE SECOND OPTICAL HARMONIC WITHIN GYROTROPIC (CHIRAL) MEDIA

A phenomenological analysis of nonlinear optical processes in macroscopically isotropic noncentrosymmetric (chiral) media was carried out in Refs. 7 and 27. A specific feature of the theoretical approach to describing the experiment reported in this paper is that the experimentally recorded second-harmonic signal cannot be ascribed to only one isolated optical nonlinearity mechanism. We showed earlier[16,17] that the second-harmonic signal in a bacteriorhodopsin solution excited with femtosecond light pulses is determined by the interference of at least two coherent contributions, due to nonlinear susceptibilities of different orders, $\chi^{(2)}$ and $\chi^{(4)}$.

We first consider possible contributions to the recorded signal by second-order nonlinear processes. In the electric-dipole approximation, the nonlinear polarization $\mathbf{P}^{NL}$ of the medium at the sum frequency $\omega_1 + \omega_2$, induced by two noncollinear plane monochromatic waves $\mathbf{E}^a(\omega_1)$ and $\mathbf{E}^b(\omega_2)$, can be represented as

$$P_i^{(2)D}(\omega_1+\omega_2) = D_2 \chi_{ijk}^{(2)D}(\omega_1+\omega_2;\omega_1,\omega_2) \times E_j^a(\omega_1) E_k^b(\omega_2), \quad (2)$$

where $E_j^a(\omega_1)$ and $E_k^b(\omega_2)$ are the vector components of the electric fields of the corresponding light waves, and $D_2$ is a factor that takes possible frequency degeneracy into account. Repeated Cartesian indices represent summation from 1 to 3.

In the case of frequency degeneracy, i.e., if the frequencies $\omega_1$ and $\omega_2$ coincide, all the components of the electric-dipole susceptibility tensor $\chi_{ijk}^{(2)D}$ go to zero as a consequence of permutation symmetry over the last two subscripts, even in a noncentrosymmetric isotropic medium.[7,28] However, with noncollinear interaction, in the nondegenerate case, when $\omega_1 \neq \omega_2$, radiation can be generated at the sum frequency because of the frequency dispersion of the second-order nonlinear electric-dipole polarizability of the molecules studied here.[28] In analyzing second-harmonic generation in a focused beam of femtosecond laser pulses (as in our experiments—see above) with relatively large spectral width (in the given paper, about 8 nm at a wavelength of 820 nm), the noncollinear frequency-nondegenerate character of the electric-dipole interaction must be taken into account. The various spectral components of the focused femtosecond pulse can interact among themselves, generating radiation at the sum frequency. The spectrum of the recorded signal will be localized at the second-harmonic frequency, and has finite width.

The finite spatial size of the molecular system, comparable with the wavelength of the light, suggests that spatial dispersion effects of the nonlinear susceptibilities can play an appreciable role in the nonlinear optical processes under consideration. This is especially promoted by the presence of spiral components in the structure of the bacteriorhodopsin molecule, which produce a substantial magnetic moment in addition to the electric-dipole moment. The presence of a nonzero magnetic moment induced by the electric field of the light wave, as well as an electric dipole moment oscillating at the second-harmonic frequency, induced by the combined action of the electric and magnetic fields of the light wave, causes a magnetic-dipole component to appear in the resulting electromagnetic field at the second-harmonic frequency, which is determined by the nonlinear susceptibility $\chi^{(2)M}$. The relatively large spatial size of the bacteriorhodopsin molecule[23] can also engender a substantial contribution determined by the quadrupole susceptibility $\chi^{(2)Q}$ of the medium of interest.[5]

When the spatial-dispersion effects of the nonlinear susceptibilities are allowed for, the induced nonlinear polarization of the medium, along with the purely electric-dipole contribution of Eq. (2), will contain ''nonlocal'' contributions corresponding to the magnetic-dipole and electric-quadrupole interactions:

$$P_i^{(2)M}(2\omega) = D_2 \chi_{ijk}^{(2)M}(2\omega;\omega,\omega) E_j^a(\omega) H_k^b(\omega), \quad (3)$$

$$P_i^{(2)Q}(2\omega) = D_2 \chi_{ijkl}^{(2)Q}(2\omega;\omega,\omega) \nabla_j E_k^a(\omega) E_l^b(\omega), \quad (4)$$

where $H_k^b(\omega)$ is the magnetic field component corresponding to the light wave, and $\nabla_j$ is the differentiation operator along the corresponding coordinate, with the indicated operator acting on each of the fields lying to the right of it.

Equations (3) and (4) are written for the degenerate case, in which the frequencies coincide, $\omega_1 = \omega_2 = \omega$, on the assumption of noncollinearity of the interaction. Such a representation is assumed to be justified in zeroth order of the theory of the frequency dispersion of the optical susceptibilities, since nonzero components of the second-order magnetic-dipole and electric-quadrupole susceptibility tensors exist in the degenerate version of the interaction for

isotropic media. We point out that nonlocal second-order processes can occur in centrosymmetric media[29] and can be associated in the same way with the ''chiral-insensitive'' components of the hyperpolarizability tensor of the chiral molecules being studied. Therefore, a nongyrotropic contribution will be present in the second-harmonic signal along with the ''chiral-sensitive'' contributions.

Femtosecond laser radiation having a high peak intensity provides a unique capability for studying processes associated with high-order nonlinear optical susceptibilities, in particular those with fourth-order nonlinear susceptibility. In the electric-dipole approximation, the induced nonlinear polarization $\mathbf{P}^{(4)D}$ of the medium at the second-harmonic frequency, due to the indicated processes, can be expressed in terms of the components of the fourth-order electric-dipole nonlinear susceptibility tensor $\chi^{(4)D}_{ijklm}$ in the form[7]

$$P_i^{(4)D}(2\omega)=D_4\chi^{(4)D}_{ijklm}(2\omega;\omega,\omega,\omega,-\omega)\times E_j^a(\omega)E_k^a(\omega)E_l^a(\omega)E_m^{b*}(\omega), \tag{5}$$

where $D_4$ is a 4-factor that allows for frequency degeneracy, and the asterisk denotes the complex conjugate.

Unlike the second-order electric-dipole susceptibility tensor in Eq. (2), the tensor $\chi^{(4)D}_{ijklm}$ in a macroscopically noncentrosymmetric medium has nonzero components even in the degenerate case, and therefore Eq. (5) is written in the zeroth approximation of frequency dispersion. The nonlocal fourth-order contributions can also be neglected, since the nonlinear susceptibilities $\chi^{(4)M}$ and $\chi^{(4)Q}$ corresponding to these processes are of the order of $d/\lambda$ with respect to the corresponding fourth-order electric-dipole susceptibility ($d$ is the characteristic size of the molecule, and $\lambda$ is the wavelength). Therefore, we shall restrict ourselves in what follows to the consideration of nonlocal second-order contributions.

In the nonlinear optics of high-order nonlinearities,[30] it is well known that besides the contribution of the ''direct'' nonlinear optical process of higher order (in our case, the fourth order, due to the nonlinear polarization of Eq. (5)), there exist ''cascade'' processes (of the same order in the optical field), which result from the simultaneous presence of several lower-order nonlinearities.[7] In particular, a process can occur as follows:

1. In the first stage, a second-harmonic field (frequency $2\omega$) is generated because of a quadratic nonlinearity of the form of Eqs. (2)–(4).

2. In the second stage of the cascade, the second-harmonic field is mixed with the field of the fundamental frequency $\omega$ at a third-order dipole nonlinearity according to the scheme $2\omega=2\omega+\omega-\omega$.

According to Refs. 31, 32, and 33, chap. 4, the second-harmonic field generated in a liquid in stage 1 can be written

$$\mathbf{E}(2\omega)=\frac{4\pi\mathbf{k}(\mathbf{k}\cdot\mathbf{P}^{(2)}(2\omega))}{|\mathbf{k}|^2\varepsilon(2\omega)}+\frac{4\pi}{(c\mathbf{k}/2\omega)^2-\varepsilon(\omega)}\times\left\{\mathbf{P}^{(2)}(2\omega)-\frac{\mathbf{k}}{|\mathbf{k}|^2}(\mathbf{k}\cdot\mathbf{P}^{(2)}(2\omega))\right\}, \tag{6}$$

where

$$\mathbf{P}^{(2)}(2\omega)=\mathbf{P}^{(2)D}(2\omega)+\mathbf{P}^{(2)Q}(2\omega)+\mathbf{P}^{(2)M}(2\omega), \tag{7}$$

the Cartesian components of vectors $\mathbf{P}^{(2)D}(2\omega)$, $\mathbf{P}^{(2)Q}(2\omega)$, and $\mathbf{P}^{(2)M}(2\omega)$ are defined in Eqs. (2), (3), and (4), respectively; $\mathbf{k}=2\mathbf{k}_\omega$ is the wave vector of a wave with nonlinear polarization at the doubled frequency; and $\varepsilon(2\omega)$ and $\varepsilon(\omega)$ are the permittivity of the liquid at the second-harmonic frequency and the fundamental frequency, respectively.

The second-harmonic field of Eq. (6) is mixed with the fundamental-frequency field $\mathbf{E}(\omega)$ in the second stage of the cascade, with the generation of fourth-order nonlinear cascade polarization at the second-harmonic frequency, of the following form:[33]

$$\begin{aligned}\mathbf{P}^{(4)\mathrm{cas}}(2\omega)=D_3\{&\chi^{(3)}_{1122}(2\omega;2\omega,\omega,-\omega)\times\mathbf{E}(2\omega)(\mathbf{E}(\omega)\mathbf{E}^*(\omega))\\&+\chi^{(3)}_{1212}(2\omega;2\omega,\omega,-\omega)\mathbf{E}(\omega)\times(\mathbf{E}(2\omega)\mathbf{E}^*(\omega))+\chi^{(3)}_{1221}(2\omega;2\omega,\omega,-\omega)\\&\times\mathbf{E}^*(\omega)(\mathbf{E}(2\omega)\mathbf{E}(\omega))\}.\end{aligned} \tag{8}$$

Here $D_3$ is a frequency degeneracy factor (in this case, $D_3=6$), and $\chi^{(3)}_{1122}(2\omega;2\omega,\omega,-\omega)$, etc. are the corresponding linearly independent components of the third-order (electric-dipole) nonlinear optical susceptibility tensor.

Another possible process contributing to the effective nonlinear fourth-order susceptibility is the following:

(1) In the first stage, self-action of the fundamental-frequency wave occurs (because of third-order dipolar nonlinear susceptibility), due to cubic nonlinear polarizability of the form

$$P_i^{(3)}(\omega)=D_3\chi^{(3)D}_{ijkl}(\omega;\omega,\omega,-\omega)E_j(\omega)E_k(\omega)E_i^*(\omega), \tag{7a}$$

where $D_3$ is a frequency degeneracy factor, and $\chi^{(3)D}_{ijkl}(\omega;\omega,\omega,-\omega)$ are the components of the third-order (electric-dipole) nonlinear optical susceptibility tensor.

(2) In the second stage of the cascade, the second-harmonic field is generated (because of second-order nonlinear susceptibility) with the participation of waves of the fundamental frequency partially modified by the self-action process, which is determined by nonlinear polarization of the form of Eq. (7a). In this stage, the nonlinear polarization induced in the medium at the second-harmonic frequency can be written in the form

$$P_i^{(4)\mathrm{cas}}(2\omega)=\chi^{(2)}_{ijk}(2\omega;\omega,\omega)E_j'(\omega)E_k(\omega). \tag{8a}$$

Here $\chi^{(2)}_{ijk}(2\omega;\omega,\omega)$ are the components of the second-order nonlinear optical susceptibility tensor, and the prime denotes the optical field of the fundamental frequency, modified by the self-action effect of Eq. (7a).

The self-action described by the nonlinear polarization of Eq. (7a) can manifest itself in phase self-modulation of the fundamental-frequency field, thereby broadening its spectrum, and in distortion of the phase front of the light wave, causing self-focusing/self-defocusing of the light beam[34] and, if the input wave is elliptically polarized, causing self-rotation of the polarization ellipse. Of all the enumerated

consequences of the self-action of the fundamental-frequency wave, the last effect might play the greatest role under our experimental conditions.

The cascade polarization of Eq. (8) can be comparable in magnitude to the direct fourth-order polarization of Eq. (5). Writing Eq. (8) in Cartesian coordinates and using Eq. (6), expressions can be obtained for the fourth-order cascade nonlinearities, which are too involved to include here.

The total fourth-order nonlinearity can be written as

$$\chi^{(4)}_{ijklm}(2\omega;\omega,\omega,\omega,-\omega)=\chi^{(4)D}_{ijklm}(2\omega;\omega,\omega,\omega,-\omega)+\chi^{(4)\mathrm{cas}}_{ijklm}(2\omega;\omega,\omega,\omega,-\omega). \tag{9}$$

Initially assuming that the nonlinear susceptibilities are complex quantities (because of the possible electronic resonance associated with absorption at one or several of the wavelengths involved in the interaction), it is necessary to take into account not only the amplitudes but also the phases of the individual components, as well as the possible interference of the different contributions. This circumstance is successfully used, for example, in coherent anti-Stokes Raman spectroscopy (CARS)[33] to analyze and separate the informative coherent contributions in the CARS signal.[35] It must be kept in mind in this case that, besides the interference of coherent processes of the same order (electric-dipole and magnetic-dipole), processes that have different orders of nonlinearity can also interfere.

Let us assume that the recorded signal at the second-harmonic frequency is generated by two coherent sources, the first of which is a sum of the contributions due to nonlinear susceptibilities $\chi^{(2)D}$, $\chi^{(2)M}$, and $\chi^{(2)Q}$, while the second is due to nonlinear susceptibility $\chi^{(4)}$ (direct and cascade processes). Fundamental-frequency radiation with intensity $I_\omega$ and complex field amplitude $E(\omega)$ induces in the medium two waves with nonlinear polarization and complex amplitudes $P^{(2)}(2\omega)=\chi^{(2)}E^2(\omega)$ and $P^{(4)}(2\omega)=\chi^{(4)}E^3(\omega)E^*(\omega)$ that interfere with each other. The amplitude of the total nonlinear polarization of the medium can then be written as

$$P_{NL}(2\omega)=P^{(2)}(2\omega)+P^{(4)}(2\omega), \tag{10}$$

where $P^{(2)}(\omega)$ is determined by Eq. (7), $P^{(4)}(2\omega)$ is determined by the sum of the nonlinear polarization waves given in Eqs. (5), (8), and (8a), and the second-harmonic signal intensity is $I_{2\omega}\propto|P_{NL}(2\omega)|^2$.

Strictly speaking, the dependence of the second-harmonic field intensity is determined not only by the nonlinear polarizability $P_{NL}(2\omega)$ and consequently by the nonlinear susceptibilities (direct and cascade) but also by the linear optical susceptibilities at the fundamental frequency, $\chi^{(1)}(\omega)$, and the doubled frequency, $\chi^{(1)}(2\omega)$, since the ''coherence length'' depends on these latter susceptibilities during second-harmonic generation. This quantity in our experiments was $l_{\mathrm{coh}}=\pi/\Delta k=\lambda_\omega/4(n_{2\omega}-n_\omega)\approx 15\ \mu$m when $\lambda_\omega=820$ nm and $l'_{\mathrm{coh}}\approx 5\ \mu$m when $\lambda_\omega=590$ nm ($n_\omega$ and $n_{2\omega}$ are the refractive indices of the bacteriorhodopsin solution at the corresponding frequencies). However, the corresponding factor in the expression for the second-harmonic intensity is determined by the linear optical parameters of the solution and the beam-focusing parameters at the fundamental frequency, which remained unchanged during our experiments; therefore, in describing the dependence of the second-harmonic intensity on the polarization state and intensity of the fundamental-frequency wave, we shall be interested only in the corresponding dependence of the nonlinear optical polarization.

In general, the complex nonlinear susceptibilities $\chi^{(2)}$ and $\chi^{(4)}$ can be represented as $\chi^{(n)}=|\chi^{(n)}|\exp(i\alpha_n)$, where $\alpha_n$ is a phase factor that reflects the complex character of the corresponding quantity. The intensity of the second-harmonic signal in this case can be written

$$I_{2\omega}\propto|P_{NL}(2\omega)|^2\propto|\chi^{(2)}|^2\ I_\omega^2+2|\chi^{(2)}||\chi^{(4)}|\cos\alpha I_\omega^3+|\chi^{(4)}|^2\ I_\omega^4, \tag{11}$$

where $\alpha=\alpha_4-\alpha_2$ reflects the phase shift between the complex values of $\chi^{(2)}$ and $\chi^{(4)}$.

The dependence of the total second-harmonic signal determined by Eq. (11) has a complex character and contains not only second- and fourth-order contributions but also an interference term that determines a contribution that depends cubically on the initial radiation intensity. Note that either constructive or destructive interference can occur, depending on the sign of cos $\alpha$. From an analysis of the experimentally measured dependence $I_{2\omega}=f(I_\omega)$, we established in Ref. 17 that when $\lambda_\omega=820$ nm, $\alpha=153\pm 2°$, so that cos $\alpha=-0.89$, and the interference was destructive.

A more rigorous treatment requires that the tensor character of the nonlinear susceptibilities defined by Eqs. (2)–(5) and (9) be taken into account. In this case, the induced nonlinear polarizations in Eq. (10) are vector quantities, and their direction and amplitude will depend on the polarization state of the initial radiation. Then Eq. (11) remains in force, if by $\chi^{(2)}$ and $\chi^{(4)}$ we denote the polarization-dependent effective nonlinear susceptibilities of the corresponding orders. The total second-harmonic intensity will also have a complex polarization dependence, with the character of the latter being determined not only by the relationship among the components of the nonlinear susceptibility tensors, but also by the intensity of the initial radiation.

## 5. DISCUSSION OF THE EXPERIMENTAL RESULTS

It must be pointed out, above all, that the form of the polarization dependence shown in Figs. 3 and 4 substantially differs from the corresponding dependence predicted by hyper-Rayleigh scattering theory.[36] In fact, the hyper-Rayleigh scattering signal intensity should be maximized when the exciting radiation has linear polarization and is oriented in a plane coincident with that of the maximum transmittance of the analyzer, whereas in our experiments at high intensities of nonresonant exciting radiation, as well as with resonant excitation, the maximum is reached when the polarization is circular. The authors of Refs. 19 and 22 associate the anomalously high level of hyper-Rayleigh scattering that they observed with the coherent character of the radiation of dipoles located in the same membrane, so that the fact that a maximum of the second-harmonic signal is

present for circular polarization of the input radiation is noted as a distinguishing feature of such an interaction.[21,22] The intensity of the second-harmonic radiation in this case is proportional to the square of the number of molecules in an individual membrane. The characteristic size of the membranes in our experiments was less by a factor of 10–20, so that the individual fragments contained about 200–500 bacteriorhodopsin molecules. The lower concentration of bacteriorhodopsin molecules corresponded to approximately the same number of fragments per unit volume of the suspension as in the papers cited above. The efficiency of hyper-Rayleigh scattering should have been reduced by about a factor of $10^4$ in this case, and should be below the sensitivity limits of the recording apparatus in our experiments.

Two more experimental facts must also be pointed out: For a sufficiently high intensity of the exciting radiation (curve *1* in Fig. 3), the second-harmonic signal intensity has different values for left- and right-circular polarization of the input radiation; this is probably associated with the chiral properties of the test medium (here an analogy can be drawn with linear circular dichroism). Moreover, as our experiments showed, for plane-polarized input radiation, the second-harmonic radiation was elliptically polarized (or partially depolarized), so that rotation of the principal axis of the polarization ellipse from the linear polarization direction of the input radiation was observed (an analogy can be drawn with linear optical rotation). This rotation, however, cannot be explained by linear optical rotation at the fundamental frequency, which, as we have mentioned, was small. These two facts are not explained by the extended theory of hyper-Rayleigh scattering of Refs. 21 and 22, in terms of which a suspension of purple membranes is represented by a centrosymmetric medium.

We now point out the other principal features of the experimental technique that we used. The process of second-harmonic generation was observed by means of femtosecond laser pulses having high peak power (all the way to $4\times10^6$ W) with relatively low mean power of the radiation (no more than 200 mW). The exciting radiation had a relatively high spectral width ($\delta\lambda/\lambda\approx10^{-2}$). The experiment was carried out using focused beams, so that the waves involved in the excitation process substantially differed from plane waves.

A theoretical analysis of second-harmonic generation in a gyrotropic medium consisting of arbitrarily oriented chiral molecules requires that one take into account local (electric-dipole) second- and fourth-order nonlinear contributions, as well as nonlocal (quadratic and magnetic-dipole) second-order contributions. The most general approach to formulating the basic principles of this analysis is outlined in the preceding section. However, a more complete and detailed theoretical analysis of the mutual influence of all the indicated contributions requires separate consideration. In this section, we wish to concentrate on a qualitative analysis of the basic characteristics of the recorded processes.

The generation of the second harmonic of monochromatic radiation due to nonlinear second-order susceptibility is forbidden even in a gyrotropic isotropic medium.[7] When the initial spectrum in noncollinear geometry has a large width, interaction of different spectral components of the exciting radiation (the generation of sum frequencies) becomes important. In the linear approximation of quadratic nonlinearity theory,[28] it can be expected that the effective value of the second-order electric-dipole nonlinear susceptibility will be proportional to $\delta\lambda/\Delta\lambda$, where $\delta\lambda$ is the spectral width of the exciting radiation and $\Delta\lambda$ is the offset of the wavelength $\lambda$ of the exciting radiation from the maximum of the absorption band. The magnetic-dipole and quadrupole susceptibilities are of order $d/\lambda$, where $d$ is the characteristic size of a single molecule. Taking into account our experimental parameters, it can be seen that the local and nonlocal second-order contributions can have comparable values.

An important difference between local and nonlocal nonlinearities is their differing sensitivity to the symmetry of the medium and the geometry of the experiment. Nonlinear polarization induced in a gyrotropic medium by magnetic-dipole and quadrupole interactions is identical for left- and right-circularly polarized input radiation, whereas polarization induced by electric-dipole interaction should change sign when the sign of the ''twist'' of the light wave's polarization changes. This circumstance can probably explain the intensity difference between the second-harmonic signal and the polarization dependences in Fig. 3 for $\varphi=45°$ and $\varphi=135°$: the various contributions interfere constructively in the first case, and destructively in the second.

The polarization dependences shown in Fig. 3 change as the peak power of the nonresonant interaction at the sample changes. The greatest changes in second-harmonic signal intensity occur in the case of plane-polarized input radiation ($\varphi=0°$ and 180°). This experimental fact can be described in terms of the interference interaction of the second- and fourth-order optical nonlinearities of the medium determined by Eq. (11): as the intensity of the exciting pulse increases, the contribution of the fourth-order nonlinearity to the second-harmonic signal increases by comparison with the contribution of the second-order nonlinearity, with the interference of the various contributions having a destructive character. The destructive character of the interference of the second- and fourth-order processes is confirmed in our other measurements.[16,17] The dependence of the second-harmonic intensity on the intensity of input radiation with femtosecond excitation substantially differed from quadratic, decreasing significantly at high enough excitation intensities. In this case, the dips in the polarization dependence at $\varphi=0°$ and $\varphi=180°$ were absent when nanosecond pulses were used in the same sample, whereas the dependence of the second-harmonic intensity on the excitation intensity was close to quadratic.[17]

The transition to resonant excitation conditions in our experiments was accompanied by an increase in the second-harmonic generation efficiency for circular polarization of the input radiation by comparison with the conversion efficiency of plane-polarized radiation. This fact can be explained by recalling that for fourth-order processes, resonances play a role not only at the fundamental frequency and at the second-harmonic frequency, but also at intermediate Raman frequencies (in our case, at frequency $3\omega$).[7]

This suggests that processes associated with fourth-order

nonlinear susceptibility experience more efficient amplification in this case. For resonant excitation using nanosecond laser pulses at a wavelength of $\lambda_\omega = 532$ nm, we were unable to detect a second-harmonic signal; this can be interpreted as a reduction of the role of second-order processes when one goes to resonant conditions. The lack of a difference between intensity maxima of the second-harmonic signal for waves with right- and left-circular polarization in Fig. 4 is also consistent with the assumption that the contribution of nonlinear second-order processes is small.

In fact, as one goes from right-circular polarization of the exciting radiation to left-circular polarization, the sign of the second-harmonic electric field generated by electric-dipole nonlinear susceptibilities (including fourth order) changes, by virtue of the chiral properties of the medium. Since the second-harmonic intensity can be detected in the experiment, such a sign change can be reflected in the signal only in the presence of a ''reference'' signal of the same frequency that is insensitive to right–left circular polarization of the initial radiation, as the second-harmonic radiation generated by second-order nonlocal contributions might be. It therefore stands to reason that the dependence shown in Fig. 4 reflects the polarization properties of the second-harmonic signal generated by the electric-dipole (local) fourth-order nonlinear optical susceptibility of the test sample under resonant excitation conditions.

Qualitative analysis of the experimental data shown here thus makes it possible to conclude that the nonlinear optical signal observed in our experiments at the second-harmonic frequency can be described by a coherent superposition of different second- and fourth-order nonlinearities, of magnetic-dipole, electric-dipole, and electric-quadrupole nature.

## 6. CONCLUSION

The dependence of the second-harmonic generation intensity in a finely dispersed suspension of purple membranes containing bacteriorhodopsin on the polarization state of the exciting radiation has been measured, using laser pulses of femtosecond width. The femtosecond pulses possessed sufficiently high peak power and low average energy to make it possible to observe nonlinear optical processes resulting from nonlinear susceptibility not only of second order, but also of higher (fourth) order, without damaging the object under study. A strengthening of the radiation-conversion efficiency and an increase in the role of higher-order nonlinear susceptibilities in the process of second-harmonic generation was observed as one goes to resonant excitation conditions.

This work was carried out with the partial financial support of the Russian Fund for Fundamental Research (Grants No. 96.02.16596 and 96.03.34067). The Laboratoire de Physico Chimie de l'Atmosphère is a member of the Center d'Etude et de Recherche des Lasers et Application. A.V.B., A.V.P., and A.P.Sh. thank the Université du Littoral, Dunkerque, France for providing positions for the period during which the experiments were carried out. One of the authors (A.V.B.) acknowledges Samsung Electronics for partial support of his participation in the joint experiment. The authors are grateful to A. Yu. Resnyanskiĭ (M. V. Lomonosov Moscow State University) and J. Burie (Université du Littoral, Dunkerque) for help in carrying out the experiments.

[1] Y. R. Shen, *Principles of Nonlinear Optics*, Wiley, New York (1984).
[2] S. A. Akhmanov, N. I. Koroteev, and I. L. Shumay, *Nonlinear Optical Diagnostics of Laser-Excited Semiconductor Surfaces*, V. S. Letokhov, C. V. Shank, Y. R. Shen, and H. Walther (eds.), Harwood Academic Publishers, London (1989).
[3] S. A. Akhmanov and R. V. Khokhlov, *Problems of Nonlinear Optics*, VINITI, Moscow (1964).
[4] G. A. Reider and T. F. Heinz, *Photonic Probes of Surfaces*, P. Halevi (ed.), Elsevier Science, Amsterdam (1995), p. 415.
[5] B. Koopmans, A. M. Janner, H. T. Jonkman *et al.*, Phys. Rev. Lett. **71**, 3569 (1993).
[6] B. Koopmans, A. Anema, H. T. Jonkman *et al.*, Phys. Rev. B **48**, 2759 (1993).
[7] N. I. Koroteev, Zh. Éksp. Teor. Fiz. **106**, 1260 (1994) [JETP **79**, 681 (1994)].
[8] L. Velluz, M. Le Grand, and M. Grosjean, *Optical Circular Dichroism: Principles, Measurements, and Applications*, Academic Press, New York (1965).
[9] A. B. Rubin, *Biophysics*, [in Russian], Vysshaya Shkola, Moscow (1987).
[10] J. D. Byers, H. I. Lee, T. Petralli-Mallow *et al.*, J. Phys. Chem. **97**, 1383 (1993).
[11] T. Verbiest, M. Kauranen, A. Persoons *et al.*, J. Am. Chem. Soc. **116**, 9203 (1994).
[12] J. J. Maki, M. Kauranen, and A. Persoons, Phys. Rev. B **51**, 51 (1993).
[13] L. Hecht and L. Barron, Chem. Rev. **225**, 525 (1994).
[14] P. M. Rentzepis, J. A. Giordmaine, and K. W. Wecht, Phys. Rev. Lett. **16**, 792 (1966).
[15] A. P. Shkurinov, A. V. Dubrovskii, and N. I. Koroteev, Phys. Rev. Lett. **70**, 1085 (1993).
[16] A. V. Balakin, D. Boucher, E. Fertein *et al.*, submitted to Opt. Commun.
[17] A. V. Balakin, D. Boucher, N. I. Koroteev *et al.*, JETP Lett. **64**, 718 (1996).
[18] P. K. Schmidt and G. W. Rayfield, Appl. Opt. **33**, 4289 (1994).
[19] E. Hendrickx, K. Clays, A. Persoons *et al.*, J. Am. Chem. Soc. **117**, 3547 (1995).
[20] Q. Song, C. Wan, and C. K. Johnson, J. Phys. Chem. **98**, 1999 (1994).
[21] D. L. Andrews, P. Allcock, and A. A. Demidov, Chem. Phys. Lett. **190**, 1 (1995).
[22] P. Allcock, D. L. Andrews, S. R. Meech *et al.*, Phys. Rev. A **53**, 2788 (1996).
[23] T. A. Ceska and R. Henderson, J. Mol. Biol. **213**, 539 (1990).
[24] C. L. Tang and H. Rabin, Phys. Rev. B **3**, 4025 (1971).
[25] D. L. Andrews, in *Modern Nonlinear Optics*, part 2, M. Evans and S. Kielich (eds.), Wiley, New York (1993), p. 545.
[26] G. Binning, C. F. Quate, and Ch. Gerber, Phys. Rev. Lett. **56**, 930 (1986).
[27] N. I. Koroteev, in *Frontiers in Nonlinear Optics. The Sergei Akhmanov Memorial Volume*, H. Walther, N. Koroteev, and M. Scully (eds.), Inst. of Physics Publishing, Bristol (1993), p. 228.
[28] J. A. Giordmaine, Phys. Rev. A **138**, 1599 (1965).
[29] G. Ashwell, G. Jeffries, D. G. Hamilton *et al.*, Nature (London) **375**, 385 (1995).
[30] S. A. Akhmanov, in *Nonlinear Spectroscopy: Proceedings, Enrico Fermi Course 64*, N. Bloembergen (ed.), North Holland Publishing Co., Amsterdam (1977).
[31] C. Flytzanis, in *Quantum Electronics, a Treatise*, H. Rabin and C. L. Tang (ed.), Academic Press, New York (1975), part A, p. 9.
[32] C. Flytzanis and N. Bloembergen, in *Progress in Quantum Electronics*, J. H. Sanders and S. Stenholm (eds.), Pergamon Press, Oxford (1976), vol. 4, part III, p. 271.
[33] S. A. Akhmanov and N. I. Koroteev, *Methods of Nonlinear Optics in the Spectroscopy of Scattered Light*, [in Russian], Nauka, Moscow (1981).
[34] A. Brodeur, F. A. Ilkov, and S. L. Chin, Opt. Commun. **129**, 193 (1996).
[35] V. F. Kamalov, N. I. Koroteev, A. P. Shkurinov *et al.*, J. Phys. Chem. **93**, 5645 (1989).
[36] P. D. Maker, Phys. Rev. A **1**, 923 (1970).

Translated by W. J. Manthey

# Tunneling ionization of Rydberg molecules

B. A. Zon*)

*Voronezh State University, 394693 Voronezh, Russia*
(Submitted 20 January 1997)
Zh. Éksp. Teor. Fiz. **112**, 115–127 (July 1997)

The theory of tunneling ionization of atoms is generalized to ionization of symmetric top molecules, either polar or nonpolar. Low-lying excited states of molecules, for which the ordinary Born–Oppenheimer approximation holds, and high-lying excited states, for which the inverse Born–Oppenheimer approximation holds, are discussed. Ionization in a constant external field is analyzed, as is ionization in an alternating field. It is shown that the orientation of the molecule's axis along the field does not lead to any appreciable increase in the ionization probability as compared to other orientations. 
[S1063-7761(97)01007-X]

## 1. INTRODUCTION

The qualitative explanation of tunneling constituted one of the triumphs of quantum mechanics. However, the quantitative theory of this phenomenon for atoms and molecules proved to be difficult. The reason lies primarily in the presence of a long-range Coulomb potential perturbing the electron's motion in the continuous spectrum. Rigorous results were obtained only for the hydrogen atom, since in parabolic coordinates the variables in the Schrödinger equation for an electron moving in the Coulomb field and a constant external electric field separate.[1,2]

With the development of laser physics the problem of describing the tunneling effect came to the foreground thanks to the famous paper of Keldysh,[3] who showed that the tunneling ionization regime emerges in an alternating electric field as well.

Modern quantitative theory of tunneling began with the paper of Smirnov and Chibisov,[4] who used parabolic coordinates to describe electron motion in the tunneling region of an arbitrary atom. The matching of this region with the vicinity of the atomic core takes place on a surface on which the sum of the energy of the electron–core interaction and the interaction of the electron with the external field is considerably lower than the total electron energy. Clearly, these ideas are similar to the well-known $R$-matrix method in nuclear reaction theory,[5] which recently has found wide application in the theory of multiphoton processes in atoms and molecules.[6]

The next important step was taken by Perelomov *et al.*,[7] who found that Keldysh's tunneling ionization regime follows, to within a multiplicative factor, from Smirnov and Chibisov's formulas if (a) the field is formally assumed to be alternating and (b) one averages over the field's period.

In Refs. 4 and 7 the tunneling probability was found to within a constant that determines the asymptotic behavior of the radial wave function of an atomic electron. This constant was calculated by Ammosov, Delone, and Kraĭnov,[8] who used the semiclassical approximation and the model potential method. The resulting theory of tunneling in atoms provided a quantitative explanation of a number of experimental results[9,10] and became known as ADK theory. A description of the modern state of ADK theory and a comparison of its results with experimental data can be found in Ref. 11. A more recent work in this field is the paper of Faisal,[12] where an attempt was made to describe tunneling detachment of two electrons from an atom.

The present paper develops the theory of tunneling for molecules. In contrast to atoms, an electron in a molecule has no definite value either of orbital angular momentum or the projection of that momentum on the direction of an external field, quantities that play an important role in the Smirnov–Chibisov theory. Furthermore, calculations of the electronic wave functions of molecules done by quantum chemistry methods lead to complicated and implicit dependencies of these functions of the angular variables. At present, relatively simple expressions for the angular parts of the electronic wave functions for molecules, which are largely not centrally symmetric, are known only for two models: an electron in the field of short-range potentials,[13,14] and an electron in the field of a point dipole[15–17] (a review of these topics can be found in Ref. 18). In the present paper the description is done using the second model, since it corresponds better to real molecules.

Note that the existing tunneling theory is essentially single-particle, so that, strictly speaking, it can be applied only to Rydberg states. However, vibrational and rotational excitations play an important role in molecules even in the one-electron approximation. This fact and the lack of spherical symmetry complicate the theory of tunneling in molecules considerably.

In Secs. 2 and 3 we discuss the asymptotic behavior of the electronic wave function of a polar molecule inside the barrier. In Sec. 4 we calculate the probability of tunneling ionization in a constant uniform external field. Such a field is an interesting object in its own right, in connection particularly with problems of ZEKE spectroscopy.[19,20] However, the formulas cannot be applied to high-lying excited Rydberg states of molecules (and these are studied primarily by ZEKE-spectroscopy methods) because the adiabatic, or Born–Oppenheimer, approximation is used in deriving them: the separation of the Rydberg states in this region must be considerably greater than the frequencies of rotational tran-

sitions of the molecular core.[1)] This leads to a situation in which the inverse Born–Oppenheimer approximation is valid.[18] The corresponding tunneling effect is discussed in Sec. 5.

Section 6 is devoted to a transition from a constant field to an alternating one by the method developed in Ref. 7 and mentioned earlier. In Sec. 7 we discuss the effect of orientation of the molecule. Finally, in Sec. 8 we examine the effect of the dipole moment on the tunneling probability. We also calculate the probability of ionization of nonpolar molecules.

The influence of collective effects associated with rotational and vibrational excitations of the molecule on the tunneling probability is not studied in this paper. These problems merit a separate investigation.

Throughout the paper we use natural units.

## 2. WAVE FUNCTION IN THE *in*-REGION

The effect of an external electric field on the motion of an electron moving inside the barrier created by the field can be ignored.[2] In this region the electron moves in the field of the short-range molecular core, the Coulomb field of the residual ion with charge $z$, and the field of the point dipole **d** at rest. A fixed orientation of the dipole corresponds to the Born–Oppenheimer approximation.

In the one-particle approximation one must assume that the core possesses no electron angular momentum that affects the motion of the electron. Then the electronic wave function can be written as[15,16]

$$\Psi_{in}(\mathbf{r})=R_{\nu lm}(r)Z_{lm}(\theta,\phi). \tag{1}$$

Here $m$ is the conserved projection of the electron angular momentum on the direction of the dipole moment. For symmetric top molecules, in which the dipole moment is directed along the symmetry axis (as we assume below), $m=0$ corresponds to the $\Sigma$ state, $m=1$ to the $\Pi$ state, etc.

The angular functions $Z_{lm}(\theta,\phi)$ are eigenfunctions of the equation

$$[L^2+2d\ \cos\ \theta]Z_{lm}(\theta,\phi)=\lambda_{lm}(\lambda_{lm}+1)Z_{lm}(\theta,\phi),$$

where **L** is the orbital angular momentum operator.

The number $l$ is not the electron's orbital angular momentum since it is not conserved in the absence of spherical symmetry, but it labels the eigenvalues $\lambda_{lm}$ in such a way that $\lambda_{lm}\to l\geqslant|m|$ as $d\to 0$.

The functions $Z_{lm}$ can be expanded in the usual spherical harmonics,

$$Z_{lm}(\theta,\phi)=\sum_{ll'} a^m_{ll'}Y_{l'm}(\theta,\phi), \tag{2}$$

where the coefficients $a^m_{ll'}$ obey linear algebraic equations with a tridiagonal matrix that follow from the above eigenvalue equation. The solution of these equations based on perturbation theory in $d$ provides a good description of the exact solutions up to the critical value of the dipole moment, at which a ''fall-to-center'' effect[2] occurs and the point-dipole approximation becomes invalid. Some values of $\lambda_{lm}$ obtained by this method are

$$\lambda_{00}\simeq-\frac{2}{3}d^2,\quad \lambda_{10}\simeq 1+\frac{2}{15}d^2,\quad \lambda_{20}\simeq 2+\frac{2}{105}d^2,$$

$$\lambda_{11}\simeq 1-\frac{1}{15}d^2,\quad \lambda_{21}\simeq 2+\frac{1}{105}d^2,\quad \lambda_{31}\simeq 3+\frac{1}{210}d^2,$$

$$\lambda_{22}\simeq 2-\frac{2}{105}d^2,\quad \lambda_{32}\simeq 3,\quad \lambda_{42}\simeq 4+\frac{4}{3465}d^2.$$

The $\nu$ in (1) can be interpreted as the electron's effective principal quantum number. If we allow for the short-range core in the quantum defect method,[21,22] we can write the radial wave function in (1) in the form[23]

$$R_{\nu lm}(r)=(-1)^n C^{-1/2}_{\nu lm}\frac{z^{1/2}}{\nu r}W_{\nu,\lambda+1/2}\left(\frac{2zr}{\nu}\right),$$

$$C_{\nu lm}=\left(1+\frac{d\mu_{\nu lm}}{d\nu}\right)\Gamma(\nu-\lambda)\Gamma(\nu+\lambda+1), \tag{3}$$

$$\nu\equiv\nu_{nlm}=n+\lambda+1-\mu_{\nu lm},\quad \lambda\equiv\lambda_{lm},$$

$n=0,1,2,\ldots$ is the radial quantum number, $\mu_{nlm}$ is the quantum defect due to the core, and $W$ is the Whittaker function. The energy of the state (1) is given by the usual Rydberg formula

$$E=-\frac{z^2}{2\nu^2}. \tag{4}$$

The radial function (3) differs from the corresponding atomic function in that the second index in the Whittaker function is not half an odd integer. Note that the quantum defect method immediately gives the value of the normalization constant $C_{\nu lm}$.

Calculating tunneling probability requires knowing the asymptotic behavior of the radial function (3) as $r\to\infty$. If we use the well-known asymptotic behavior of the Whittaker function,[24] we obtain

$$R_{\nu lm}(r)\simeq(-1)^n\frac{2z^{3/2}}{\nu^2}C^{-1/2}_{\nu lm}\exp\left(-\frac{zr}{\nu}\right)\left(\frac{2zr}{\nu}\right)^{\nu-1}. \tag{5}$$

The wave function is written here in the molecular system of coordinates. Let us go over to the system of coordinate linked to the external field.

## 3. WAVE FUNCTION IN THE LABORATORY COORDINATE SYSTEM

We direct the $x_3$ axis of the laboratory system of coordinates along the electric field **F**. The $\tilde{x}_3$ axis of the molecular system of coordinates is directed along the dipole **d**. Let $\beta$ be the angle between **d** and **F**; the $x_1$ and $\tilde{x}_1$ axes lie in the same plane as the $x_3$ and $\tilde{x}_3$ axes. Then the $x_2$ and $\tilde{x}_2$ axes prove to be coincident, while the transformations of the other coordinates are

$$x_3=\tilde{x}_3\ \cos\ \beta-\tilde{x}_1\ \sin\ \beta,\quad x_1=\tilde{x}_3\ \sin\ \beta+\tilde{x}_1\ \cos\ \beta.$$

In addition to Cartesian coordinates, we introduce parabolic coordinates in the laboratory reference frame:

$$x_1=\sqrt{\xi\eta}\ \cos\ \varphi,\quad x_2=\sqrt{\xi\eta}\ \sin\ \varphi,\quad x_3=\frac{1}{2}(\xi-\eta),$$

$$\xi = r + x_3, \quad \eta = r - x_3, \quad \varphi = \arctan\frac{x_2}{x_1}.$$

Let us examine the region near the $x_3$ axis, which determines the penetrability of the barrier. As noted earlier, in the theory of tunneling the external field is assumed to be weak compared to the atomic field, so that there exists a range of distance,

$$\frac{\nu}{z} \ll \xi \ll \xi_0 \equiv \frac{z^2}{\nu^2 F},$$

where the field of the molecular core can be neglected but the external field is still not strong enough. At $\xi \simeq \xi_0$ the wave functions in the *in*- and *out*-regions are matched.

At such distances the substantial values of the coordinate $\eta$ are determined by the decay of the wave function (5):

$$\eta \leqslant \eta_0 \equiv \frac{\nu}{z}.$$

Thus, in the matching region there is a smallness parameter

$$\zeta \equiv \left(\frac{\eta_0}{\xi_0}\right)^{1/2} \simeq \left(\frac{\nu^3 F}{z^3}\right)^{1/2} \equiv \left(\frac{F}{F_0}\right)^{1/2} \ll 1, \tag{6}$$

where $F_0 = z^3/\nu^3$ is the atomic field strength in a state with effective principal quantum number $\nu$.

In parabolic coordinates, the dependence of the functions (2) on the angle $\phi = \arctan(\tilde{x}_2/\tilde{x}_1)$ is given by

$$e^{im\phi} \simeq \left[\frac{\sin\beta + 2\zeta(\cos\beta\,\cos\varphi + i\,\sin\varphi)}{\sin\beta + 2\zeta(\cos\beta\,\cos\varphi - i\,\sin\varphi)}\right]^{m/2} \equiv \Phi_m(\beta;\varphi).$$

The function $\Phi_m$ is simpler in (a) the region $G$ of large values of $\beta$, where $\beta, \pi - \beta \gg \zeta$:

$$\Phi_m(\beta;\varphi) \simeq 1; \tag{7}$$

(b) the region $L_+$ of small values of $\beta$, where $\beta \ll \zeta$:

$$\Phi_m(\beta;\varphi) \simeq e^{im\varphi}; \tag{8}$$

and (c) the region $L_-$ where $\pi - \beta \ll \zeta$:

$$\Phi_m(\beta;\varphi) \simeq e^{-im\varphi}. \tag{9}$$

We transform the Legendre polynomials in a similar manner:

$$P_l^{|m|}(\cos\theta) = P_l^{|m|}\left(\frac{\tilde{x}_3}{r}\right) \simeq P_l^{|m|}[(1 - 2\zeta^2)\cos\beta + 2\zeta\,\sin\beta\,\cos\varphi].$$

If we exclude the neighborhood of the zeroes of the Legendre polynomials in $G$,

$$P_l^{|m|}(\cos\theta) \simeq P_l^{|m|}(\cos\beta). \tag{10}$$

In $L_+$ (see Ref. 24),

$$P_l^{|m|}(\cos\theta) \simeq (-1)^m \frac{(l+|m|)!}{|m|!(l-|m|)!}\left(\frac{\eta}{\xi}\right)^{|m|/2}. \tag{11}$$

Finally in $L_-$ the Legendre polynomial (1) acquires an additional phase factor $(-1)^l$.

If we employ the expressions (7)–(11), the Eqs. (1), (2), and (5) can be written in the following form (we have dropped all inessential phase factors):

$$\Psi_{in}(\beta \in G; \xi, \eta, \varphi) = \frac{B^G_{\nu lm}}{\sqrt{2\pi}} e^{-z\xi/2\nu}\left(\frac{\xi}{2}\right)^{\nu-1} e^{-z\eta/2\nu}, \tag{12}$$

$$\Psi_{in}(\beta \in L_\pm; \xi, \eta, \varphi) = B^L_{\nu lm} e^{-z\xi/2\nu}\left(\frac{\xi}{2}\right)^{\nu-1-|m|/2} e^{-z\eta/2\nu}(2\eta)^{|m|/2}\frac{e^{\pm im\varphi}}{\sqrt{2\pi}}, \tag{13}$$

$$B^G_{\nu lm} = \frac{2^\nu z^{\nu+1/2}}{\nu^{\nu+1}} C^{-1/2}_{\nu lm} \sum_{l'} a^m_{ll'} \times \left[\frac{2l'+1}{2}\frac{(l'-|m|)!}{(l'+|m|)!}\right]^{1/2} P^{|m|}_{l'}(\cos\beta), \tag{14}$$

$$B^L_{\nu lm} = \frac{2^\nu z^{\nu+1/2}}{\nu^{\nu+1}|m|!}\, C^{-1/2}_{\nu lm} \sum_{l'} a^m_{ll'}\left[\frac{2l'+1}{2}\frac{(l'+|m|)!}{(l'-|m|)!}\right]^{1/2}. \tag{15}$$

If we neglect the difference between $B^G$ and $B^L$, we can interpret (12) as a special case of (13) with $m = 0$. Comparing (13) with Eq. (7) of Ref. 4, we notice that the two yield the same dependence on $\xi$, $\eta$, and $\varphi$ and differ only by a phase factor.

## 4. PROBABILITY OF TUNNELING

Finding the probability of tunneling requires calculating the barrier penetration factor, which can be done in parabolic coordinates. Since the strength of the external field is assumed to be low compared to that of the atomic field, the classical turning points are positioned at large $\xi$, and at such distances the centrifugal term in the electron energy can be ignored.[2,4,7] The electron–dipole interaction decays with distance in the same way as the centrifugal interaction. Hence the effect of the dipole moment on the electron motion under and outside the barrier can be ignored.

Thus, to obtain the probability of tunneling ionization of molecules we can employ the results of calculating this effect in atoms. As noted above, the difference here is only in the value of the normalization constant.

The final expression for the ionization rate (ionization probability per unit time) has the form

$$W_{\nu lm}(\beta \in G) = \left(\sum_{l'} a^m_{ll'}\left[\frac{(2l'+1)(l'-|m|)!}{(l'+|m|)!}\right]^{1/2} P^{|m|}_{l'}(\cos\beta)\right)^2 \times \frac{S_{2\nu}(F)}{C_{\nu lm}}, \tag{16}$$

$$W_{\nu lm}(\beta \in L_\pm) = \left(\sum_{l'} a^m_{ll'}\left[\frac{(2l'+1)(l'+|m|)!}{(l'-|m|)!}\right]^{1/2}\right)^2 \frac{S_{2\nu-|m|}(F)}{|m|!\,C_{\nu lm}}, \tag{17}$$

where the constant $C_{\nu lm}$ has been defined in (3), and

$$S_p(F)\equiv\frac{z^2}{\nu^3}\left(\frac{\nu^2F}{4z^3}\right)^{1-p}\exp\left(-\frac{2z^3}{3\nu^3F}\right). \tag{18}$$

Clearly, (17) follows from (16) at angles

$$\beta,\pi-\beta=\beta_{\min}\equiv\left(\frac{F}{4F_0}\right)^{1/2}\equiv\left(\frac{\nu^3F}{4z^3}\right)^{1/2}\ll 1. \tag{19}$$

Recall that the small parameter $\beta_{\min}$ reflects the presence of a boundary in the cone of orientations of the molecule's axis with respect to the external field, and Eqs. (17) and (16) are valid inside and outside the cone, respectively. Bearing in mind the substitution (19) for the regions $L_{\pm}$, we examine only Eq. (16).

## 5. HIGH-LYING EXCITED RYDBERG STATES

We now repeat all the calculations for high-lying excited Rydberg states, for which the inverse Born–Oppenheimer approximation mentioned in the Introduction is valid.

In the Appendix we show that in this approximation the wave function of a molecule is of the type (1), with all the arguments of the function referred to the laboratory coordinate system. The asymptotic behavior of the radial function is determined, apart from notation, by Eqs. (3) and (5). The angular function now depends explicitly on the Euler angles $\Omega$ that specify the orientation of the molecular core in the laboratory coordinate system, and is given by Eq. (A4).

Clearly, the main contribution in the vicinity of the $x_3$ axis is provided by the component of the function (A4) corresponding to $\sigma=0$. Hence, introducing into (A4) the substitution

$$Y_{l'\sigma}\left(\frac{\mathbf{r}}{r}\right)\Big|_{\vartheta\to 0}\to\sqrt{\frac{2l'+1}{2}}\,\delta_{\sigma 0},$$

we obtain

$$Z^{jkl}_{JM}\left(\frac{\mathbf{r}}{r},\Omega\right)\simeq D^j_{kM}(\Omega)\sum_{l'}\sqrt{\frac{(2j+1)(2l'+1)}{16\pi^2}}\times a^{Jjk}_{ll'}C^{JM}_{jMl'0}.$$

When we calculate the electron flux along the $x_3$ axis with such molecular wave functions, bilinear combinations of the core wave functions ($D$-functions) emerge. Integrating over Euler angles, we obtain for the ionization rate

$$W_{\nu slM}=\left(\sum_{l'}\sqrt{2l'+1}a^s_{ll'}C^{JM}_{jMl'0}\right)^2\frac{S_{2\nu}(F)}{C_{sl}},$$

$$C_{sl}=\left(1+\frac{d\mu_{\nu sl}}{d\nu}\right)\Gamma(\nu-\lambda_{sl})\Gamma(\nu+\lambda_{sl}+1), \tag{20}$$

$$\nu\equiv\nu_{sl}=n+\lambda_{sl}+1-\mu_{sl},\quad s\equiv\{Jjk\},\quad n=0,1,2,\ldots.$$

Equation (20) provides an explicit expression for the dependence of the ionization rate on the magnetic quantum number $M$. The dependence of the rate on the core quantum numbers $j$ and $k$ and the molecule's total angular momentum $J$ is contained in the $\lambda_{sl}$ and $a^s_{ll'}$.

## 6. TUNNELING IN AN ALTERNATING FIELD

In an alternating field of frequency $\omega$, tunneling emerges at small values of the Keldysh parameter[3]:

$$\gamma\equiv\frac{z\omega}{\nu F}\ll 1, \tag{21}$$

where $F$ is the electric field amplitude.

Below we assume that the inequality (21) is valid for low-lying excited Rydberg states, which satisfy the Born–Oppenheimer approximation. For the condition (21) to be true for the high-lying Rydberg states considered in Sec. 5, it is necessary, at least, that the frequency $\omega$ of the external field be much lower than the molecule's rotation frequency.

To obtain the probability of ionization in an alternating and linearly polarized field, we can, following the results of Ref. 7, make the replacement $F\to F\cos\omega t$ in (16) and (17) and average over the field's period. Calculating the integrals in (16) and (17) by the saddle-point method leads to an additional factor,

$$\left(\frac{3F}{\pi F_0}\right)^{1/2}\ll 1. \tag{22}$$

With allowance for this factor, Eq. (17) becomes the well-known ADK result,[2)] provided that one neglects $d\mu/d\nu$ and allows for the fact that neglecting the core's dipole moment leads to

$$a_{ll'}\to\delta_{ll'},\quad \lambda_{lm}\to l\geqslant|m|. \tag{23}$$

The difference is that in the ADK theory the detached electron is described by the method of the Simons model potential,[25] which corresponds to the formal introduction of a nonintegral orbital angular momentum $l^*$, while in the present paper to describe electron motion we use the quantum defect method, being more consistent theoretically. However, numerically both methods leads to essentially identical results.[22]

In the same manner we can examine arbitrary (elliptical) polarization of the alternating field. Not wishing to go into further details, thoroughly discussed in Refs. 7 and 8, we only note that in a circularly polarized field the ionization probability is given by the same Eqs. (16) and (17), since in such a field the instantaneous electric field does not change. Hence, for elliptical polarization, Eqs. (16) and (17) acquire an additional factor that varies between unity and the value specified by (22).

We now turn to some limits and special cases of the above formulas.

## 7. ROLE OF ORIENTATION OF THE MOLECULE

When the orientation of the molecule with respect to the polarization of the external field is random, Eq. (16) must be averaged over the angle $\beta$. Using the orthonormality of the coefficients $a^m_{ll'}$ and Stirling's formula for the gamma function $\Gamma$ in the $C_{\nu lm}$, we obtain, with allowance for the factor (22),

$$\overline{W}_{\nu lm}=\frac{\sqrt{3}e^{2\nu}}{\pi\sqrt{\pi}(\nu^2-\lambda^2)^{\nu}}\left(1+\frac{d\mu_{\nu lm}}{d\nu}\right)^{-1}$$

$$\times\left(\frac{\nu-\lambda}{\nu+\lambda}\right)^{\lambda+1/2} S_{2\nu-1/2}(F),$$

$$\nu\equiv\nu_{nlm}, \quad \lambda\equiv\lambda_{lm}. \tag{24}$$

Similarly, Eq. (20) can be averaged over the magnetic quantum number $M$:

$$\overline{W}_{\nu sl}=\frac{e^{2\nu}S_{2\nu}(F)}{2\pi(\nu^2-\lambda^2)^{\nu}}\left(1+\frac{d\mu_{\nu sl}}{d\nu}\right)^{-1}\left(\frac{\nu-\lambda}{\nu+\lambda}\right)^{\lambda+1/2},$$

$$\nu\equiv\nu_{nsl}, \quad \lambda\equiv\lambda_{sl}. \tag{25}$$

With a high-intensity light field, in which tunneling is observed, the distribution of orientations of molecules is highly anisotropic,[26] with the result that Eq. (24) can be used only for estimates. However, this equation can be employed to simplify (16) and (17):

$$W_{\nu lm}(\beta)=\overline{W}_{\nu lm}\left(\sum_{l'} a^{m}_{ll'}\left[\frac{(2l'+1)(l'-|m|)!}{(l'+|m|)!}\right]^{1/2}\times P^{|m|}_{l'}(\cos\beta)\right)^2. \tag{26}$$

Concluding this section, we note that, for example, Seidman *et al.*[27] allowed for a strong dependence of the ionization rate on the molecule's orientation with respect to the external field. According to the above equations, no such dependence emerges for $\Sigma$ states.

## 8. ROLE OF THE DIPOLE MOMENT AND THE IONIZATION OF NONPOLAR MOLECULES

Since the probability of tunneling ionization of atoms at moderate values of $\nu$ depends strongly on the electron's orbital angular momentum,[8] one should expect a considerable change in the ionization probability for molecules because of the presence of a dipole moment. The point here is that the dipole moment can force the Rydberg electron to go into a state with orbital angular momentum such that the ionization probability is a maximum.

However, as the above equations imply, the ionization probability is determined by the conserved quantum number $\lambda$, and the fact that it is not an integer is due primarily to the effect of the dipole moment on the ionization probability. The difference between the coefficients $a_{ll'}$ and their limiting values (23) can be neglected to a high accuracy. Then, for instance, Eq. (26) takes the much simpler form

$$W_{\nu lm}(\beta)=\frac{(2l+1)(l-|m|)!}{(l+|m|)!}\,\overline{W}_{\nu lm}[P^{|m|}_{l}(\cos\beta)]^2. \tag{27}$$

For nonpolar molecules (23) is strictly observed. In such molecules the difference between the initial hydrogen-like states of the Rydberg electron is determined solely by the short-range molecular core, which is manifested by the difference in the corresponding quantum defects. Here the probability of Rydberg states becoming ionized is given by Eq. (27), with $\lambda_{lm}=l$ in the expression (24) for $\overline{W}_{\nu lm}$.

Similarly, Eq. (20) for nonpolar molecules becomes

$$W_{\nu slM}=(2l+1)(C^{JM}_{jMl0})^2\,\frac{S_{2\nu}(F)}{C_{sl}}, \tag{28}$$

where $\lambda_{sl}$ must also be replaced by $l$ in the expression for $C_{sl}$.

Equation (28) also describes tunneling for extremely high-lying excited Rydberg states of polar molecules, which for typical molecules begin at $\nu\sim 60$. In such states the interaction of the Rydberg electron and the dipole moment proves to be weaker than the splitting of the core's molecular levels, a splitting responsible for the formation of a dipole moment, for example, $\Lambda$-doubling in diatomic molecules.[28]

In conclusion, we note that the dependence of the ionization probability of atoms and molecules on the external field strength is given by the same function (18). This agrees with the experimental data obtained by Chin *et al.*[29] in observing the ionization of $H_2$ and $N_2$ molecules by the light of a carbon dioxide laser. The fact that these molecules do not dissociate, although the dissociation potential is several times lower than the ionization potential, finds its explanation in the model adopted in Ref. 30.

I am deeply grateful to N. B. Delone, V. P. Kraĭnov, and V. E. Chernov for discussions and useful remarks.

## APPENDIX: THE GREEN'S FUNCTION OF A MOLECULE IN THE INVERSE BORN–OPPENHEIMER APPROXIMATION

In Ref. 23 the quantum defect method was used to build the one-electron Green's function of a polar molecule in the Born–Oppenheimer approximation. Here we derive a similar quantity when the inverse Born–Oppenheimer approximation is valid. The significant difference is that now the Green's function is not one-particle, since the rotational degrees of freedom of the molecular core must be taken into account.

We write the molecule's Hamiltonian in the form

$$H=H^{+}+T+V_c-\frac{z}{r}-\frac{\mathbf{d}\cdot\mathbf{r}}{r^3}. \tag{A1}$$

Here $\mathbf{r}$ is the radius vector of the Rydberg electron, $T$ is the electron's kinetic energy, $\mathbf{d}$ is the dipole moment of the core, and $V_c$ is the short-range part of the core potential, which determines the values of the quantum defects. The core is assumed to be axisymmetric, and its Hamiltonian $H^{+}$ is

$$H^{+}=Bj^2+(C-B)j^2_{\zeta}, \tag{A2}$$

where $B$ and $C$ are the rotational constants.

Following Ref. 17, we introduce the molecular angular functions $Z^{jk}_{JM}(\mathbf{r}/r,\Omega)$, which are eigenfunctions of the equation

$$\left(L^2+2\,\frac{\mathbf{d}\cdot\mathbf{r}}{r}\right)Z^{jk}_{JM}\left(\frac{\mathbf{r}}{r},\Omega\right)=\lambda(\lambda+1)Z^{jk}_{JM}\left(\frac{\mathbf{r}}{r},\Omega\right). \tag{A3}$$

Here $\Omega=\{\alpha,\beta,\gamma\}$ denotes the set of Euler angles, which determine the orientation of the molecular core in the laboratory system of coordinates, and $\mathbf{L}$ is the orbital angular momentum operator of the electron.

The functions in Eq. (A3) correspond to molecular states with total angular momentum $J$ and projection $M$ on an axis

fixed in space, and with a core angular momentum $j$ and projection $k$ on the molecule's axis of symmetry. They can be represented in the form

$$Z_{JM}^{jk}\left(\frac{\mathbf{r}}{r},\Omega\right)=\sqrt{\frac{2j+1}{8\pi^2}}\sum_{l\sigma m}a_l^s C_{jml\sigma}^{JM}D_{km}^{j}(\Omega)Y_{l\sigma}\left(\frac{\mathbf{r}}{r}\right). \tag{A4}$$

The coefficients $a_l^s$ in (A4) satisfy a finite system of linear algebraic equations,[17] in contrast to the infinite system of equations that emerge in the Born–Oppenheimer approximation; they are normalized by the condition

$$\sum_l |a_l^s|^2=1. \tag{A5}$$

The eigenvalues $\lambda$ in Eq. (A3) and the coefficients $a_l^s$ in (A4) depend on the set of quantum numbers $s\equiv\{Jj\Omega\}$.

The functions $Z_{JM}^{jk}$ form a complete set, so that the Green's function determined by the equation

$$(H-E)G_E(\mathbf{r},\Omega;\mathbf{r}',\Omega')=\delta(\mathbf{r}-\mathbf{r}')\delta(\Omega-\Omega'), \tag{A6}$$

can be expanded in these functions:

$$G_E(\mathbf{r},\Omega;\mathbf{r}',\Omega')=\sum_{sM}g_s(E;r,r')Z_{JM}^{jk}\left(\frac{\mathbf{r}}{r},\Omega\right)\times Z_{JM}^{jk*}\left(\frac{\mathbf{r}'}{r'},\Omega'\right). \tag{A7}$$

If we also expand the delta functions,

$$\delta(\mathbf{r}-\mathbf{r}')=\frac{\delta(r-r')}{rr'}\sum_{l\sigma}Y_{l\sigma}\left(\frac{\mathbf{r}}{r}\right)Y_{l\sigma}^{*}\left(\frac{\mathbf{r}'}{r'}\right),$$

$$\delta(\Omega-\Omega')=\sum_{jkm}\frac{2j+1}{8\pi^2}D_{km}^{j}(\Omega)D_{km}^{j*}(\Omega'), \tag{A8}$$

we can use the above equations to derive an equation for the radial Green's function:

$$\left[-\frac{1}{2r^2}\frac{d}{dr}\left(r^2\frac{d}{dr}\right)+\frac{\lambda_s(\lambda_s+1)}{2r^2}+V_c-\varepsilon\right]g_s(E;r,r')=\frac{\delta(r-r')}{rr'},\quad \varepsilon=E-Bj(j+1)-(C-B)k^2. \tag{A9}$$

This equation is no different from those studied in Refs. 31 and 23. Outside the core, where the potential $V_c$ affects only the scattering phases (in the continuous spectrum) or the quantum defects (in the discrete spectrum), but not the radial dependence of the wave functions, the solution can be expressed in terms of the Whittaker function and, with the change in notation

$$\lambda_{lm}\rightarrow\lambda_{sl}, \tag{A10}$$

is given by Eqs. (7) and (9) of Ref. 23.

The second index in $\lambda_{sl}$ in (A10) labels the eigenvalues of Eq. (A3) in such a way that $\lambda_{sl}\rightarrow l$ as $d\rightarrow 0$. The same index must be attached to the coefficients $a_l^s$,[3)] after which all the equations of the ordinary and inverse Born–Oppenheimer approximations coincide. The only difference here is the way in which the angular parts (2) and (A4) are defined.

In particular, the electronic radial wave function is found by calculating the residue of the Green's function, and is given by Eqs. (3) and (5), in which the change of notation (A10) and the change in indices $\{lm\}\rightarrow\{ls\}$ must be introduced in the appropriate places.

*)E-mail: zon@niif.vucnit.voronezh.su

1)A similar relationship for vibrational transitions introduces essentially no changes into the theory, except for the rare cases of inversion degeneracy of rotational spectra and the like.[18]

2)We note in passing that Eqs. (20) and (21) of Ref. 8 contain an error: the factor $(n^*+l^*)/(n^*-l^*)$ must be replaced by $(n^*-l^*)/(n^*+l^*)$.

3)Since the system of equations (A3) is finite, it can be solved exactly in the simplest cases. Explicit expressions for $\lambda_{sl}$ and $a_{ll'}^s$ in some of these cases can be found in Ref. 17.


[1] C. Lanczos, Z. Phys. **68**, 204 (1931).
[2] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-relativistic Theory*, 3rd ed., Pergamon Press, Oxford (1977).
[3] L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1945 (1964) [Sov. Phys. JETP **20**, 1307 (1965)].
[4] B. M. Smirnov and M. I. Chibisov, Zh. Éksp. Teor. Fiz. **49**, 841 (1965) [Sov. Phys. JETP **22**, 585 (1966)].
[5] A. M. Lane and R. G. Thomas, Rev. Mod. Phys. **30**, 257 (1958).
[6] C. J. Joachain, in *Report on 7th Intern. Conf. on Multiphoton Processes*, Garmisch-Partenkirchen, Germany (1996).
[7] A. M. Perelomov, V. S. Popov, and M. V. Terent'ev, Zh. Éksp. Teor. Fiz. **50**, 1393 (1966) [Sov. Phys. JETP **23**, 924 (1966)].
[8] M. V. Ammosov, N. B. Delone, and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **91**, 2008 (1986) [Sov. Phys. JETP **64**, 1191 (1986)].
[9] S. Augst, D. Strickland, D. D. Meyerhofer, S. L. Chin, and J. H. Eberly, Phys. Rev. Lett. **63**, 2212 (1989).
[10] W. Xiong and S. L. Chin, Zh. Éksp. Teor. Fiz. **99**, 481 (1991) [Sov. Phys. JETP **72**, 268 (1991)].
[11] N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms*, Springer-Verlag, Berlin (1994).
[12] F. H. M. Faisal, Phys. Lett. A **187**, 180 (1994).
[13] Yu. N. Demkov and V. N. Ostrovskiĭ, *The Method of Zero-Radius Potentials in Atomic Physics* [in Russian], Leningrad State Univ. Press, Leningrad (1975).
[14] G. F. Drukarev, *Collisions of Electrons with Atoms and Molecules* [in Russian], Nauka, Moscow (1978).
[15] B. A. Zon, Zh. Éksp. Teor. Fiz. **102**, 36 (1992) [Sov. Phys. JETP **75**, 19 (1992)].
[16] J. K. G. Watson, Mol. Phys. **81**, 227 (1994).
[17] B. A. Zon, Phys. Lett. A **203**, 373 (1995).
[18] B. A. Zon, Laser Phys. **No. 3**, (1997).
[19] K. Müller-Dethlefs and E. W. Schlag, in *Ann. Rev. Phys. Chem.*, Vol. 42, Palo Alto, Calif. (1991), p. 109.
[20] F. Merkt and T. P. Softley, Int. Rev. Phys. Chem. **12**, 205 (1993).
[21] M. J. Siton, Rep. Prog. Phys. **70**, 167 (1982).
[22] L. P. Rapoport, B. A. Zon, and N. L. Manakov, *The Theory of Multiphoton Processes in Atoms* [in Russian], Atomizdat, Moscow (1978).
[23] V. E. Chernov and B. A. Zon, J. Phys. B **29**, 4164 (1996).
[24] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Sums, and Products*, Academic Press, New York (1980).
[25] G. Simons, J. Chem. Phys. **55**, 756 (1971).
[26] D. Normand, S. Dobosz, M. Lezius, P. D'Olivera, and M. Schmidt, in *Report on 7th Intern. Conf. on Multiphoton Processes*, Garmisch-Partenkirchen, Germany (1996).
[27] T. Seidman, M. Yu. Ivanov, and P. B. Corkum, Phys. Rev. Lett. **75**, 2819 (1995).

[28] D. L. Dorofeev and B. A. Zon, Zh. Éksp. Teor. Fiz. **110**, 882 (1996) [JETP **83**, 485 (1996)].
[29] S. L. Chin, Y. Liang, J. E. Decker, F. A. Ilkov, and M. V. Ammosov, J. Phys. B **25**, L249 (1992).
[30] B. A. Zon, Chem. Phys. Lett. **262**, 744 (1996).
[31] B. A. Zon, N. L. Manakov, L. P. Rapoport, Dokl. Akad. Nauk SSSR **188**, 560 (1969) [Sov. Phys. Dokl. **14**, 904 (1969)].

Translated by Eugene Yankovsky

# Features of the photon statistics of two coupled electromagnetic field modes under the conditions of strong mode coupling

M. E. Veĭsman[1)] and S. Yu. Kalmykov

*Moscow Physicotechnical Institute, 141700 Dolgoprudnyi, Moscow Region, Russia*
(Submitted 23 December 1996)
Zh. Éksp. Teor. Fiz. **112**, 128–136 (July 1997)

The probabilities of transitions between the Fock states of two electromagnetic field modes under the influence of coupling between modes of finite duration are investigated. It is shown that the transition probability is a strongly oscillating function of the mode numbers of the photons. Under conditions in which the coupling frequency exceeds the geometric mean of the mode frequencies (strong coupling), large numbers of photons are excited in the mode with the lower frequency. The excitation of a two-dimensional radiation field oscillator and the ''red'' asymmetry of the transition probabilities can be attributed to instability of the classical two-dimensional oscillator with strong mode coupling. © *1997 American Institute of Physics.* [S1063-7761(97)01107-4]

## 1. INTRODUCTION

The distribution functions of photons in nonclassical optical states have been studied in recent years by a number of investigators.[1] Interest in the photon statistics of nonclassical light was stimulated by the work in Ref. 2, in which the oscillating behavior of the distribution function of photons in a single-mode squeezed coherent state was detected for the first time, and it was suggested that such behavior of these functions be regarded as a sign of the nonclassical nature of the state under consideration. The squeezing of quantum fluctuations of a radiation field has been achieved experimentally using various quantum optical systems.[3] In particular, two-mode squeezing was used in Refs. 4 and 5. The distribution functions of photons in a two-dimensional squeezed vacuum[6] and the photon statistics in a two-dimensional squeezed coherent state of general form with complex displacement and squeezing parameters[7] were investigated in Refs. 6 and 7.

In the present work we find the transition probabilities between Fock states of two electromagnetic field modes with differing constant frequencies under the influence of coupling between modes of finite duration. Mode coupling can appear when coherent light propagates in a nonlinear medium with a refractive index that depends on the amplitude of the field. It is presumed that the ''coordinates'' of a two-dimensional field oscillator are coupled over the course of a restricted time interval and that the coupling frequency is a constant real quantity.[8] This quantum-mechanical model of two-mode light is exactly solvable by a method that utilizes linear integrals of the motion[9,10] of a particular quantum system with a quadratic nonstationary Hamiltonian.

Integrals of the motion that are quadratic with respect to the position and momentum operators were found for a one-dimensional quantum oscillator with a variable frequency in Ref. 11. Integrals of the motion that are linear in the position and momentum operators were obtained for a one-dimensional parametric oscillator in Refs. 12 and 13 and for a nonstationary multidimensional oscillator in Ref. 14.

Quantum integrals of the motion for the system under consideration are found in explicit form in Sec. 2. The desired transition probability can be expressed in terms of a Hermite polynomial of four variables with vanishing arguments.[15] Under the conditions of strong mode coupling, in which the coupling frequency exceeds the geometric mean of the mode frequencies, the classical two-dimensional oscillator is an unstable system that executes infinite motion in phase space, because its potential energy is nonpositive definite. It is shown that strong mode coupling of finite duration leads to strong excitation of a two-dimensional quantum oscillator with considerable ''red'' asymmetry of the transition probability as a function of the mode numbers of the photons. It is significant that these anomalies of the photon statistics are not observed in weak mode coupling, even though the coupling frequency can be low in comparison with the mode frequencies.

## 2. INTEGRALS OF THE MOTION

An explicit expression can be obtained for the transition probabilities using the general theory of multidimensional quantum systems with arbitrary quadratic Hamiltonians, which is based on their dynamic symmetries and linear integrals of the motion.[9,10] If a unitary evolution operator $\hat{U}(t)$ of the system exists, then $2N$ integrals of the motion $\hat{\mathbf{p}}_0$ and $\hat{\mathbf{q}}_0$ ($N$ is the number of degrees of freedom in the system) can be constructed via a canonical transformation of the momentum and position operators $\hat{\mathbf{p}}$ and $\hat{\mathbf{q}}$. The integrals of the motion are the initial momenta and the coordinates of the system. According to the Stone–von Neumann theorem, the operators $\hat{\mathbf{p}}_0$ and $\hat{\mathbf{q}}_0$ form a complete set.[9] The propagator (i.e., the evolution operator in a certain representation) of a nonstationary quadratic system can be expressed explicitly in terms of the elements of the canonical transformation matrix.

Two light modes with nonstationary mode coupling are described by a Hamiltonian of the form

$$\hat{H}(t)=\frac{1}{2}\hat{\mathbf{Q}}\mathbf{B}\hat{\mathbf{Q}}. \tag{1}$$

 1063-7761/97/070068-05$10.00  

Here $\hat{\mathbf{Q}}=(\hat{p}_1,\hat{p}_2,\hat{q}_1,\hat{q}_2)$ is the 4-vector of operators of the quadrature components, and $\mathbf{B}(t)$ is the $4\times4$ matrix

$$\mathbf{B}=\begin{pmatrix}\mathbf{I}_2 & 0\\ 0 & \mathbf{B}_2\end{pmatrix},\quad \mathbf{B}_2=\begin{pmatrix}\omega_1^2 & \lambda(t)\\ \lambda(t) & \omega_2^2\end{pmatrix},$$

where $\mathbf{I}_2$ is a $2\times2$ unit matrix. The coupling coefficient is a piecewise constant function, $\lambda(0<t<T)=\omega_0^2$, $\lambda(t<0,\ t>T)=0$, and $T$ is the mode coupling duration. In the general case $\omega_1\neq\omega_2$. All frequencies are real, and the masses of the oscillators are set to unity. The introduction of normal coordinates by a canonical replacement that preserves the commutation relations is impossible when $\omega_1\neq\omega_2$; therefore, the problem is solved in the original variables.

According to Refs. 9 and 10, the time-dependent quantum integrals of the motion are linear superpositions of quadrature operators

$$\hat{\mathbf{I}}(t)=\begin{pmatrix}\hat{\mathbf{p}}_0(t)\\ \hat{\mathbf{q}}_0(t)\end{pmatrix}=\mathbf{\Lambda}(t)\hat{\mathbf{Q}},\quad \mathbf{\Lambda}(t)=\begin{pmatrix}\mathbf{\Lambda}_1(t) & \mathbf{\Lambda}_2(t)\\ \mathbf{\Lambda}_3(t) & \mathbf{\Lambda}_4(t)\end{pmatrix}. \quad (2)$$

The simplectic matrix $\mathbf{\Lambda}(t)$ does not alter the commutation relations $[\hat{I}_a,\hat{I}_b]=[\hat{Q}_a,\hat{Q}_b]=-i\hbar\Sigma_{ab}$, where $a,b=1,...,4$, and has the property $\mathbf{\Lambda}^{-1}=-\mathbf{\Sigma}\mathbf{\Lambda}^T\mathbf{\Sigma}$, where

$$\mathbf{\Sigma}=\begin{pmatrix}0 & \mathbf{I}_2\\ -\mathbf{I}_2 & 0\end{pmatrix},$$

and the superscript $T$ denotes matrix transposition. A necessary and sufficient condition for invariance of the operators $\hat{\mathbf{I}}(t)$ (Eq. (2)) is their evolution according to the matrix equation

$$\dot{\mathbf{\Lambda}}=\mathbf{\Lambda}\mathbf{\Sigma}\mathbf{B} \quad (3)$$

with the initial condition $\mathbf{\Lambda}(0)=\mathbf{I}_4$, where $\mathbf{I}_4$ is the $4\times4$ unit matrix. The solution of system (3) with a piecewise constant coupling coefficient has the form

$$\mathbf{\Lambda}(t\leqslant T)=\mathbf{T}_+\otimes\mathbf{\Xi}_+-\mathbf{T}_-\otimes\mathbf{\Xi}_-,$$

where the symbol $\otimes$ denotes the tensor product of the auxiliary matrices

$$\mathbf{T}_\pm=\begin{pmatrix}\cos(\Omega_\pm t) & \Omega_\pm\sin(\Omega_\pm t)\\ -\Omega_\pm^{-1}\sin(\Omega_\pm t) & \cos(\Omega_\pm t)\end{pmatrix},$$

$$\mathbf{\Xi}_\pm=\frac{1}{a_+-a_-}\begin{pmatrix}a_\pm & 1\\ 1 & -a_\mp\end{pmatrix},$$

which are expressed in terms of the parameters $\omega_\pm^2=\omega_1^2\pm\omega_2^2$, $\Omega_\pm^2=(\omega_+^2\pm\sqrt{\omega_-^4+4\omega_0^4})/2$, $a_\pm=(\Omega_\pm^2-\omega_2^2)/\omega_0^2$, and $a_+a_-=-1$. The matrices $\mathbf{\Xi}_\pm$ satisfy the relations $\mathbf{\Xi}_\pm\mathbf{B}_2=\Omega_\pm^2\mathbf{\Xi}_\pm$ and have the properties $\mathbf{\Xi}_\pm^2=\pm\mathbf{\Xi}_\pm$ and $\mathbf{\Xi}_+\mathbf{\Xi}_-=0$. The eigenvalues $\Omega_\pm$ coincide with the eigenvalues of the Hamilton equations for a two-dimensional coupled oscillator with Hamiltonian (1). Under strong coupling conditions ($\omega_0^2>\omega_1\omega_2$), a classical two-dimensional coupled oscillator has a nonpositive definite potential energy ($\det\mathbf{B}_2<0$) and is an unstable system. In this case the eigenvalue $\Omega_-$ is purely imaginary, and is responsible for the unstable behavior of the elements of the canonical transformation matrix $\mathbf{\Lambda}$. In the opposite case of weak coupling, the classical system executes finite motion, the eigenvalues of the problem are real, and the elements of $\mathbf{\Lambda}$ are bounded.

Hamiltonian (1), expressed in terms of the ladder operators $\hat{a}_i=(\hat{q}_i\sqrt{\omega_i}+i\hat{p}_i/\sqrt{\omega_i})/\sqrt{2\hbar}$ and $\hat{a}_i^\dagger=(\hat{q}_i\sqrt{\omega_i}-i\hat{p}_i/\sqrt{\omega_i})/\sqrt{2\hbar}$, where $i=1,2$, has the form

$$\hat{H}(t)=\frac{1}{2}\hat{\mathbf{a}}\mathbf{D}\hat{\mathbf{a}},$$

where $\hat{\mathbf{a}}=(\hat{a}_1,\hat{a}_2,\hat{a}_1^\dagger,\hat{a}_2^\dagger)$ is a 4-vector, $\mathbf{D}=\mathbf{K}^T\mathbf{B}\mathbf{K}$,

$$\mathbf{K}=\sqrt{\frac{\hbar}{2}}\begin{pmatrix}-i\mathbf{E}_\omega & i\mathbf{E}_\omega\\ \mathbf{E}_\omega^{-1} & \mathbf{E}_\omega^{-1}\end{pmatrix},\quad \mathbf{E}_\omega=\begin{pmatrix}\sqrt{\omega_1} & 0\\ 0 & \sqrt{\omega_2}\end{pmatrix}.$$

The linear invariants $\hat{\mathbf{b}}=(\hat{b}_1,\hat{b}_2,\hat{b}_1^\dagger,\hat{b}_2^\dagger)$ can be constructed via a homogeneous canonical transformation of the photon creation and annihilation operators

$$\hat{\mathbf{b}}=\mathbf{\Omega}\hat{\mathbf{a}}, \quad (4)$$

where $\mathbf{\Omega}=\mathbf{K}^{-1}\mathbf{\Lambda}\mathbf{K}$. The equation $\dot{\mathbf{\Omega}}=-i\hbar^{-1}\mathbf{\Omega}\mathbf{\Sigma}\mathbf{D}$ with initial condition $\mathbf{\Omega}(0)=\mathbf{I}_4$ is a necessary and sufficient condition for the invariance of $\hat{\mathbf{b}}$ ($\mathbf{\Omega}(0)=\mathbf{I}_4$). The matrix

$$\mathbf{\Omega}=\begin{pmatrix}\zeta & \eta\\ \eta^* & \zeta^*\end{pmatrix} \quad (5)$$

consists of the $2\times2$ blocks $\zeta$ and $\eta$, which have the form

$$\zeta=\frac{1}{2}(\mathbf{E}_\omega^{-1}\mathbf{\Lambda}_1\mathbf{E}_\omega+\mathbf{E}_\omega\mathbf{\Lambda}_4\mathbf{E}_\omega^{-1})+\frac{i}{2}(\mathbf{E}_\omega^{-1}\mathbf{\Lambda}_2\mathbf{E}_\omega^{-1}-\mathbf{E}_\omega\mathbf{\Lambda}_3\mathbf{E}_\omega),$$

$$\eta=\frac{1}{2}(-\mathbf{E}_\omega^{-1}\mathbf{\Lambda}_1\mathbf{E}_\omega+\mathbf{E}_\omega\mathbf{\Lambda}_4\mathbf{E}_\omega^{-1})+\frac{i}{2}(\mathbf{E}_\omega^{-1}\mathbf{\Lambda}_2\mathbf{E}_\omega^{-1}+\mathbf{E}_\omega\mathbf{\Lambda}_3\mathbf{E}_\omega).$$

### 3. TRANSITION PROBABILITIES

The transformation (4) preserves the Bose commutation relations $[\hat{b}_i,\hat{b}_j]=[\hat{b}_i^\dagger,\hat{b}_j^\dagger]=0$ and $[\hat{b}_i,\hat{b}_j^\dagger]=\delta_{ij}$ ($\delta_{ij}$ is the Kronecker delta), making $\hat{\mathbf{b}}^\dagger$ and $\hat{\mathbf{b}}$ the formal creation and annihilation operators at some arbitrary time. The eigenvector $|\mathbf{n},t\rangle$ of the formal operators of the number of particles ($\hat{b}_1^\dagger\hat{b}_1$, $\hat{b}_2^\dagger\hat{b}_2$) in the coupled modes, which satisfies the relation $\hat{b}_i^\dagger\hat{b}_i|\mathbf{n},t\rangle=n_i|\mathbf{n},t\rangle$, where $i=1,2$, is a formal two-dimensional Fock state of two modes with nonstationary coupling. It is important to note that a discrete complete set of Fock states for a system with a nonstationary quadratic Hamiltonian can always be defined in a similar manner, even if the energy spectrum becomes continuous at a certain moment.

In particular, the formation of a coupled state is impossible in the two-dimensional potential of classical oscillator (1) with strong mode coupling ($\omega_0^2>\omega_1\omega_2$), and the energy spectrum of quantum system (1) is continuous. Nevertheless, there is still a complete set of state vectors $|\mathbf{n},t\rangle$, which are

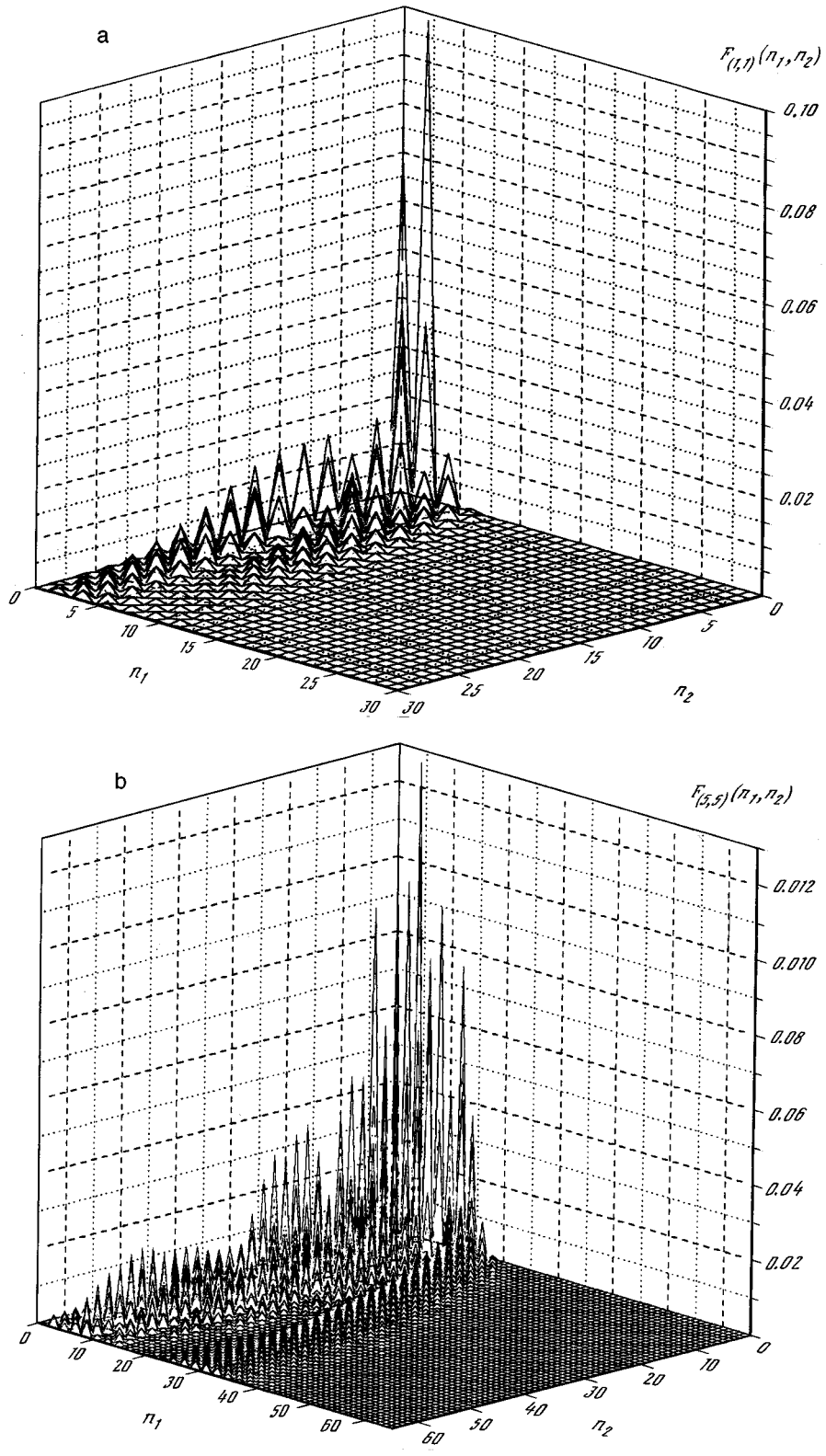



FIG. 1. Transition probability $F_{\mathbf{m}}(\mathbf{n},t)$ between Fock states of modes with very different frequencies ($\omega_1/\omega_2=3$) under strong coupling conditions ($\omega_0/\omega_2=2$, mode coupling duration $T\omega_2=\pi/2$). Case of ''red'' asymmetry of the transition probabilities from Fock states with symmetric occupation numbers: a — $|1,1,t\rangle$, b — $|5,5,t\rangle$.

eigenvectors for the integrals of the motion ($\hat{b}_1^\dagger\hat{b}_1$, $\hat{b}_2^\dagger\hat{b}_2$) and have all the algebraic properties of Fock states. It is assumed that the state $|\mathbf{n},t\rangle$ reproduces the temporal evolution of the original Fock state of the uncoupled modes $|n_1\rangle|n_2\rangle$ at some arbitrary time and coincides with it at the initial time ($t=0$):

$$|\mathbf{n},t=0\rangle=|n_1\rangle|n_2\rangle, \quad |n_i\rangle=\frac{(\hat{a}_i^\dagger)^{n_i}}{\sqrt{n_i!}}|0_i\rangle, \tag{6}$$

where $|0_i\rangle$ is the vacuum state vector of the $i$th mode.

When $t<0$ and $t\to\infty$, Hamiltonian (1) ceases to depend on time. In this case the initial and final states (6) of the stationary system of uncoupled modes exist, and transitions take place between them. The amplitude of the transition from the initial state $|in\rangle$ to the final state $|f\rangle$ (which is assumed to be given by (6)) is given by the matrix element

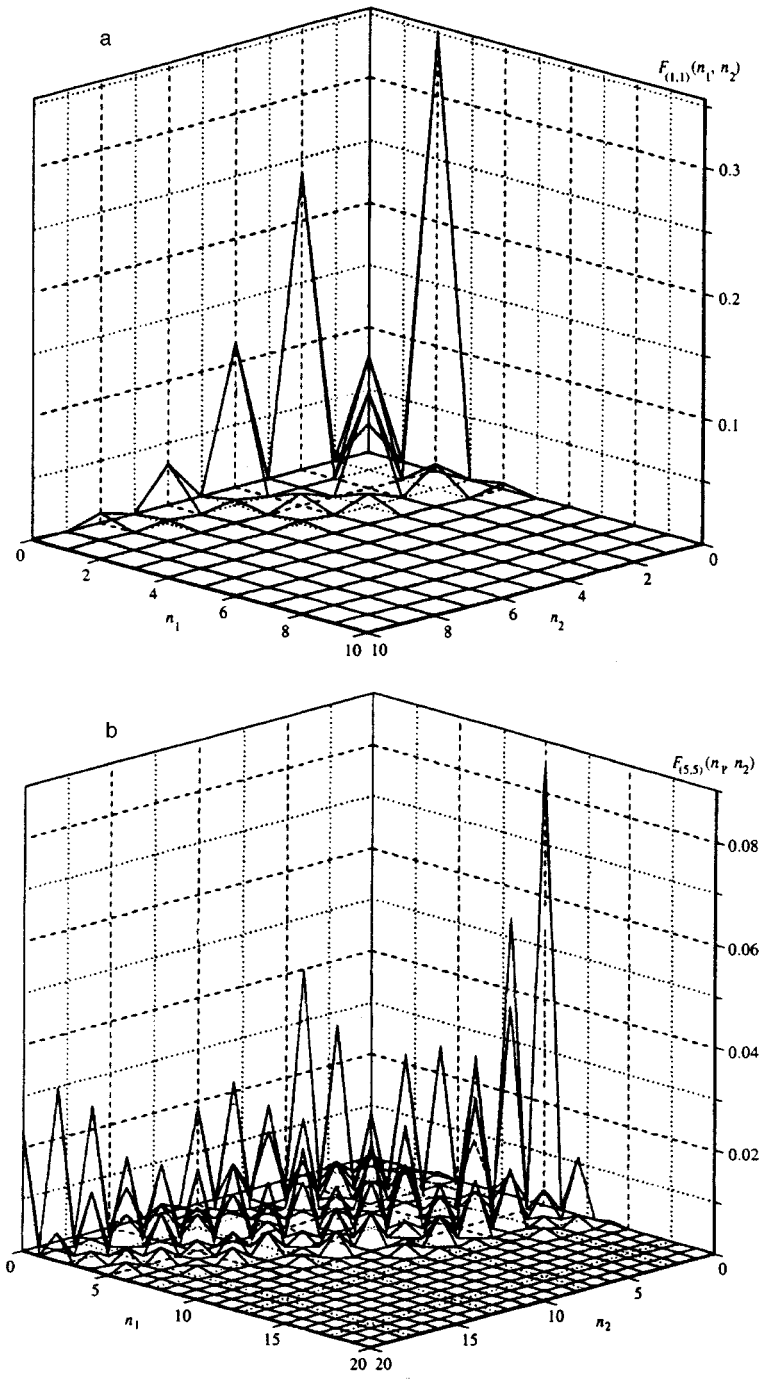



FIG. 2. Same as Figs. 1a and 1b, except that $\omega_0/\omega_2=1.5$ (weak coupling). The excitation of the field oscillators is negligible in comparison with the strong coupling case.

$\langle f|t\to\infty\rangle$, where $|t\to\infty\rangle$ is the limit of the state $|\mathbf{n},t\rangle$ as $t\to\infty$.

The amplitude $\langle\mathbf{n}|\mathbf{m},t\rangle$, which relates the initial Fock state $|\mathbf{m},t\rangle$ to the final Fock state $|\mathbf{n}\rangle$, is none other than the Green's function in the discrete Fock basis. It was represented in Refs. 10 and 16 in the form of Hermite polynomials of several variables in terms of elements of the matrix $\mathbf{\Omega}$ (Eq. (5)) of canonical transformation (4). The probability of a transition between two-mode Fock states $F_{\mathbf{m}}(\mathbf{n},t)=|\langle\mathbf{n}|\mathbf{m},t\rangle|^2$ is, by definition, the two-mode distribution function of the photons in the Fock state of the coupled modes $|\mathbf{m},t\rangle$. Here $\mathbf{m}$ is the label of the state, and $\mathbf{n}$ is a discrete vector variable. The transition probability has the form[16]

$$F_{\mathbf{m}}(\mathbf{n},t)=|\det\,\zeta(t)|^{-1}\frac{|H^{(\mathbf{Y})}_{\mathbf{nm}}(\mathbf{0})|^2}{\mathbf{n}!\,\mathbf{m}!},\tag{7}$$

where $\mathbf{m}!=m_1!m_2!$, $\mathbf{n}!=n_1!n_2!$,

$$\mathbf{Y}=\begin{pmatrix}\zeta^{-1}\eta & -\zeta^{-1}\\ -\zeta^{T^{-1}} & -\eta^*\zeta^{-1}\end{pmatrix}$$

is a $4\times4$ block matrix, and the $2\times2$ matrices $\zeta$ and $\eta$ are blocks of the simplex matrix $\mathbf{\Omega}$.

The examples of the distributions of the numbers of photons presented below were obtained numerically using the familiar recurrence relation for Hermite polynomials of $N$ variables

$$H^{\{\mathbf{R}\}}_{n_1 \ldots n_i+1 \ldots n_N}(\mathbf{y}) = \left( \sum_{k=1}^{N} R_{ik} y_k \right) H^{\{\mathbf{R}\}}_{n_1 \ldots n_i \ldots n_N}(\mathbf{y}) - \sum_{k=1}^{N} R_{ik} n_k H^{\{\mathbf{R}\}}_{n_1 \ldots n_k-1 \ldots n_N}(\mathbf{y}),$$

which starts with $H^{\{\mathbf{R}\}}_0 = 1$, $H^{\{\mathbf{R}\}}_{0 \ldots 1_i \ldots 0}(\mathbf{y}) = \Sigma^N_{k=1} R_{ik} y_k$. This relation was obtained by differentiating the product of the function for $N$-dimensional Hermite polynomials[15]

$$\exp\left( -\frac{1}{2}\mathbf{aRa} + \mathbf{aRy} \right) = \sum_{\mathbf{n}=0}^{\infty} \frac{\alpha^{\mathbf{n}}}{\mathbf{n}!} H^{\{\mathbf{R}\}}_{\mathbf{n}}(\mathbf{y}),$$

where $\alpha^{\mathbf{n}} = \alpha_1^{n_1} \alpha_2^{n_2} \ldots \alpha_N^N$, $\Sigma_{\mathbf{n}=0}( \ldots ) = \Sigma_{n_1=0} \ldots \Sigma_{n_N=0}( \ldots )$, and $\boldsymbol{\alpha} = \alpha_1, \ldots, \alpha_N$ is an $N$-dimensional vector with complex components.

The Hamiltonian (1) is represented in dimensional variables. Since all the dimensional parameters appear in the expression for the transition probability in the form of dimensionless combinations, some convenient frequency (say, a frequency in the optical range, viz., $10^{15}$ s$^{-1}$) can be taken as the nominal frequency unit.

Figures 1 and 2 show the transition probabilities $F_{\mathbf{m}}(\mathbf{n},t)$ (Eq. (7)) for $m_1 = m_2 = 1$ (Figs. 1a and 2a), $m_1 = m_2 = 5$ (Figs. 1b and 2b) photons in the initial state $|\mathbf{m},0\rangle$. Figure 1 corresponds to strong coupling of asymmetric modes ($\omega_1/\omega_2 = 3$, $\omega_0/\omega_2 = 2$, and mode coupling duration $T\omega_2 = \pi/2$). The transition probability has a pronounced asymmetric character, the maxima of the distribution being concentrated along the $n_2$ axis, which corresponds to the low-frequency mode. The excitation of large numbers of photons in the ''red'' mode occurs exclusively under strong coupling conditions ($\omega_0^2 > \omega_1\omega_2$), in which the matrix elements of the canonical transformation (4) increase exponentially. This anomaly of photon statistics is not observed under weak coupling conditions, although the coupling frequency cannot be small in comparison with the mode frequencies. This is clear in Figs. 2a and 2b, where the parameters are the same as in Figs. 1a and 1b, except that $\omega_0/\omega_2 = 1.5$ (weak coupling).

The transition probability (7) undergoes strong even-odd oscillations, since an even-dimensional Hermite polynomial (four-dimensional in the present case) with a zero argument vanishes for odd sums of its indices. This selection rule applies to arbitrary quadratic Hamiltonians of type (1) without linear ''current'' terms (these terms should add a complex displacement vector to the canonical transformation (4), making the argument of the Hermite polynomial nonzero). When the number of photons in the initial Fock state increases (see Fig. 1b), the transition probability acquires a slowly oscillating envelope in any cross section $F(n_1,n_2,t)$ for a given $n_1$ (or $n_2$). This is reminiscent of the photon statistics in a squeezed single-mode state[2] (see also the distribution functions of photons in a squeezed two-mode vacuum[6]).

## 4. CONCLUSIONS

Although the general expressions for the transition probabilities in multimode states of quadratic systems have been known for a long time,[9,10,16] the investigation of the photon statistics of specific multidimensional nonstationary systems is still of interest. The present work involved an investigation of the probabilities of transitions between Fock states of two modes with differing frequencies under the influence of strong mode coupling, where the formation of a coupled state in the two-dimensional potential of the oscillator is impossible. Excitation of the low-frequency mode has been detected under the conditions of strong mode coupling (the probability of the excitation of large numbers of photons increases). The oscillating character of the transition probability as a function of the photon mode numbers has been demonstrated, confirming the highly nonclassical nature of the formal Fock states of coupled modes.

We thank V. I. Man'ko for valuable discussions. S. Yu. K. thanks the Soros Educational Program (ISSEP) for its support (in the form of a Soros graduate stipend, Grant No. a96-1913).

[1)]e-mail: bme@hedric.msk.su

[1] V. Bužek and P. L. Knight, in *Progress in Optics*, E. Wolf (ed.), Elsevier, Amsterdam (1995), Vol. 34, p. 1; V. Bužek, G. Adam, and G. Drobný, Ann. Phys. (N.Y.) **245**, 37 (1996).
[2] W. Schleich and J. A. Wheeler, Nature (London) **326**, 574 (1987); W. Schleich, J. A. Wheeler, and D. F. Walls, Phys. Rev. A **38**, 1177 (1988).
[3] ''Squeezed states of the electromagnetic field,'' J. Opt. Soc. Am. B **4** (10), (1987).
[4] C. M. Caves and B. L. Schumaker, Phys. Rev. A **31**, 3068 (1985); B. L. Schumaker and C. M. Caves, Phys. Rev. A **31**, 3093 (1985).
[5] S. M. Barnett and P. L. Knight, J. Opt. Soc. Am. B **2**, 467 (1985).
[6] G. Schrade, V. M. Akulin, V. I. Man'ko, and W. P. Schleich, Phys. Rev. A **48**, 2398 (1993).
[7] M. Selvadoray, M. Sanjay Kumar, and R. Simon, Phys. Rev. A **49**, 4957 (1994).
[8] V. V. Dodonov, O. V. Man'ko, and V. I. Man'ko, Trudy Fiz. Inst. Akad. Nauk SSSR **191**, 224 (1989) [*Theory of Nonstationary Quantum Oscillators (Proceedings of the Lebedev Physics Institute, Academy of Sciences of the USSR, Vol. 198)*, M. A. Markov (ed.), Nova Science, Commack, N. Y. (1992)].
[9] I. A. Malkin and V. I. Man'ko, *Dynamic Symmetries and Coherent States of Quantum Systems* [in Russian], Nauka, Moscow (1979).
[10] V. V. Dodonov and V. I. Man'ko, Trudy Fiz. Inst. Akad. Nauk SSSR **183**, 182 (1987) [*Invariants and Evolution of Nonstationary Quantum Systems (Proceedings of the Lebedev Physics Institute, Academy of Sciences of the USSR, Vol. 183)*, M. A. Markov (ed.), Nova Science, Commack, N. Y. (1989)].
[11] H. R. Lewis, Phys. Rev. Lett. **18**, 510 (1967); H. R. Lewis and W. B. Riesenfeld, J. Math. Phys. (N.Y.) **10**, 1458 (1969).
[12] I. A. Malkin and V. I. Man'ko, Phys. Lett. A **32**, 243 (1970).
[13] I. A. Malkin, V. I. Man'ko, and D. A. Trifonov, Phys. Rev. D **2**, 1371 (1970).
[14] I. A. Malkin, V. I. Man'ko, and D. A. Trifonov, J. Math. Phys. (N.Y.) **14**, 576 (1973).
[15] *Bateman Manuscript Project. Higher Transcendental Functions*, A. Erdélyi (ed.), McGraw-Hill, New York (1953); V. V. Dodonov and V. I. Man'ko, J. Math. Phys. (N.Y.) **35**, 4277 (1994).
[16] V. V. Dodonov, O. V. Man'ko, and V. I. Man'ko, Phys. Rev. A **50**, 813 (1994); V. V. Dodonov, O. V. Man'ko, and V. I. Man'ko, Kratk. Soobshch. Fiz. , **43** (1994).

Translated by P. Shelnitz

# Optical pumping in a $\Lambda$-system by parametric luminescence light

D. V. Kupriyanov and I. M. Sokolov

*St. Petersburg State Technical University, 195251 St. Petersburg, Russia*
(Submitted 19 November 1996)
Zh. Éksp. Teor. Fiz. **112**, 137–162 (July 1997)

We examine the optical pumping effect in an ensemble of three-level atoms with a $\Lambda$ configuration of the energy sublevels excited by parametric luminescence light in the squeezed state. We derive quantum kinetic equations that describe the evolution of the density matrix of atoms irradiated by low-intensity squeezed light with a finite-width spectrum. In particular, we show that because of the quantum statistical properties of the squeezed light there can be a redistribution of atoms among the lower energy sublevels, despite the equality of the intensities of the spectral components of the light that resonantly excites optical transitions in the $\Lambda$-system. The relation of the optical pumping effect to the correlation and spatial–temporal spectral properties of squeezed light is discussed in detail. Finally, we show that the effects are closely linked to the finiteness of the width of the squeezed-light spectrum. © *1997 American Institute of Physics.* [S1063-7761(97)01207-9]

## 1. INTRODUCTION

The excitation of atoms by polarized or spectrally selective radiation leads to optical pumping, i.e., to the redistribution of atoms among the Zeeman, hyperfine, and other metastable sublevels, and to the creation of nonequilibrium states of atomic ensembles polarized in the internal angular momentum. Starting with the classic work of Kastler,[1] for a long time the main source of electromagnetic radiation in optical pumping experiments were spectral sources of incoherent radiation: gas-discharge tube, etc. The necessary selectivity of excitation was achieved by using optical polarizers and spectral filters. Such sources, being fairly simple and reliable, are widely used. The theory of the optical pumping effect and the various physical phenomena observed in such experiments are thoroughly described in a number of review articles (see, e.g., Refs. 2–4). Research into the interaction between radiation and an atomic ensemble polarized in angular momentum makes it possible to study the subtle features of the dynamics of collisions and other elementary processes affecting the polarization of atoms. Because of the large relaxation times and small widths of Zeeman and hyperfine transition lines, the optical pumping method combined with EPR methods has found wide application in measuring weak magnetic fields and in manufacturing quantum clocks.

The use of coherent sources of radiation has significantly broadened the area of research into physical phenomena associated with the interaction of radiation and matter. In particular, an important step in understanding the role of coherent effects in the optical pumping process was the discovery of coherent trapping of populations,[5–7] which vividly demonstrated that the behavior of a multilevel atomic ensemble controlled by coherent radiation differs significantly from the behavior of such an ensemble excited by the radiation of an incoherent spectral source. Because of the interaction with coherent fields, the optical pumping in stationary conditions is between states that are coherent superpositions of the lower atomic states. The current state of the problem and a description of the various applications of the effect of coherent population trapping can be found, say, in the review article in Ref. 8.

New horizons in studies of the optical pumping effect and related phenomena opened thanks to the discovery and manufacture of sources of squeezed light. Such radiation, which is neither coherent nor incoherent (in the ordinary sense of the word), possesses statistical properties that reveal the quantum nature of the electromagnetic field. It comes as no surprise, then, that the search for, and study of, various manifestations of the quantum properties of squeezed radiation, both in measurements and in various interactions with atomic and other media, is the topic of detailed studies in optics and spectroscopy. Possible spectroscopic applications of squeezed radiation were successfully demonstrated in experiments.[9–11] Some quantum statistical effects caused by the interaction of atoms with squeezed radiation have been predicted. For instance, in Refs. 12 and 13, a variation in the decay constants for excited atomic states placed in the field of broadband squeezed radiation with zero average amplitudes (''squeezed vacuum'') has been predicted. It has been shown theoretically[14,15] and corroborated by experiments[11] that in a two-photon optical transition the presence of strong correlations between the modes of squeezed radiation leads to a linear dependence of the population of the excited level on the light intensity in the low-intensity limit, while in accordance with classical ideas this dependence should be quadratic. A review of preliminary results (mainly theoretical) concerning the interaction of squeezed radiation with atoms can be found in Ref. 16.

In the present paper we discuss the optical pumping effect in a three-level $\Lambda$-system excited by low-intensity two-mode squeezed radiation. For definiteness we will speak of parametric luminescence light, although all our results are equally valid for other sources of squeezed light that has a low intensity and a low degree of squeezing. The principal feature in our analysis is the allowance for quantum correlations in the light interacting with the atomic subsystem from the viewpoint of their effect on the evolution of the atomic

 1063-7761/97/070073-14$10.00  

density matrix. We show that the squeezing effect in the light irradiating the medium strongly influences the equations of motion and the distribution of atoms among the sublevels. In other words, we show that in the case where light has essentially quantum statistical properties, the equation for the atomic density matrix can be reduced neither to the known optical pumping equations, valid for an incoherent spectral source, nor to the Bloch equations, describing the interaction with coherent radiation. These equations, occupying a position between these two cases, so to say, possess a number of specific properties that do not agree with the common ideas about the excitation of a three-level system by coherent or incoherent radiation.

The low-intensity approximation used in the present paper is justified not only from the standpoint of real experimental conditions but also because it allows examination of the effects caused by the finiteness of the width of the squeezed light spectrum. In most works that analyze the effect of squeezed light on the dynamics of a multilevel quantum system, the spectrum of the squeezed light was assumed to be broadband, i.e., the width was considered to be much larger than the natural widths of the excited levels.[14–19] On the other hand, the known sources of squeezed light have a finite spectral width, comparable to the natural width of the levels participating in the quantum transitions. More than that, the analysis done in Refs. 20–22 for two-level systems revealed, using this situation as an example, that the quantum statistical features in the behavior of the populations and the medium polarization, caused by the effect of the thermostat formed by the squeezed light, change significantly if the width of the fluctuation spectrum becomes comparable to the widths of the atomic levels.

Earlier the nontrivial features of the optical pumping effect in the excitation of multilevel quantum systems by light with nonclassical statistical properties and the effect of the finiteness of the width of the fluctuation spectrum were reported and briefly discussed in Ref. 23. In the present paper we use the fairly simple multilevel $\Lambda$-system to study not only the qualitative picture of the phenomenon but also to do quantitative estimates. We also note that the closely related problem, the interaction of a three-level ladder system with broadband squeezed radiation, was considered in Ref. 17, and the effect of the finiteness of the spectrum width in this case was discussed in Ref. 24. The finiteness of the fluctuation spectrum is usually taken into account by the method that uses a stochastic differential equation for the density matrix, successfully employed in studies of two-level atomic transitions.[20–22] However, the use of this method for three-level systems and more general multilevel systems is hindered by considerable technical difficulties. In this connection we note that a special methodological feature of the present work is the use of the diagram technique, which makes possible a graphical analysis of the problems caused by the finiteness of the fluctuation spectrum and their effect on the equation for the density matrix.

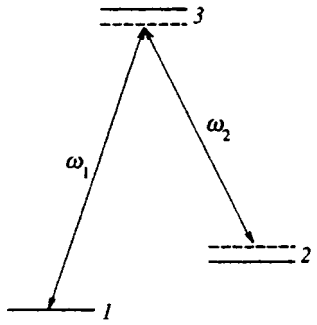



FIG. 1. Diagram of the proposed excitation of a three-level $\Lambda$-system by parametric luminescence light. The emission of two phase-associated modes characterized by median frequencies $\omega_1$ and $\omega_2$ is in quasiresonance with the optical transitions $1\rightarrow 3$ and $2\rightarrow 3$.

## 2. ANALYSIS OF THE DIAGRAMMATIC EQUATIONS FOR THE DENSITY MATRIX OF THREE-LEVEL ATOMS

In this section we perform a QED analysis of equations that describe the evolution of the density matrix of three-level atoms when the medium is irradiated by low-intensity spontaneous parametric luminescence light. Such radiation is formed in the parametric scattering process: a decay of one or two coherent pump wave photons into a pair of photons because of the three- or four-wave interaction in a nonlinear optical medium (see Ref. 25). We assume that the light emitted by the source and irradiating the atoms is a superposition of various phase-associated modes generated near median frequency $\omega_1$ and $\omega_2$. Parametric luminescence light is an example of squeezed light in which two squeezed quadrature components are superpositions of distinct-frequency components of two phase-associated modes. Here the average field amplitudes near the two frequencies $\omega_1$ and $\omega_2$ are zero. We show that in the given case, where the light has essentially quantum statistical properties (its correlation characteristics cannot be described by correlation functions with $c$-number complex-valued amplitudes), the equations for the density matrix cannot be reduced to the well-known optical pumping equations, which are valid for an incoherent spectral source. A condition important for our analysis and which a parametric source of radiation must meet consists in the assumption that the emitted light obeys quasi-Gaussian statistics. This assumption, which is valid for a light source consisting of a large number of scatterers participating in parametric oscillation, means that the averages of the product of an arbitrary number of field operators separate into products of all possible pair averages. Unlike ordinary Gaussian statistics, among the pair averages there are averages of products of operators of not only different frequencies but also of operators of the same frequency.

The diagram of the active levels and the proposed excitation of a three-level $\Lambda$-system is depicted in Fig. 1. Quasiresonant light interacts with atoms on the optical transitions $1\rightarrow 3$ and $2\rightarrow 3$. Two median frequencies $\omega_1$ and $\omega_2$, which are assumed to be close to the frequencies $\omega_{31}$ and $\omega_{32}$ of the corresponding optical transitions, are isolated in the spectrum of the driving radiation. The states 1 and 2 are assumed to be separated by a low-frequency energy interval $\omega_{32}$ and

are characterized by a short time for the populations of these sublevels to relax to their equilibrium values. In practice this means that the states 1 and 2 are either sublevels of a hyperfine structure or Zeeman sublevels separated by a magnetic or electric field.

We derive the equations for the density matrix of the atomic subsystem by employing Keldysh's diagram technique.[26,27] The atomic density matrix in the coordinate representation is determined in this technique by one-time components of the Green's function $G^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$:

$$G^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\mp i\langle\Psi^{\dagger}_{n_2}(\mathbf{r}_2t_2)\Psi_{n_1}(\mathbf{r}_1t_1)\rangle, \quad (2.1)$$

where $\Psi_{n_1}(\mathbf{r}_1t_1)$ and $\Psi^{\dagger}_{n_2}(\mathbf{r}_2t_2)$ are the atomic annihilation and creation operators in the Heisenberg representation, and the quantum numbers $n_1$ and $n_2$ correspond to internal states and in the case of the three-level system under consideration assume values 1, 2, and 3. The upper sign in (2.1) corresponds to Bose statistics and the lower sign to Fermi statistics. The equation for the atomic density matrix is set up directly by using an equation that the Green's function (2.1) must satisfy and that can be derived diagrammatically. Here, employing the relationship that links the semiclassical Wigner density matrix $\rho_{n_1n_2}=\rho_{n_1n_2}(\mathbf{p},\mathbf{r},t)$ and the function $G^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,r_2)$,

$$\rho_{n_1n_2}(\mathbf{p},\mathbf{r},t)=\pm i\int d^3\xi e^{-\frac{i}{\hbar}\mathbf{p}\boldsymbol{\xi}} \times G^{(-+)}_{n_1n_2}\left(\mathbf{r}+\frac{\boldsymbol{\xi}}{2},t;\mathbf{r}-\frac{\boldsymbol{\xi}}{2},t\right), \quad (2.2)$$

and assuming that the dependence of the density matrix on the atomic radius vector $\mathbf{r}$ and momentum $\mathbf{p}$ is smooth, we can obtain the desired equation via the following transformation:[27]

$$\frac{\partial}{\partial t}\rho_{n_1n_2}(\mathbf{p},\mathbf{r},t)+\frac{\mathbf{p}}{m}\nabla_{\mathbf{r}}\rho_{n_1n_2}(\mathbf{p},\mathbf{r},t) =\mp\int d^3\xi e^{-\frac{i}{\hbar}\mathbf{p}\boldsymbol{\xi}}[\hat{G}^{-1}_{02}{}^{*}-\hat{G}^{-1}_{01}] \times G^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2). \quad (2.3)$$

Here the action of the differential operators $\hat{G}^{-1}_{0j}$ with $j=1,2$ is defined as follows:

$$\hat{G}^{-1}_{0j}=i\frac{\partial}{\partial t_j}+\frac{\hbar}{2m}\Delta_j, \quad (2.4)$$

where $m$ is the mass of the atom, and the Laplace operator acts on the spatial coordinate $\mathbf{r}_j$. On the right-hand side of Eq. (2.3), after the operator $\hat{G}^{-1}_{01}$ and the complex-conjugate operator $\hat{G}^{-1}_{02}{}^{*}$ have acted it is assumed that the variables must be transformed in accordance with (2.2). The meaning of (2.3) is that after the analytic expression for the Green's function $G^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$, obtained as result of analyzing the corresponding diagrammatic equation, has been plugged into the right-hand side, we arrive at an equation describing the variation of the density matrix with time, an equation that under certain conditions can prove to be closed with respect to the atomic density matrix, i.e., have the form of a kinetic equation.

We study the interaction of atoms with the electromagnetic field in the dipole approximation and in the rotating wave approximation. The corresponding operator defined in the interaction representation has the form

$$V(t)=-\sum_{m,n}\int d^3r d_{mn}E^{(-)}(\mathbf{r}t)\Psi^{0\dagger}_m(\mathbf{r}t)\Psi^0_n(\mathbf{r}t)+\text{H.c.} \quad (2.5)$$

Here $d_{mn}$ is the dipole-moment matrix element of the transition between the lower states $n=1,2$ and the excited state $m=3$, and $E^{(-)}(\mathbf{r}t)$ is the operator of the negative-frequency field-strength component. We use the definition of frequency components of the field adopted in Ref. 28, $E^{(\pm)}\propto e^{\pm i\omega t}$, i.e., the plus and minus are associated with photon creation and annihilation, respectively. The time evolution of the operators $\Psi^{0\dagger}_m(\mathbf{r}t)$, $\Psi^0_n(\mathbf{r}t)$, and $E^{(\pm)}(\mathbf{r}t)$ is determined by the Heisenberg equations with unperturbed atom and field Hamiltonians: $H_0=H_a+H_f$. In the interaction representation the Green's function $G^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ can be written

$$G^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2) =\mp i\langle\tilde{T}(\hat{S}^{-1}\Psi^{0\dagger}_{n_2}(\mathbf{r}_2t_2))T(\Psi^0_{n_1}(\mathbf{r}_1t_1)\hat{S})\rangle, \quad (2.6)$$

where the evolution operator $\hat{S}$ is defined as

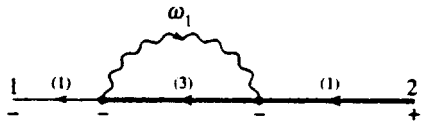
$$\hat{S}=\hat{S}(\infty,-\infty)=T\exp\left(-\frac{i}{\hbar}\int_{-\infty}^{\infty}V(t)dt\right). \quad (2.7)$$

In these expressions, $T$ and $\hat{T}$ are the operators of time ordering and antiordering. The expansion of the evolution operators in (2.6) in perturbation series generates a corresponding expansion for the Green's function $G^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$, which can be depicted diagrammatically. The definition and properties of the Green's functions for atoms and for the field, which appear in diagrammatic expansions and correspond to different types of time ordering, are discussed in Appendices A and B.

We examine the various diagram contributions obtained by summing diagram sequences using the function $G^{(-+)}_{11}\times(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ as an example and establish, in accordance with (2.3), the corresponding contributions in the equation describing the time variation of the density matrix $\rho_{11}(\mathbf{p},\mathbf{r},t)$. Bearing in mind Eq. (2.3), we must analyze in the diagram expansion of the Green's function $G^{(-+)}_{11}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ only those diagrams leading to analytic contributions that do not vanish as a result of the action of the operators $\hat{G}^{-1}_{01}$ and $\hat{G}^{-1}_{02}{}^{*}$. To simplify the calculations, we use an optically thin medium and do not take into account the effect of the atoms on the state of the field. What is important for our analysis is the assumption that the intensity of the irradiating light is low, which limits the possible set of self-energy diagrams. Physically this assumption means that the light intensity must be low compared to the saturation intensity on the $1\rightarrow3$ and $2\rightarrow3$ transitions (the stimulated transition rate is low compared to the rate $\gamma$ of the spontaneous decay of the excited state). At the same time, the inten-

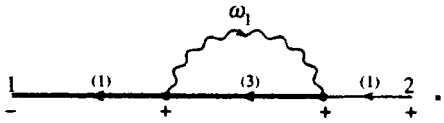
sity must be sufficiently high to induce a redistribution of populations among the sublevels 1 and 2 (the optical pumping effect). There exists, however, a much more important criterion of when the intensity of the irradiating light is low. This criterion will be formulated later.

Some of the diagram sequences can be represented by the sums




(2.8)

or




(2.9)

Here and in what follows light lines depict unperturbed Green's functions, heavy lines the exact atomic Green's functions, and wavy lines the Green's functions of the irradiating light (see Appendices A and B). The diagram vertices correspond to dipole-moment matrix elements. For convenience, an atomic line is labeled by the number of the state and a photon line by the corresponding median frequencies. Note that some of the terms in (2.8) coincide with some of the terms in (2.9). But as a result of the subsequent transformation (2.3) only one of the diagram sums, (2.8) or (2.9), provides a finite contribution, depending on whether we consider the action of the operator $\hat{G}_{01}^{-1}$ or of the operator $\hat{G}_{02}^{-1*}$. In the equation for the density matrix the diagrams describe a stimulated process in which the atoms leave state 1, with the rate of depopulation proportional to the intensity of the irradiating light:

$$\left(\frac{\partial\rho_{11}}{\partial t}\right)_{d.p.} = -\int_0^\infty \frac{d\omega}{2\pi}\frac{2\pi\omega}{c\hbar}|d_{13}|^2 \times\frac{\gamma}{(\omega-\omega_{31}-\mathbf{k}_1\mathbf{p}/m)^2+\gamma^2/4}J_1(\omega)\rho_{11}(\mathbf{p},\mathbf{r},t). \tag{2.10}$$

Here $J_1(\omega)$ is the photon spectral flux density near the frequency $\omega_1$ in the direction characterized by the wave vector $\mathbf{k}_1$; it can be expressed in terms of the time spectrum of the field correlation function $\Phi^{(+-)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ defined in Appendix B, formula (B7). If we write the correlation function as

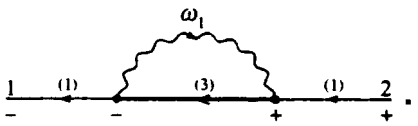
$$\Phi^{(+-)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\Phi_{11}^{(+-)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2) + \Phi_{22}^{(+-)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2), \tag{2.11}$$

where the first term corresponds to the contribution of the radiation at frequency $\omega_1$ and the second term to the contribution at frequency $\omega_2$, then $J_1(\omega)$ is represented by the spectral decomposition of the function $\Phi_{11}^{(+-)}(\mathbf{r},t_1;\mathbf{r},t_2)=\Phi_{11}^{(+-)}(t_1-t_2)$ as follows:

$$J_1(\omega)=\frac{c}{2\pi\hbar\omega}\int d\tau e^{i\omega\tau}\Phi_{11}^{(+-)}(\tau). \tag{2.12}$$

In the general equations of optical pumping for the case of a multilevel atom, the polarization of the ground state described by a depopulation term of the form (2.10) is sometimes called depopulation pumping.[2,29]

The diagram sum of the form




(2.13)

describes the spontaneous and stimulated transitions of atoms from level 3 to state 1 in the equation for the density matrix. If we assume that the excited state 3 can be populated only through stimulated transitions from sublevel 1 or 2 and that the intensity of optical pumping is low, then allowing for inverse stimulated $3\rightarrow1$ transitions only makes the accuracy of the calculations too high. Thus, the diagram (2.13) leads to the following term in the equation for the density matrix $\rho_{11}$:

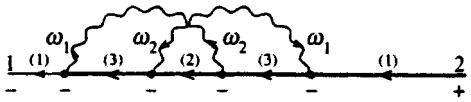
$$\left(\frac{\partial\rho_{11}}{\partial t}\right)_{r.p.}=\gamma_{31}\rho_{33}(\mathbf{p},\mathbf{r},t). \tag{2.14}$$

Here $\gamma_{31}$ is the rate of spontaneous decay from level 3 to level 1. In the general theory of the optical pumping of a multilevel atom, the polarization of the ground state described by the repopulation term of type (2.14) is sometimes called repopulation pumping.[2,29]

By employing the diagram technique we have derived the well-known contributions entering into an ordinary kinetic balance equation. This was done primarily so as later to compare these terms with additional terms that emerge in the equation for the density matrix of atoms interacting with spontaneous parametric luminescence light. Here we must note, however, that the Markovian nature of the evolution of the atomic density matrix, determined by the derivatives (2.10) and (2.14), can be related to the fact that in interpreting the diagrams we used the expressions of the atomic Green's functions in terms of the atomic Wigner density matrix via Eqs. (A4)–(A7) of Appendix A. A necessary prerequisite for such a procedure is the requirement that the diagonal components of the density matrix for states 1 and 2 do not change on the time scale on which the field correlation functions decay and that the time variation of the population of the excited state on the same scale be determined solely by the spontaneous decay process. These assumptions, which in the case of the contributions (2.8), (2.9), and (2.13) are actually reduced to the requirement that the width of the exciting light spectrum be larger than the stimulated transition rates and the relaxation times for states 1 and 2, are usually true, so that the expressions (2.10) and (2.14) are fairly accurate. What is important here is that the width of the spectrum can be much smaller than the atomic natural linewidth. In this case the kinetic stage in the evolution of the system is determined by the stimulated transition rates

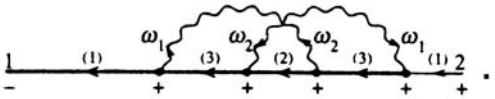
and the relaxation times for states 1 and 2. The time variation of the density matrix of the excited state, $\rho_{33}$, on the slow time scale will adiabatically follow the components $\rho_{11}$ and $\rho_{22}$ of the ground-state density matrix.

Some of the diagram sequences containing the anomalous Green's functions of the electromagnetic field (see the definitions (B3)–(B6) of Appendix B) can be represented by the sums




(2.15)

or




(2.16)

Just as with the diagrams (2.8) and (2.9), some of the terms in (2.15) coincide with some of the terms in (2.16). But as a result of the subsequent transformation (2.3) only one of the diagram sums provides a finite contribution, depending on whether we consider the action of the operator $\hat{G}_{01}^{-1}$ or of the operator $\hat{G}_{02}^{-1*}$. The equation for the density matrix acquires the following term:

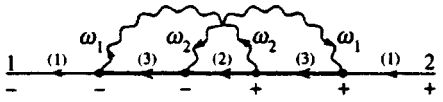
$$\left(\frac{\partial\tilde{\rho}_{11}}{\partial t}\right)_{d.p.} = \pm\int d^3\xi e^{-\frac{i}{\hbar}\mathbf{p}\boldsymbol{\xi}}\int\int\int d^3r_4dt_4d^3r_5dt_5d^3r_6dt_6$$
$$\times\frac{1}{\hbar^4}|d_{13}|^2|d_{23}|^2 iG_{33}^{(--)}\left(\mathbf{r}+\frac{\boldsymbol{\xi}}{2},t;\mathbf{r}_4,t_4\right)$$
$$\times iG_{22}^{(--)}(\mathbf{r}_4,t_4;\mathbf{r}_5,t_5)iG_{33}^{(--)}(\mathbf{r}_5,t_5;\mathbf{r}_6,t_6)$$
$$\times iG_{11}^{(-+)}\left(\mathbf{r}_6,t_6;\mathbf{r}-\frac{\boldsymbol{\xi}}{2},t\right)$$
$$\times\Phi_{12}^{(++)}\left(\mathbf{r}+\frac{\boldsymbol{\xi}}{2},t;\mathbf{r}_5,t_5\right)$$
$$\times\Phi_{21}^{(--)}(\mathbf{r}_4,t_4;\mathbf{r}_6,t_6)+\text{c.c.}, \qquad (2.17)$$

where c.c. stands for the complex-conjugate contribution. Here we used the notation for the atomic Green's functions and the field anomalous correlation functions defined in Appendices A and B, with the anomalous correlation functions, which are defined in (B9) and (B10), represented in the form

$$\Phi^{(--)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\Phi_{12}^{(--)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$$
$$+\Phi_{21}^{(--)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2), \qquad (2.18)$$
$$\Phi^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\Phi_{12}^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$$
$$+\Phi_{21}^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2),$$

where the first term corresponds to the contribution of beats near the median frequencies $\omega_1$ and $\omega_2$ in the time arguments $t_1$ and $t_2$, and the second terms to the contribution of beats near the median frequencies $\omega_2$ and $\omega_1$.

Another diagram sum determined by anomalous correlation functions,




(2.19)

leads to an analytic expression of the following form in the equation for the density matrix:

$$\left(\frac{\partial\tilde{\rho}_{11}}{\partial t}\right)_{r.p.} = \pm\int d^3\xi e^{-\frac{i}{\hbar}\mathbf{p}\boldsymbol{\xi}}$$
$$\times\int\int\int d^3r_4dt_4d^3r_5dt_5d^3r_6dt_6$$
$$\times\frac{1}{\hbar^4}|d_{13}|^2|d_{23}|^2 iG_{33}^{(--)}\left(\mathbf{r}+\frac{\boldsymbol{\xi}}{2},t;\mathbf{r}_4,t_4\right)$$
$$\times iG_{22}^{(-+)}(\mathbf{r}_4,t_4;\mathbf{r}_5,t_5)iG_{33}^{(++)}(\mathbf{r}_5,t_5;\mathbf{r}_6,t_6)$$
$$\times iG_{11}^{(++)}\left(\mathbf{r}_6,t_6;\mathbf{r}-\frac{\boldsymbol{\xi}}{2},t\right)\Phi_{12}^{(++)}\left(\mathbf{r}\right.$$
$$\left.+\frac{\boldsymbol{\xi}}{2},t;\mathbf{r}_5,t_5\right)\Phi_{21}^{(--)}(\mathbf{r}_4,t_4;\mathbf{r}_6,t_6)+\text{c.c.}, \qquad (2.20)$$

where the notation is the same as is in (2.17). Note that actually neither (2.17) nor (2.20) depends on the quantum statistics of the atoms, which follows from the properties of the atomic Green's functions and the assumed weakness of the degeneracy effects in the semiclassical setting. Here all the functions of the form $G^{(--)}$ and $G^{(++)}$ must be considered in these expressions as referring to the single-particle case, i.e., they should actually be the unperturbed advanced or retarded Green's functions for one atom. Allowance for the corrections to the advanced and retarded Green's functions that reflect the atom–field interaction leads only to excessive accuracy in the present calculations.

We denote the contributions (2.17) and (2.20) in the equation describing the time variation of the density matrix, by analogy with (2.10) and (2.14), as additional contributions to depopulation and repopulation processes. From the classical viewpoint they correspond to higher-order corrections in the light intensity and can be dropped. An important fact that forces us to keep these contributions is that in a quantum setting involving squeezed light with a low degree of squeezing (parametric luminescence light is just such light) the anomalous correlation functions are much larger than the normal correlation functions. This situation is graphically illustrated by condition (B15) in Appendix B, which is valid for dimensionless Fourier transforms of the correlation functions. Generally speaking, the contributions (2.17) and (2.20) are of the same order of smallness in the light intensity as the contributions (2.10) and (2.14) and should also be taken into account when the interaction of atoms and the pump light is described in the low-intensity limit. The criterion of weakness of the pump light in this case is the low degree of squeezing of the radiation emitted by the parametric source, which ensures the validity of (B15). This condition can be

met if we use an optically thin nonlinear parametric medium as the source.[25] We also note that the appearance in the equations describing optical pumping by squeezed light of terms corresponding to the linear dependence of two-stage transitions on the intensity of the irradiating light in the event of atom excitation by low-intensity light has been reported by Parkins.[23]

Another important feature setting the contributions (2.17) and (2.20) apart from (2.10) and (2.14) is that it is more difficult to indicate the terms corresponding to the atomic density matrix. For the contributions (2.17) and (2.20) to acquire the form of terms in a kinetic equation we must justify the use of the representation (A4)–(A7) in the atomic Green's functions. Generally, when the correlation time $\tau_c$ determining the decay of the field anomalous correlation functions is comparable to or exceeds the transition times determining the evolution of the atomic subsystem, the use of this representation is unjustified. Moreover, at large correlation times the diagrams (2.15), (2.16), and (2.19) may become more complicated in view of the possible coupling of single-particle atomic lines into unfactorable two-particle Green's functions through the interaction with the thermostat. The terms (2.17) and (2.20) in the equation for the density matrix generally lead to a non-Markovian evolution of the atomic subsystem. The evolution may become Markovian in the limit $\tau_c \ll \tau_0$, where $\tau_0$ is the characteristic relaxation time for the difference in populations of the lower sublevels. Note that the question of whether the evolution of the density matrix of the atomic subsystem is Markovian or non-Markovian is related to the quantum statistical properties of the pump radiation.

There is one more important property of the diagrams (2.15), (2.16), and (2.19). It is related to the fact that they can be interpreted as diagrams containing vertex corrections to the initial point vertex of the interaction corresponding to the dipole moments of the transitions. Vertex corrections also exist for ordinary incoherent Gaussian radiation, but their contribution is small due to the smallness of the degeneracy parameter for this radiation. For quantum radiation with a low degree of squeezing (the case examined here), the degeneracy parameter is small, too, but the vertex corrections are nevertheless important because they can be expressed in terms of the field anomalous correlation functions. Clearly, setting up an equation for the atomic subsystem in the case of radiation with an arbitrary degree of squeezing requires allowing for all vertex corrections and analyzing higher-order terms in the interaction with the optical pumping radiation in the perturbation-theory series. The possibility of keeping only the corrections (2.15), (2.16), and (2.19) to the interaction vertices and, as a result, of deriving a relatively simple equation for the atomic density matrix is a special feature of the case of radiation with a low degree of squeezing.

The final equation for the density matrix $\rho_{11}=\rho_{11}(\mathbf{p},\mathbf{r},t)$ can be written

$$\frac{\partial}{\partial t}\rho_{11}(\mathbf{p},\mathbf{r},t)+\frac{\mathbf{p}}{m}\nabla_{\mathbf{r}}\rho_{11}(\mathbf{p},\mathbf{r},t)=\left(\frac{\partial\rho_{11}}{\partial t}\right)_{d.p.}+\left(\frac{\partial\rho_{11}}{\partial t}\right)_{r.p.}$$
$$+\left(\frac{\partial\tilde{\rho}_{11}}{\partial t}\right)_{d.p.}+\left(\frac{\partial\tilde{\rho}_{11}}{\partial t}\right)_{r.p.}+\left(\frac{\partial\rho_{11}}{\partial t}\right)_{R}, \qquad (2.21)$$

where the last term is responsible for relaxation processes. Note that the diagram expansion (and its analytic interpretation) in the case of the equation for the density matrix $\rho_{22}$ can be examined in exactly the same way as in the case with $\rho_{11}$ due to the symmetry of the problem with respect to states 1 and 2. The corresponding analytic expression can be obtained from (2.10), (2.14), (2.17), and (2.20), with the labels 1 and 2 are interchanged. The variation of the density matrix $\rho_{33}$ of the excited state also require no separate analysis since it is, obviously, described by the sum of contributions for the lower sublevels 1 and 2 multiplied by $-1$, due to the conservation of the total population of all sublevels in the optical interaction cycle. A simple analysis of the possible diagrams shows that the equations for the off-diagonal components of the density matrix, $\rho_{n_1n_2}=\rho_{n_1n_2}(\mathbf{p},\mathbf{r},t)$ $(n_1 \neq n_2)$, are homogeneous and unrelated to the populations in the system. This means that there is no excitation of the coherence between the sublevels of the $\Lambda$-system when only spontaneous parametric luminescence light is used. Exciting the optical coherences $\rho_{31}$ and $\rho_{32}$ and the radio-frequency coherence $\rho_{21}$ requires using external sources of light or radio-frequency field. The dynamics of the density matrix determined by Eq. (2.21) for $\rho_{11}$ and similar equations for $\rho_{22}$ and $\rho_{33}$ must differ dramatically from the dynamics described by the cases of incoherent and coherent optical pumping. It is also important to note that the corresponding differences in the dynamics of the evolution of populations of three-level atoms are due primarily to the quantum nature of light and cannot be explained from the classical viewpoint.

### 3. THE KINETIC EQUATION APPROXIMATION

Let us transform Eq. (2.21) and similar equations for $\rho_{22}$ and $\rho_{33}$ to a system of kinetic equations. Usually, in going over to the kinetic equation approximation, it is assumed that the excitation spectrum of incoherent radiation (squeezed, in particular) is broadband compared to the characteristic relaxation rates in the system (see Refs. 12 and 16). In this section we derive a system of kinetic equations under other conditions, specific to the three-level system considered here, in which the characteristic relaxation time $\tau_0$ of the ground state (sublevels 1 and 2) is much longer than the natural lifetime $\gamma^{-1}$ of the excited state 3. Let us assume that the width of the spectrum of both modes of the excitation radiation, $\Delta\omega$, and the reciprocal relaxation time $\tau_c^{-1}$ (which determines the width of the spectrum of the anomalous correlation functions) are much smaller than $\gamma$. At the same time, the evolution has a kinetic stage only in the limit $\Delta\omega$, $\tau_c^{-1} \gg \tau_0^{-1}$. These inequalities are sufficient for the kinetic approximation to describe the evolution of the system for times comparable to $\tau_0$. Let us limit our discussion to the case of cold atoms evenly distributed over the interaction volume. Here we assume that the Doppler shift of optical transition lines is much smaller than the natural width $\gamma$. We estimate the effect of atomic motion at the end of our analysis. Under these conditions the kinetic equation can be intro-

duced directly for the diagonal components of the density matrix of the internal state, $\rho_{nn}(t)$ $(n=1,2,3)$:

$$\rho_{nn}(t)=\frac{1}{n_0}\int\frac{d^3pd^3r}{(2\pi\hbar)^3}\rho_{nn}(\mathbf{p},\mathbf{r},t), \tag{3.1}$$

where $n_0$ is the atom number density.

Using the representation of the atomic Green's functions given in Approximation A, we represent the contributions (2.17) and (2.20) as terms of a kinetic equation. As a result, the variation of the populations $\rho_{nn}=\rho_{nn}(t)$ $(n=1,2,3)$ in the $\Lambda$-system is described by the following system of kinetic equations:

$$\begin{aligned}
&\frac{\partial}{\partial t}\rho_{11}=-w_1\rho_{11}+\gamma_{31}\rho_{33}+\tilde{w}_1\rho_{11}+\tilde{\gamma}_{21}\rho_{22}-\Gamma_{12}\rho_{11}+\Gamma_{21}\rho_{22},\\
&\frac{\partial}{\partial t}\rho_{22}=-w_2\rho_{22}+\gamma_{32}\rho_{33}+\tilde{w}_2\rho_{22}+\tilde{\gamma}_{12}\rho_{11}-\Gamma_{21}\rho_{22}+\Gamma_{12}\rho_{11},\\
&\frac{\partial}{\partial t}\rho_{33}=-\frac{\partial}{\partial t}\rho_{11}-\frac{\partial}{\partial t}\rho_{22}.
\end{aligned} \tag{3.2}$$

We placed the various contributions on the right-hand sides of these equations in the same sequence as in Eq. (2.21). In particular, the last two terms describe collisional (nonradiative) transitions between the sublevels 1 and 2, characterized by the constants $\Gamma_{12}$, $\Gamma_{21}\sim\tau_0^{-1}$. The last equation in (3.2) means that the $\Lambda$-system is closed, so that the rate of spontaneous decay of the excited state is the sum of the rates of decay into states 1 and 2: $\gamma=\gamma_{31}+\gamma_{32}$.

Let us discuss in detail the various contributions in Eqs. (3.2) that describe light-induced transitions in the $\Lambda$-system. By $w_n$ $(n=1,2)$ we denote the ordinary stimulated transition rates, which correspond to the use of the classical time-dependent perturbation theory:

$$w_n=w_n(\omega_n)=\sigma'_{n3}(\omega_n)J_n. \tag{3.3}$$

These transitions are characterized by an absorption cross section $\sigma'_{n3}(\omega)$ of the $n\to3$ transition and a photon flux density in the $n$th mode,

$$J_n=\frac{c}{2\pi\hbar\omega_n}\Phi_{nn}^{(+-)}(\mathbf{r},t;\mathbf{r},t). \tag{3.4}$$

The cross section $\sigma'_{n3}(\omega)$ is the real part of the complex-valued interaction cross section of the $n\to3$ transition:

$$\begin{aligned}
\sigma_{n3}(\omega)&=\sigma'_{n3}(\omega)+i\sigma''_{n3}(\omega)\\
&=\frac{4\pi\omega}{c\hbar}|d_{n3}|^2\frac{1}{-i(\omega-\omega_{3n})+\gamma/2}.
\end{aligned} \tag{3.5}$$

By representing the $w_n$ in the form (3.3) we imply that excitation occurs in the vicinity of the median frequency $\omega_n\sim\omega_{3n}$ in a spectral interval small compared to the natural atomic linewidth.

The additional depopulation pumping terms emerging as a result of excitation of the atomic subsystem through the field anomalous correlation functions are characterized by the rates $\tilde{w}_n$ $(n=1,2)$:

$$\begin{aligned}
\tilde{w}_n&=\tilde{w}_n(\omega_1,\omega_2,\Omega)\\
&=\tfrac{1}{4}(\sigma_{13}(\omega_1)\sigma_{23}(\omega_2)G_{n\bar{n}}(\Omega)\\
&\quad+\sigma_{13}^*(\omega_1)\sigma_{23}^*(\omega_2)G_{n\bar{n}}(\Omega)),
\end{aligned} \tag{3.6}$$

where $\Omega=\omega_1-\omega_2$, and $\bar{n}=2,1$ for $n=1$ or 2, respectively. The electromagnetic field enters into this expression through the spectral functions $G_{12}(\Omega)$ and $G_{21}(\Omega)$ defined as follows:

$$\begin{aligned}
G_{12}(\Omega)&=\frac{c^2}{(2\pi\hbar)^2\omega_1\omega_2}\int_0^\infty d\tau e^{-i\omega_{21}\tau}\\
&\quad\times\Phi_{12}^{(++)}(\mathbf{r},t_1;\mathbf{r},t_2)\Phi_{21}^{(--)}(\mathbf{r},t_1;\mathbf{r},t_2),\\
G_{21}(\Omega)&=\frac{c^2}{(2\pi\hbar)^2\omega_1\omega_2}\int_0^\infty d\tau e^{i\omega_{21}\tau}\\
&\quad\times\Phi_{21}^{(++)}(\mathbf{r},t_1;\mathbf{r},t_2)\Phi_{12}^{(--)}(\mathbf{r},t_1;\mathbf{r},t_2),
\end{aligned} \tag{3.7}$$

where $\omega_{21}$ is the transition frequency between the sublevels 2 and 1. In (3.7) we have allowed for the fact that the integrands, which depend only on the difference in the time arguments, $\tau=t_1-t_2$, oscillate at frequency $\Omega$. Actually the functions $G_{12}(\Omega)$ and $G_{21}(\Omega)$ depend on $\Omega$ through the difference $\Omega-\omega_{21}$. The two functions are linked by the symmetry relation $G_{21}(\Omega)=G_{12}^*(\Omega)$.

Note the two facts that set the additional rate parameter $\tilde{w}_n$ apart from the ordinary rate of stimulated transitions to the excited state, $w_n$, defined by (3.3). First, $\tilde{w}_n$ has no fixed sign, i.e., can be either positive or negative, and can, in this way, effectively decrease or increase the rate at which light-induced atoms leave the ground state. This should come as no surprise, since these depopulation rate constants are higher-order corrections to $w_n$, so that only the difference $w_n-\tilde{w}_n$ must be positive. But since both quantities are of the same order in the light intensity, the difference can be much smaller or much larger than the ordinary rate $w_n$ of stimulated transitions. Second, the rate parameter $\tilde{w}_n$ depends in a resonant manner not only on the frequencies $\omega_1$ and $\omega_2$ but also on their difference $\Omega$. The resonant dependence on $\Omega$ is represented by the corresponding dependence of the functions $G_{12}(\Omega)$ and $G_{21}(\Omega)$, which are the spectral decompositions of the product of the anomalous correlation functions at the atomic transition frequency $\omega_{21}$. Resonant buildup occurs where the frequency of the natural beats of the integrands in (3.7) becomes comparable to the transition frequency $\omega_{21}$. Here the resonance width is determined by the decay time of the anomalous correlations in the exciting light.

The additional repopulation contributions, which describe the direct redistribution of atoms among the sublevels 1 and 2 induced by the introduction through the field anomalous correlation functions, are characterized in Eqs. (3.2) by the rate constants $\tilde{\gamma}_{21}$ and $\tilde{\gamma}_{12}$. These constants have the form

$$\tilde{\gamma}_{n\bar{n}}=\tilde{\gamma}_{n\bar{n}}(\omega_1,\omega_2,\Omega)$$
$$=\tfrac{1}{4}(\sigma_{n3}(\omega_n)\sigma^*_{\bar{n}3}(\omega_{\bar{n}})G_{\bar{m}m}(\Omega)$$
$$+\sigma^*_{n3}(\omega_n)\sigma_{\bar{n}3}(\omega_{\bar{n}})G^*_{\bar{m}m}(\Omega)), \qquad (3.8)$$

$n=1, 2$, $\bar{n}=2, 1$, and the notation is the same as in (3.6). Clearly, $\tilde{\gamma}_{21}=\tilde{\gamma}_{12}$. The constants (3.8) can have an arbitrary sign (just as (3.6) can). However, one must allow for the fact that the corresponding contribution in the kinetic equations (3.2) must be considered with the repopulation terms describing the spontaneous decay of the excited state, and their total contribution must correspond to the total positive probability of the light-induced transitions between the sublevels 1 and 2.

The system of kinetic equations (3.2) obtained in this section describes the evolution of populations in the $\Lambda$-system and differs dramatically from the equations describing the optical pumping of an atomic ensemble by an incoherent spectral source[2] and from the Bloch equations describing optical pumping by light from a coherent source. We note again that there is no way in which the given equations, valid at low intensities of the irradiating light with a zero average amplitude, can be derived on the basis of classical ideas about fluctuations of the field strength. All the additional rate parameters, $\tilde{w}_1$, $\tilde{w}_2$, $\tilde{\gamma}_{21}$, and $\tilde{\gamma}_{12}$, are quantities proportional to the exciting light intensity, just as the ordinary stimulated transition rates $w_1$ and $w_2$ are. Equations (3.2) have a fairly narrow applicability range because of the stringent conditions imposed on the source of parametric radiation. Below we formulate these condition and then formally solve the system of equations for the stationary case.

## 4. STATIONARY SOLUTION: THE EFFECT OF OPTICAL PUMPING WITH EQUAL INTENSITIES OF THE EXCITING LIGHT IN THE LEGS OF THE $\Lambda$-SYSTEM

In the stationary case the solution of the system of equations (3.2), which describes the distribution of atoms among the sublevels of the $\Lambda$-system, parametrically depends on the frequencies $\omega_1$ and $\omega_2$ and on the difference $\Omega$. From the standpoint of classical perturbation theory, a characteristic feature of the excitation scheme being discussed is that at equal differences $\omega_1-\omega_{31}=\omega_2-\omega_{32}$, $\Omega=\omega_{21}$, and at equal excitation intensities in the legs of closed the $\Lambda$-system ($\gamma=\gamma_{31}+\gamma_{32}$) the optical pumping effect must be reduced to an equalization of populations on the sublevels 1 and 2, provided that these sublevels are close to each other ($\omega_{21}\ll\omega_{31},\omega_{32}$) and that $\Gamma_{12}$, $\Gamma_{21}\ll w_1, w_2\ll\gamma$. This fundamental property of the three-level $\Lambda$-system becomes invalid in the case described by the system of kinetic equations (3.2).

We prove the above statement by solving the system of equations (3.2) for the resonant case with $\omega_1=\omega_{31}$, $\omega_2=\omega_{32}$, and $\Omega=\omega_{21}$. Comparing (3.6) and (3.8), we notice that

$$\tilde{w}_1=\tilde{w}_2=\tilde{\gamma}_{12}=\tilde{\gamma}_{21}\equiv\tilde{w}. \qquad (4.1)$$

After plugging (4.1) into (3.2) it is easy to obtain an exact analytic solution for the stationary system of equations derived from (3.2). This is unnecessary, however, since keeping only the leading terms in the expansion of the solution in a series in light intensity is sufficient. To this accuracy we obtain

$$\rho_{11}\to\frac{\gamma_{31}}{w_1}\frac{w_1w_2-(w_2+w_1)\tilde{w}}{\gamma_{31}w_2+\gamma_{32}w_1}+\frac{\tilde{w}}{w_1},$$
$$\rho_{22}\to\frac{\gamma_{32}}{w_2}\frac{w_1w_2-(w_2+w_1)\tilde{w}}{\gamma_{31}w_2+\gamma_{32}w_1}+\frac{\tilde{w}}{w_2}, \qquad (4.2)$$
$$\rho_{33}\to\frac{w_1w_2-(w_2+w_1)\tilde{w}}{\gamma_{31}w_2+\gamma_{32}w_1},$$

where the $\to$ symbol means that we need only the leading terms in the expansions, assuming $\Gamma_{12}$, $\Gamma_{21}\ll w_1$, $w_2$, $\tilde{w}\ll\gamma$. In view of this, the normalization condition for the components of the density matrix in (4.2) means that the sum of $\rho_{11}$ and $\rho_{22}$ is unity to within quantities of order $w_1/\gamma$, $w_2/\gamma$, and $\tilde{w}/\gamma$.

When the sublevels 1 and 2 are close to each other and the intensities of the exciting light in the legs of the $\Lambda$-system are the same, $\omega_{21}\ll\omega_{31}$, $\omega_{32}$ and $J_1=J_2$, the following condition must be met:

$$w_1\gamma_{32}=w_2\gamma_{31}, \qquad (4.3)$$

since the left- and right-hand sides contain the same products of dipole-moment matrix elements. If condition (4.3) is met, the solutions (4.2) assume the following simple form:

$$\rho_{11}=\frac{1}{2}+\frac{\tilde{w}}{2}\left(\frac{1}{w_1}-\frac{1}{w_2}\right),$$
$$\rho_{22}=\frac{1}{2}+\frac{\tilde{w}}{2}\left(\frac{1}{w_2}-\frac{1}{w_1}\right), \qquad (4.4)$$
$$\rho_{33}=\frac{w_1+w_2}{2\gamma}-\frac{\tilde{w}}{2\gamma_{31}}-\frac{\tilde{w}}{2\gamma_{32}}.$$

These solutions show that when the optical-transition dipole moments are distinct, $|d_{13}|\neq|d_{23}|$, and hence $w_1\neq w_2$, the atoms may become redistributed among the sublevels 1 and 2 with probabilities differing from 1/2, even though $J_1=J_2$. This optical pumping effect is determined solely by the quantum properties of parametric radiation and is based on the fact that $\tilde{w}$ is a quantity of the same order as $w_1$ and $w_2$. Below we estimate these quantities together, based on the general properties of field correlation functions listed in Appendix B. We also formulate more carefully the requirements that the source of parametric radiation must meet in the range in which the system of kinetic equations (3.2) is valid.

Let us start by estimating $\tilde{w}$, which in accordance with (4.1), (3.6), and (3.8) has the form

$$\tilde{w}=\frac{1}{4}\sigma^0_{13}\sigma^0_{23}(G_{12}(\omega_{21})+G^*_{12}(\omega_{21})), \qquad (4.5)$$

where

$$\sigma^0_{13}=\frac{8\pi\omega_{31}}{c\hbar\gamma}|d_{13}|^2,\quad \sigma^0_{23}=\frac{8\pi\omega_{32}}{c\hbar\gamma}|d_{23}|^2 \qquad (4.6)$$

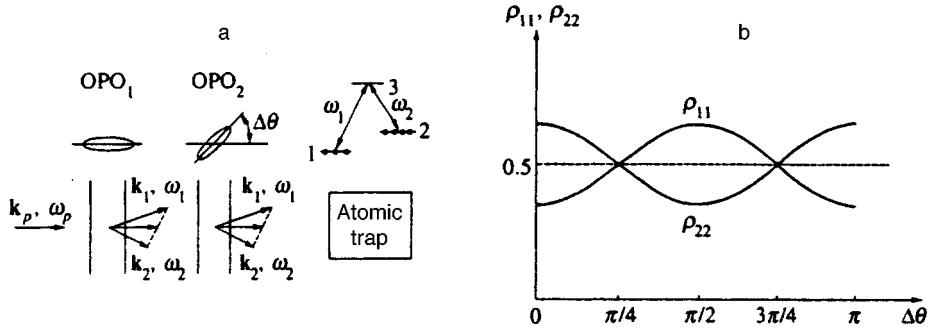



FIG. 2. (a) Schematic of the proposed experiment in optically pumping three-level atoms by radiation emitted by two optical parametric oscillators, $OPO_1$ and $OPO_2$, characterized by different time ordering in the creation of photons in the two associated modes $\mathbf{k}_1\omega_1$ and $\mathbf{k}_2\omega_2$. The precooled atoms are kept in an atomic trap. Photons of the associated modes are created as a result of the decay of pump photons $\mathbf{k}_p\omega_p$ in the four-wave mixing process. The difference in orientation of the squeezing ellipses is determined by the relative phase difference $\Delta\theta=\theta_1-\theta_2$ of the anomalous correlation functions. (b) Distribution of populations in the system of sublevels 1 ($\rho_{11}$) and 2 ($\rho_{22}$) of the $\Lambda$-system driven by radiation emitted by $OPO_1$ and $OPO_2$, as a function of the relative phase difference $\Delta\theta$. The distribution forms in the limit of low intensity of the optical pumping radiation.

are the photon resonant absorption cross sections. We examine the spectral function $G_{12}(\Omega)$, which determines the rate parameter $\tilde{w}$. The causality principle imposes an important restriction on the type of anomalous correlation functions entering into the definition (3.7). Let us assume that the photons created in a parametric process are ordered in time: the photon in mode $\omega_1$ is created either earlier or later than the photon in mode $\omega_2$. This means that the anomalous correlation functions $\Phi_{21}^{(--)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$, which are defined in (2.18) and refer to a physically specified process of parametric oscillation, will vanish only for $t_1<t_2$ or only for $t_1>t_2$. Correspondingly, the functions $\Phi_{12}^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ will vanish in the opposite cases. In this way the product of these functions, which enters into the definition of $G_{12}(\Omega)$ taken for the same parametric scatterer, will vanish. But if these correlation functions are taken for two different scatterers with different time ordering in the creation of photons in modes $\omega_1$ and $\omega_2$, the integrands in (3.7) are finite. For definiteness and generality of method, we assume that the medium is irradiated by light from two optical parametric oscillators characterized by different time ordering in the generation of the $\omega_1$ and $\omega_2$ modes and connected in series, as depicted in Fig. 2a. The two parametric oscillators are driven by the same pump radiation and are assumed to be transparent for the generated squeezed radiation. The case where there is no time ordering in the creation of photons in modes $\omega_1$ and $\omega_2$ can obviously be considered a particular one if we assume that the radiation intensities and the phases of the anomalous correlation functions from both sources are the same.

Using the Fourier transform (B12) for the correlation functions, we can write their frequency components $\Phi_{21}^{(--)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ and $\Phi_{12}^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ defined in (2.18) in the form of the following Fourier decompositions:

$$\Phi_{21}^{(--)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\exp\left[-i\omega_p\left(t_1-\frac{z_1}{c}\right)-i\omega_p\left(t_2-\frac{z_2}{c}\right)\right]\times\int\frac{d^3\kappa'}{(2\pi)^3}\int_{-\infty}^{0}\frac{d\Omega'}{2\pi}\exp[-i\Omega'\times(t_1-t_2)+i\boldsymbol{\kappa}'(\mathbf{r}_1-\mathbf{r}_2)]\times\Phi^{(--)}(\boldsymbol{\kappa}',\Omega'),$$
$$\Phi_{12}^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\exp\left[i\omega_p\left(t_1-\frac{z_1}{c}\right)+i\omega_p\left(t_2-\frac{z_2}{c}\right)\right]\times\int\frac{d^3\kappa''}{(2\pi)^3}\int_{-\infty}^{0}\frac{d\Omega''}{2\pi}\exp[-i\Omega''\times(t_1-t_2)+i\boldsymbol{\kappa}''(\mathbf{r}_1-\mathbf{r}_2)]\times\Phi^{(++)}(\boldsymbol{\kappa}'',\Omega''),\tag{4.7}$$

where $\omega_p$ is the frequency of the pump light in the parametric process. According to the initial assumptions, the Fourier transforms of the correlation functions have resonant singularities near the frequencies $\Omega'$, $\Omega''\sim\Omega/2$. It has proved convenient to write the Fourier transforms as follows:

$$\Phi^{(--)}(\boldsymbol{\kappa}',\Omega')=(2\pi)^2\hbar\omega_p\delta(c\kappa_z'-\Omega')(e^{2i\theta_1(\boldsymbol{\kappa}_\perp',\Omega')}\times r_1(\boldsymbol{\kappa}_\perp',\Omega')+e^{2i\theta_2(\boldsymbol{\kappa}_\perp',\Omega')}r_2(\boldsymbol{\kappa}_\perp',\Omega')),$$
$$\Phi^{(++)}(\boldsymbol{\kappa}'',\Omega'')=(2\pi)^2\hbar\omega_p\delta(c\kappa_z''-\Omega'')(e^{-2i\theta_1(\boldsymbol{\kappa}_\perp'',\Omega'')}\times r_1(\boldsymbol{\kappa}_\perp'',\Omega'')+e^{-2i\theta_2(\boldsymbol{\kappa}_\perp'',\Omega'')}r_2(\boldsymbol{\kappa}_\perp'',\Omega'')),\tag{4.8}$$

where the terms in parentheses correspond to the contributions of the first and second parametric oscillators. In accordance with the causality principle, each of these terms must be an analytic function of $\Omega'$ and $\Omega''$ in the upper or lower half-plane. We assume that the phases $\theta_1$ and $\theta_2$ are distinct and begin to depend on their arguments only near the resonances.

Plugging (4.7) and (4.8) into (3.7), we get

$$G_{12}(\Omega)+G_{12}^*(\Omega)=\int\int\frac{d^2\kappa_\perp'}{(2\pi)^2}\frac{d^2\kappa_\perp''}{(2\pi)^2}\times\int_{-\infty}^{0}\int_{-\infty}^{0}\frac{d\Omega'}{2\pi}\frac{d\Omega''}{2\pi}2\pi\delta(\omega_{21}+\Omega'$$

$$+\Omega'')\{e^{2i(\theta_1-\theta_2)}r_1(\boldsymbol{\kappa}_\perp',\Omega')r_2(\boldsymbol{\kappa}_\perp'',\Omega'')$$
$$+e^{-2i(\theta_1-\theta_2)}r_2(\boldsymbol{\kappa}_\perp',\Omega')r_1(\boldsymbol{\kappa}_\perp'',\Omega'')\}. \tag{4.9}$$

We assume that near the resonances $\Omega'$, $\Omega''\sim\Omega/2$ the phases $\theta_1$ and $\theta_2$ vary in a coordinated manner, so that their difference remains constant. The presence of a delta function in the integrand of (4.9) implies that this expression can be finite only if the transition frequency $\omega_{21}$ coincides with $\Omega$ within a scale of order $\tau_c^{-1}$ that determines the width of the spectrum of the integrands. Integration with respect to $\kappa_\perp'$ or $\kappa_\perp''$ is limited to a scale $S_{\rm coh}^{-1}$. The quantity $\kappa_{\rm coh}^{-1}\sim S_{\rm coh}^{-1}$ is the limiting spatial scale of coherence for a parametric oscillator that allows for coordination between the phases of the radiation scattered by different volumes of the medium. The given scale can also be interpreted as the limiting spatial scale suitable for low-noise measurements involving squeezed radiation.[30] The final estimate of the rate parameter $\tilde{w}$, defined in (4.5), is

$$\tilde{w}\sim\frac{1}{2}\cos 2(\theta_1-\theta_2)\,\frac{\sigma_{13}^0\sigma_{23}^0}{S_{\rm coh}^2}\,\frac{1}{\tau_c}\langle r_1r_2\rangle, \tag{4.10}$$

where we estimated the integral (4.9) by using the average values of the integrands.

The estimates of the transition rates $w_1$ and $w_2$ can be done by reasoning along similar lines. Using the definitions of Appendix B, we can write the photon flux densities $J_1$ and $J_2$ as

$$J_1=\int\frac{d^2\kappa_\perp'}{(2\pi)^2}\int_{-\infty}^{0}\frac{d\Omega'}{2\pi}\{\lambda_1(\boldsymbol{\kappa}_\perp',\Omega')+\lambda_2(\boldsymbol{\kappa}_\perp',\Omega')\}, \tag{4.11}$$
$$J_2=\int\frac{d^2\kappa_\perp'}{(2\pi)^2}\int_{0}^{\infty}\frac{d\Omega'}{2\pi}\{\lambda_1(\boldsymbol{\kappa}_\perp',\Omega')+\lambda_2(\boldsymbol{\kappa}_\perp',\Omega')\},$$

where $\lambda_1(\boldsymbol{\kappa}_\perp',\Omega')$ and $\lambda_2(\boldsymbol{\kappa}_\perp',\Omega')$ are the degeneracy parameters that refer to the first and second oscillators, respectively. Assuming that the condition (B15) holds, we can estimate the transition rates $w_1$ and $w_2$ as follows:

$$w_1\sim\frac{\sigma_{13}^0}{S_{\rm coh}}\,\frac{1}{\tau_c}(\langle r_1^2\rangle+\langle r_2^2\rangle),$$
$$w_2\sim\frac{\sigma_{23}^0}{S_{\rm coh}}\,\frac{1}{\tau_c}(\langle r_1^2\rangle+\langle r_2^2\rangle), \tag{4.12}$$

where we have averaged the integrands in (4.11) under the assumption that the values of averages over $\Omega$ for the positive and negative parts of the spectrum are equal.

Let us now formulate more specifically the requirements that a parametric source must obey in our scheme of optically pumping the $\Lambda$-system. Effective redistribution of atoms among the sublevels 1 and 2 is possible if $w_1$, $w_2\gg\Gamma_{12}$, $\Gamma_{21}\sim\tau_0^{-1}$. Let us estimate the order of magnitude of the resonant cross section of photon absorption in both transitions by the square of the optical wavelength, $\lambda_0^2$, and the coherence area by $S_{\rm coh}\sim\lambda_0 l$, where $l$ is the length of the layer in which parametric oscillation occurs.[30] Under these conditions we find that pumping is effective when

$$\frac{\lambda_0}{l}(\langle r_1^2\rangle+\langle r_2^2\rangle)\,\frac{1}{\tau_c}\gg\frac{1}{\tau_0}. \tag{4.13}$$

Although this condition agrees with the necessary requirement that $\tau_c\ll\tau_0$, it is more stringent. In particular, the parameter $\langle r_1^2\rangle$ and $\langle r_2^2\rangle$ can be interpreted as the average values of the degeneracy parameter for the first and second sources and are small compared to unity. The ratio $\lambda_0/l$ must also be assumed small since we ignore diffraction effects and study radiation propagating along the $z$ axis with small values of the transverse wave vectors $\boldsymbol{\kappa}_\perp$. Note that the condition (4.13) must be met together with the requirement that $\tau_c\gg\gamma^{-1}$. Formally, the two requirements can be met simultaneously if we select atoms with a spin ground state as the $\Lambda$-system. The times of relaxation between the sublevels 1 and 2 in this case may reach values of order 0.1 s for typical values of the lifetimes of the excited states of order $10^{-8}-10^{-9}$ s. The most difficult task here is to select the source of parametric oscillations with the required correlation time.

The estimates (4.10) and (4.12) show that the transition rates $\tilde{w}$, $w_1$, and $w_2$ are indeed of the same order of magnitude in the intensity of the irradiating light. The light intensity enters into these expressions in the form of the average values of the degeneracy parameters in parametric oscillations, $\langle r_1^2\rangle$ and $\langle r_2^2\rangle$, and the product $\langle r_1r_2\rangle$. But the transition rate $\tilde{w}$ contains an additional smallness parameter $\lambda_0/l\ll1$. As noted earlier, this smallness parameter reflects the limited nature of our approach, since formally, remaining within it, we cannot consider parametric oscillations with the limiting value of the coherence area, $S_{\rm coh}\to\lambda_0^2$, and a length $l$ of the parametric medium comparable to $\lambda_0$. In other words, we are forbidden to consider the irradiation of atoms by light gathered from large solid angles. Nevertheless, it is natural to expect, from general physical considerations, that the transition to exceptionally small coherence areas, $S_{\rm coh}\to\lambda_0^2$, must lead to the most optimum conditions for observing the discussed quantum statistical effects in the optical pumping of a three-level atom. In this connection we believe that, despite the additional smallness of $\tilde{w}$ in relation to $w_1$ and $w_2$ present in our calculations, the actual redistribution of atoms among the sublevels 1 and 2 caused by the quantum nature of light may be essential. The above estimates also show that ignoring atomic motion in our calculations has a negative effect on observing optical pumping, since allowing for it must lead to a decrease in the absorption cross section for resonant light in comparison to $\lambda_0^2$ and hence to a relative decrease in $\tilde{w}$ in comparison to $w_1$ and $w_2$. The case of a cold atomic ensemble, which is characterized by an average Doppler line broadening small compared to the natural linewidth $\gamma$, is optimal for observing the optical pumping effect based on the quantum statistical properties of the light source.

Figure 2b depicts the populations of the sublevels 1 and 2 as functions of the relative phase difference $\Delta\theta=\theta_1-\theta_2$. For graphical reasons in Fig. 2a we show that this phase difference can be interpreted as the relative angle indicating the mutual orientation of the squeezing ellipses for the first and second sources. Geometrically, each ellipse characterizes

the range of indeterminancy for the quantum fluctuations of the field amplitude at frequencies close to $\omega_{21}/2$. A special feature of this optical pumping scheme is that the difference in the populations of the sublevels 1 and 2 depends on the relative phase difference $\theta_1-\theta_2$ and may change sign when $2(\theta_1-\theta_2)=\pi/2+\pi m$ $(m=0,\pm 1,\ldots)$. Thus, because of the asymmetry in the behavior of the quantum fluctuations of the field amplitudes in the sources an asymmetry develops in the excitation conditions in the $\Lambda$-system. This sets optical pumping by parametric luminescence light apart from optical pumping by light from a Gaussian spectral source (spontaneous emission of an atomic ensemble, thermal radiation, etc.), in which the distribution of the field fluctuations in the plane of complex-valued amplitudes is totally symmetric.

Our calculations heavily rest on the assumption that the fluctuation spectrum of the squeezed radiation is narrow. The parameters $\Delta\omega$ and $\tau_c^{-1}$, which determine the widths of the spectra of the normal and anomalous correlation functions and are actually quantities of the same order of magnitude, are assumed small compared to the natural width $\gamma$ of the excited level of the $\Lambda$-system. From the standpoint of these effects the situation becomes unfavorable if the opposite conditions hold, i.e., if $\Delta\omega,\ \tau_c^{-1}\geqslant\gamma$. Let us prove this by analyzing the relative order of magnitude of the kinetic coefficients that appear in the equations for the atomic density matrix in this case, bearing in mind that the conditions $\Delta\omega,\ \tau_c^{-1}\geqslant\gamma$ are sufficient for the kinetic equation approximation to be valid. Direct calculations readily show that as the width $\Delta\omega$ of the optical-pumping radiation spectrum increases with the integral intensity remaining constant, the stimulated transition rates $w_1$ and $w_2$ decrease in proportion to the ratio $\gamma/\Delta\omega$. The calculation of the diagrams (2.15), (2.16), and (2.19) and their transformation into the corresponding terms in the kinetic equation generates terms of the depopulation and repopulation type, similar to those that were considered earlier and have the same characteristic set of rate parameters, $\tilde{w}_1$, $\tilde{w}_2$, $\tilde{\gamma}_{12}$, and $\tilde{\gamma}_{21}$. But under the conditions we are studying here, when the width $\tau_c^{-1}$ of the anomalous correlation spectrum increases, these rate parameters decrease in proportion to the product $\gamma^2\tau_c^2$. Since $\Delta\omega\sim\tau_c^{-1}$, it is clear that the inequalities $\Delta\omega,\tau_c^{-1}\geqslant\gamma$ lead to the following inequalities between the kinetic coefficients: $w_1,\ w_2\geqslant\tilde{w}_1,\ \tilde{w}_2,\ \tilde{\gamma}_{12},\ \tilde{\gamma}_{21}$. These relationships between the kinetic coefficient are obvious even from an analysis of the general structure of the diagram contributions (2.15), (2.16), and (2.19), which constitute contributions in the perturbation series for the self-energy functions of a higher order than the diagrams (2.8) and (2.9). Indeed, the presence of an additional virtual transition in the diagrams containing anomalous correlations is characterized by additional integration over time with a natural limiting scale $\gamma^{-1}$. As a result, the corresponding correction to the self-energy functions decreases substantially because of the disappearance of correlations on this time scale.

The above reasoning implies, among other things, that the above optical pumping effect in a $\Lambda$-system is truly related to the finiteness and, more than that, to the narrowness of the fluctuation spectrum of squeezed radiation. When the radiation intensity is low, the optical pumping effect and the additional repopulation and depopulation terms in the kinetic equation must vanish when we go over to broadband squeezed radiation.

## 5. DISCUSSION

We have derived quantum kinetic equations describing the evolution of the components of the density matrix for a three-level $\Lambda$-system excited by radiation whose statistics is nonclassical. The main feature of these equations is the presence of terms reflecting the presence of anomalous correlation characteristics in the optical pumping radiation. Since the phases of the field anomalous correlation functions generated by different sources may differ, the kinetic equations acquire terms that reflect the specifically quantum interference effect when the atom is irradiated by parametric luminescence light generated by different sources. Note that this type of interference difference from the ordinary type, caused by the phase difference in the product of complex-valued amplitudes of the fields from different sources. In the case at hand the average field amplitudes are zero. And the interference of anomalous correlation functions gives rise, in particular, to an optical pumping effect, i.e., a redistribution of atoms among the sublevels of the $\Lambda$-system.

An important feature of the system of kinetic equations (3.2) is that the additional terms in these equations reflecting the quantum statistics of the radiation become important when $\omega_1-\omega_2\approx\omega_{21}$. This condition for resonance must be valid to within the spectral width $\tau_c^{-1}$ of the field correlation functions. The reader will recall that this quantity was assumed to be considerably smaller than the natural atomic linewidth $\gamma$. This fact relates the effects described by Eqs. (3.2) to the effect of coherent population trapping based on the Bloch equations.[5–8] For instance, the last equation in (4.4) implies that there is a possibility of reduction of the population of the excited sublevel 3 when the conditions for resonance are met and when the phases of the first and second parametric oscillators are chosen accordingly. Note, however, that in the case of coherent trapping the absence of atoms on the sublevel 3 is accompanied by their simultaneous trapping in a state that is a coherent superposition of states 1 and 2. No such coherent superposition of states emerges in our case. Apparently, we can say that the derived optical pumping equations, being actually balance equations, nevertheless exhibit some features of coherent Bloch equations. The possibility, which follows from the solutions (4.4), of the atoms being redistributed among the states 1 and 2 leads to the following question: can squeezed light redistribute (at least in principle) the atoms in a closed $\Lambda$-system among states 1 and 2 and at the same time lower the population (or leave it unchanged) of state 3 even if no superpositions between states 1 and 2 are formed (as happens in coherent trapping)? The above estimates show that if such an effect is possible, the atoms must be irradiated by spatially multimode squeezed radiation collected from all directions. However, an exact answer to this question lies outside the scope of our model and requires further investigation.

In conclusion we note that the assumption about the inequality of the dipole moments of the transitions, $d_{13}$

$\neq d_{23}$, which we made in describing the optical pumping effect, is not too important for observing the quantum statistical features in the system of equations (3.2). The fact that the dipole moments are equal only means that under $\Lambda$-resonance conditions the difference in the populations of sublevels 1 and 2 must be zero owing to the total symmetry of the problem, which under ordinary conditions agrees with the effects of other relaxation processes on the atoms. But when there is an additional (say, collisional) mechanism leading to the formation of a finite difference in populations between these states, the quantum statistical effects manifests themselves in the solution of the system of kinetic equations.

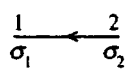
The authors would like to express their gratitude to the participants of the St. Petersburg Quantum Optics Workshop for fruitful discussions concerning the results of the present work. The research was made possible by the financial support of the Russian Fund for Fundamental Research (Grant No. 96-02-16613) and the U.S. Civilian Research and Development Foundation (Grant No. RP1-263 Prop. #2794).

## APPENDIX A: ATOMIC GREEN'S FUNCTIONS AND THEIR RELATION TO THE WIGNER DENSITY MATRIX

The unperturbed atomic Green's functions are defined as

$$G^{(0)}_{n_1n_2}{}^{(\sigma_1\sigma_2)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2) = -i\langle T_{\sigma_1\sigma_2}(\Psi^0_{n_1}(\mathbf{r}_1t_1)\Psi^{0\dagger}_{n_2}(\mathbf{r}_2t_2))\rangle, \tag{A1}$$

where the ordering operators $T_{\sigma_1,\sigma_2}$ for the possible values $\sigma_1$, $\sigma_2=\pm$ act according to the following rules: $T_{--}=T$ is the time-ordering operator, $T_{++}=\tilde{T}$ is the antiordering operator, $T_{+-}$ is the identity operator, and $T_{-+}$ is the permutation parameter. In the case of Fermi statistics, the expression (A1) must be multiplied by $-1$ if the ordering parameter $T_{\sigma_1\sigma_2}$ leads to a permutation of the $\Psi$-operators. In diagrams such functions are depicted by light lines:




$$= iG^{(0)}_{n_1n_2}{}^{(\sigma_1\sigma_2)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2). \tag{A2}$$

Clearly, if we consider, say, the decomposition of the evolution operator in (2.6) and if the atomic subsystem obeys Gaussian statistics, the average of the product of an arbitrary number of $\Psi$-operators is reduced to the product of all possible pair averages of the form (A1).

The total atomic Green's functions, which are averages of products of operators in the Heisenberg representation,

$$G^{(\sigma_1\sigma_2)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2) = -i\langle T_{\sigma_1\sigma_2}(\Psi_{n_1}(\mathbf{r}_1t_1)\Psi^{\dagger}_{n_2}(\mathbf{r}_2t_2))\rangle, \tag{A3}$$

and emerge as a result of ''dressing'' the functions (A1) with various interactions, are depicted by heavy lines in diagrams. These functions considered on small space–time scales of the difference $\mathbf{r}_1-\mathbf{r}_2$ and $t_1-t_2$ are related to the atomic Wigner density matrix as follows:

$$iG^{(--)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2) = \int \frac{d^3p}{(2\pi\hbar)^3}\exp\left[\frac{i}{\hbar}\mathbf{p}(\mathbf{r}_1-\mathbf{r}_2)-\frac{i}{\hbar}\epsilon(p)\times(t_1-t_2)-\frac{i}{2\hbar}(\epsilon_{n_1}+\epsilon_{n_2})(t_1-t_2)-\frac{1}{2}(\gamma_{n_1}+\gamma_{n_2})|t_1-t_2|\right]\times\left\{\delta_{n_1n_2}\theta(t_1-t_2)\pm\rho_{n_1n_2}\left(\mathbf{p},\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\frac{t_1+t_2}{2}\right)\right\}, \tag{A4}$$

$$iG^{(++)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2) = [iG^{(--)}_{n_1n_2}(\mathbf{r}_2,t_2;\mathbf{r}_1,t_1)]^*, \tag{A5}$$

$$iG^{(+-)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2) = \int \frac{d^3p}{(2\pi\hbar)^3}\exp\left[\frac{i}{\hbar}\mathbf{p}(\mathbf{r}_1-\mathbf{r}_2)-\frac{i}{\hbar}\epsilon(p)\times(t_1-t_2)-\frac{i}{2\hbar}(\epsilon_{n_1}+\epsilon_{n_2})(t_1-t_2)-\frac{1}{2}(\gamma_{n_1}+\gamma_{n_2})|t_1-t_2|\right]\left\{\delta_{n_1n_2}\pm\rho_{n_1n_2}\left(\mathbf{p},\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\frac{t_1+t_2}{2}\right)\right\}, \tag{A6}$$

$$iG^{(-+)}_{n_1n_2}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2) = \pm\int \frac{d^3p}{(2\pi\hbar)^3}\exp\left[\frac{i}{\hbar}\mathbf{p}(\mathbf{r}_1-\mathbf{r}_2)-\frac{i}{\hbar}\epsilon(p)(t_1-t_2)-\frac{i}{2\hbar}(\epsilon_{n_1}+\epsilon_{n_2})(t_1-t_2)-\frac{1}{2}(\gamma_{n_1}+\gamma_{n_2})|t_1-t_2|\right]\times\rho_{n_1n_2}\left(\mathbf{p},\frac{\mathbf{r}_1+\mathbf{r}_2}{2},\frac{t_1+t_2}{2}\right). \tag{A7}$$

Here $\epsilon(p)=p^2/2m$, $\epsilon_{n_1}$ and $\epsilon_{n_2}$ are internal energies of the atoms in the states $n_1$ and $n_2$, $\gamma_{n_1}$ and $\gamma_{n_2}$ are spontaneous decay rates for these state (which may be absent for the lower energy sublevels), and $\theta(\tau)$ is the Heaviside step function, equal to 1 for $\tau>0$ and to 0 for $\tau<0$. The upper sign in these equations corresponds to Bose statistics and the lower sign to Fermi statistics. Note that Eqs. (A4)–(A7), which play an important role in deriving the kinetic equation, are valid only if the evolution of the Green's functions over the space–time arguments can be considered free and the times $t_1$ and $t_2$ are close. They cease to be valid if the time difference $t_1-t_2$ becomes comparable to the characteristic transitions time in the sublevel system.

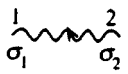
## APPENDIX B: GREEN'S FUNCTIONS AND THE ELECTROMAGNETIC-FIELD CORRELATION FUNCTIONS

In quantum electrodynamics, photon Green's functions are expressed in terms of averages of products of vector potential operators.[28,31] However, in the problem under consideration, in which there is dipole interaction between the elec-

tromagnetic field and the model sublevel system and the vector properties of the field are not specified explicitly, it is advisable to introduce Green's functions directly for the field-strength operators. Three types of these functions can be defined when the statistics is quasi-Gaussian. First, the normal Green's functions existing both in the electrodynamics of the vacuum and in the case of a Gaussian spectral source:

$$\overset{(0)}{F}{}^{(\sigma_1\sigma_2)}(\mathbf{r}_1t_1,\mathbf{r}_2t_2)=-i\langle T_{\sigma_1\sigma_2}(E^{(-)}(\mathbf{r}_1t_1)E^{(+)}(\mathbf{r}_2t_2))\rangle, \tag{B1}$$

which are the averages of time-ordered products of operators of different frequencies. These functions are depicted in diagrams by wave lines with a unidirectional arrow:




$$= i\overset{(0)}{F}{}^{(\sigma_1\sigma_2)}(\mathbf{r}_1t_1,\mathbf{r}_2t_2). \tag{B2}$$

But there is also another pair of anomalous Green's functions,

$$\overset{(-)}{F}{}^{(\sigma_1\sigma_2)}(\mathbf{r}_1t_1,\mathbf{r}_2t_2)=-i\langle T_{\sigma_1\sigma_2}(E^{(-)}(\mathbf{r}_1t_1)E^{(-)}(\mathbf{r}_2t_2))\rangle \tag{B3}$$

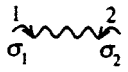
and

$$\overset{(+)}{F}{}^{(\sigma_1\sigma_2)}(\mathbf{r}_1t_1,\mathbf{r}_2t_2)=-i\langle T_{\sigma_1\sigma_2}(E^{(+)}(\mathbf{r}_1t_1)E^{(+)}\times(\mathbf{r}_2t_2))\rangle, \tag{B4}$$

which are the averages of time-ordered products of operators of the same frequency. These functions are depicted in diagrams by wavy lines with arrows pointing in opposite directions:


$$= i\overset{(-)}{F}{}^{(\sigma_1\sigma_2)}(\mathbf{r}_1t_1,\mathbf{r}_2t_2). \tag{B5}$$

and




$$= i\overset{(+)}{F}{}^{(\sigma_1\sigma_2)}(\mathbf{r}_1t_1,\mathbf{r}_2t_2). \tag{B6}$$

Note that the anomalous Green's functions (B3) and (B4) for light emitted by a parametric spectral source exist for zero average values of field strength.

A simple analysis shows that because the commutators of free fields are $c$-number functions, irrespective of the type of time-ordering in (B1), (B3), and (B4) the self-correlation properties of a quasi-Gaussian spectral source are completely determined by four correlation functions,

$$\Phi^{(+-)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\langle E^{(+)}(\mathbf{r}_1t_1)E^{(-)}(\mathbf{r}_2t_2)\rangle, \tag{B7}$$

$$\Phi^{(-+)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\langle E^{(+)}(\mathbf{r}_2t_2)E^{(-)}(\mathbf{r}_1t_1)\rangle, \tag{B8}$$

$$\Phi^{(--)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\langle E^{(-)}(\mathbf{r}_1t_1)E^{(-)}(\mathbf{r}_2t_2)\rangle, \tag{B9}$$

$$\Phi^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\langle E^{(+)}(\mathbf{r}_1t_1)E^{(+)}(\mathbf{r}_2t_2)\rangle, \tag{B10}$$

which obey the following symmetry relations:

$$\begin{aligned}&\Phi^{(-+)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=\Phi^{(+-)}(\mathbf{r}_2,t_2;\mathbf{r}_1,t_1),\\&\Phi^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)=[\Phi^{(--)}(\mathbf{r}_2,t_2;\mathbf{r}_1,t_1)]^*.\end{aligned} \tag{B11}$$

We also note that for a free field, because of the commutativity of the operators of the same frequency, the type of time ordering in the functions $\Phi^{(--)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ and $\Phi^{(++)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2)$ may be assumed arbitrary.

For light emitted by a parametric source, e.g., in a four-wave process, as a result of the monochromatic pump wave with a frequency $\omega_p$ decaying into two phase-associated modes with frequencies $\omega_p\pm\Omega$, the correlation functions (B7)–(B10) depend, after the rapidly oscillating phase factors have been isolated, only on the difference coordinates $\tau=t_1-t_2$ and $\boldsymbol{\rho}=\mathbf{r}_1-\mathbf{r}_2$. Assuming that the pump wave propagates along the $z$ axis, we do a Fourier decomposition of these functions in the difference space–time coordinates:

$$\begin{aligned}&\Phi^{(\sigma_1\sigma_2)}(\boldsymbol{\kappa},\Omega)\\&\quad=\int_{-\infty}^{\infty}d\tau\int d^3\rho e^{i\Omega\tau-i\boldsymbol{\kappa}\boldsymbol{\rho}}\\&\quad\times e^{-i\sigma_1\omega_p(t_1-z_1/c)-i\sigma_2\omega_p(t_2-z_2/c)}\\&\quad\times\Phi^{(\sigma_1\sigma_2)}(\mathbf{r}_1,t_1;\mathbf{r}_2,t_2).\end{aligned} \tag{B12}$$

The symmetry relations (B11) are transformed as follows:

$$\begin{aligned}&\Phi^{(-+)}(\boldsymbol{\kappa},\Omega)=\Phi^{(+-)}(-\boldsymbol{\kappa},-\Omega),\\&\Phi^{(++)}(\boldsymbol{\kappa},\Omega)=[\Phi^{(--)}(\boldsymbol{\kappa},\Omega)]^*.\end{aligned} \tag{B13}$$

This shows that only two functions are independent, and for these we take $\Phi^{(+-)}(\boldsymbol{\kappa},\Omega)$ and $\Phi^{(--)}(\boldsymbol{\kappa},\Omega)$.

For quasimonochromatic radiation propagating along the $z$ axis, the given pair of correlation functions can be represented as follows:

$$\begin{aligned}&\Phi^{(+-)}(\boldsymbol{\kappa},\Omega)=(2\pi)^2\hbar\omega_p\delta(c\kappa_z-\Omega)\lambda(\boldsymbol{\kappa}_\perp,\Omega),\\&\Phi^{(--)}(\boldsymbol{\kappa},\Omega)=(2\pi)^2\hbar\omega_p\delta(c\kappa_z-\Omega)\\&\qquad\times e^{2i\theta(\boldsymbol{\kappa}_\perp,\Omega)}r(\boldsymbol{\kappa}_\perp,\Omega),\end{aligned} \tag{B14}$$

where we ignored diffraction effects and allowed for the fact that the phase of the anomalous correlation function $\theta$ may, generally speaking, also depend on the transverse wave vector $\boldsymbol{\kappa}_\perp$ and the frequency $\Omega$: $\theta=\theta(\boldsymbol{\kappa}_\perp,\Omega)$. What is important here is the assumption that the spectrum of these functions contains two resonance singularities near the frequencies $\Omega\sim\pm\Omega_1$, where $\Omega_1=\omega_1-\omega_p=\omega_p-\omega_2$. This guarantees that the initial assumptions concerning the excitation of the $\Lambda$-system, formulated in Sec. 1 and depicted Fig. 1, are valid.

When the source of parametric radiation is weak (i.e., the probability of a pump photon decaying into a pair of phase-associated photons is low), the radiation intensity is characterized by a dimensionless parameter $\lambda(\boldsymbol{\kappa}_\perp,\Omega)$ (the degeneracy parameter), which is small: $\lambda(\boldsymbol{\kappa}_\perp,\Omega)\ll1$. The appearance of anomalous correlations in the emitted light is also characterized here by a small dimensionless parameter $r(\boldsymbol{\kappa}_\perp,\Omega)\ll1$. In the case of parametric conversion close to

ideal (pure parametric scattering with no Raman scattering or resonance luminescence), the two parameters are linked by the following approximate relationship:

$$\lambda(\boldsymbol{\kappa}_\perp,\Omega)\approx r^2(\boldsymbol{\kappa}_\perp,\Omega)\ll r(\boldsymbol{\kappa}_\perp,\Omega). \tag{B15}$$

This can never happen if the averaging operation in the light correlation functions is examined from the classical viewpoint. The inequality is a reflection of the purely quantum nature of spontaneous parametric luminescence, related to the rigorous time correlation in the creation of phase-associated ''twin'' photons. The condition (B15) implies, among other things, that leaving first-order contributions in $\Phi^{(+-)}(\cdots)$ in calculations of processes by perturbation techniques means leaving second-order contributions in $\Phi^{(--)}(\cdots)$ and $\Phi^{(++)}(\cdots)$. Note that from the viewpoint of the squeezing effect the light we are considering here is an example of radiation characterizes by a low degree of squeezing.


[1] A. Kastler, J. Phys. Radium **11**, 255 (1950).
[2] W. Happer, Rev. Mod. Phys. **44**, 169 (1972).
[3] A. Omont, Prog. Quantum Electron. **5**, 69 (1977).
[4] E. B. Aleksandrov, G. I. Khvostenko, and M. P. Chaĭka, *Interference of Atomic States*, Springer, New York (1993).
[5] E. Arimondo and G. Orriols, Nuovo Cimento Lett. **17**, 333 (1976).
[6] G. Alzetta, A. Gozzini, L. Moi, and G. Orriols, Nuovo Cimento B **36**, 5 (1976).
[7] G. Orriols, Nuovo Cimento B **53**, 1 (1979).
[8] B. D. Agap'ev, M. B. Gornyĭ, B. G. Matisov, and Yu. V. Rozhdestvenskiĭ, Usp. Fiz. Nauk **163**, No. 9, 1 (1993) [ Phys. Usp. **36**, 763 (1993)].
[9] M. Xiao, L.-A. Wu, and H. J. Kimble, Phys. Rev. Lett. **59**, 278 (1987).
[10] E. S. Polzik, J. Carri, and H. J. Kimble, Phys. Rev. Lett. **68**, 3020 (1992).
[11] N. Ph. Georgiades, E. S. Polzik, K. Edamatsu, H. J. Kimble, and A. S. Parkins, Phys. Rev. Lett. **75**, 3426 (1995).
[12] C. W. Gardiner, Phys. Rev. Lett. **56**, 1917 (1986).
[13] H. J. Carmichael, A. S. Lane, and D. F. Walls, Phys. Rev. Lett. **58**, 2539 (1987).
[14] J. Gea-Banacloche, Phys. Rev. Lett. **62**, 1603 (1989).
[15] J. Javanainen and P. L. Gould, Phys. Rev. **41**, 5088 (1990).
[16] A. S. Parkins, in *Modern Nonlinear Optics*, Part 2, M. Evans and S. Kielich (eds.), Wiley, New York (1993), p. 607.
[17] Z. Ficek and P. D. Drummond, Phys. Rev. A **43**, 6247 (1991); **43**, 6258 (1991).
[18] G. Yu. Kryuchkyan, Zh. Éksp. Teor. Fiz. **109**, 116 (1996) [JETP **82**, 60 (1996)].
[19] M. R. Ferguson, Z. Ficek, and B. J. Dalton, Phys. Rev. A **56**, 2379 (1996).
[20] A. S. Parkins, Phys. Rev. A **42**, 4352 (1990); **42**, 6873 (1990).
[21] J. I. Cirac and L. L. Sanchez-Soto, Phys. Rev. A **44**, 1948 (1991).
[22] H. Ritsch and P. Zoller, Phys. Rev. A **38**, 4657 (1988).
[23] A. S. Parkins, Phys. Rev. A **53**, 2893 (1996).
[24] C. W. Gardiner and A. S. Parkins, Phys. Rev. A **50**, 1792 (1994).
[25] D. N. Klyshko, *Photons and Nonlinear Optics*, Gordon & Breach, New York, 1988.
[26] L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1965)].
[27] E. M. Lifshitz and Pitaevskiĭ, *Physical Kinetics*, Pergamon Press, Oxford (1981).
[28] N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields*, Wiley, New York (1979).
[29] A. I. Okunevich, Opt. Spektrosk. **50**, 443 (1981) [Opt. Spectrosc. **50**, 239 (1981)].
[30] I. V. Sokolov, Opt. Spektrosk. **73**, 1158 (1992) [Opt. Spectrosc. **73**, 689 (1992)].
[31] V. B. Berestetskiĭ, E. M. Lifshitz, and L. P. Pitaevskiĭ, *Quantum Electrodynamics*, 3rd ed., Pergamon Press, Oxford (1991).


Translated by Eugene Yankovsky

# Nonlocality of relative diffusion

O. V. Tel'kovskaya and K. V. Chukbar

*Kurchatov Institute, 123182 Moscow, Russia*
(Submitted 31 October 1996)
Zh. Éksp. Teor. Fiz. **112**, 163–166 (July 1997)

On the basis of numerical modeling, the hypothesis of the nonlocal character of Richardson relative diffusion is tested and confirmed. © *1997 American Institute of Physics.*
[S1063-7761(97)01307-3]

Interest in the process of relative diffusion, i.e., the divergence in turbulent flow of two initially close-lying particles of a passive impurity, is chiefly associated with the possibility of avoiding here the fundamental difficulties blocking the path of a complete analytic description of turbulent mixing. Actually, since even turbulent flow, as was already pointed out in the classical papers[1,2] (see also Ref. 3), is a regular flow on a scale of the order of the spatial period of the given fluctuation, the dynamics of an impurity in an actual turbulent field with a wide spectrum (the inertial interval) possess the property of memory, which makes it extremely hard to use analytic approaches to the problem.

On the other hand, the kinetics of $T(\mathbf{l},t)$—the probability density that, in a given experiment, two initially close-lying particles are at the ends of vector $\mathbf{l}$ at time $t$—as a consequence of being used to define averaging over different implementations of turbulent motion (over different experiments), loses memory and, conversely, must possess the property of information loss (the mixing process occurs, and entropy increases).

Nevertheless, even this kinetic equation cannot be derived from first principles and has to be introduced into the theory by simply postulating it (which, as B. Russell once noted, has many advantages, coinciding with those inherent to stealing by comparison with honest labor). Usually, according to a tradition going back to Richardson (who himself introduced the very concept of relative diffusion), everyone reduces it to the usual diffusion equation and only argues about how its coefficient depends on the parameters of the problem.[1,2,4]

It was proposed in Ref. 5 that the analytic possibilities be broadened by going to the region of integral equations of the convolution type. (The most obvious physical reason for nonlocality is that the main contribution to the divergence rate of two particles lying in the inertial interval comes from a turbulent harmonic of the same scale as the current distance between them. Reference 6 was probably the first to call attention to the possibility in principle of a nonlocal variant (see also Ref. 7).) Specifically, the following model equation (in dimensionless units) was written for $T$:

$$\frac{\partial T(\mathbf{l},t)}{\partial t}=\frac{\sqrt{3}}{4\pi^2}\,\Gamma(2/3)\Delta\int\frac{T(\mathbf{l}',t)}{|\mathbf{l}-\mathbf{l}'|^{5/3}}\,d^3\mathbf{l}'. \tag{1}$$

The exponent in the kernel of the integral, equal to the Kolmogorov–Obukhov exponent, arises from the necessity of satisfying Richardson's law $\langle l\rangle \propto t^{3/2}$, which describes the variation of the characteristic diameter of a cloud of impurity in a turbulent medium,[1–4] while the complicated numerical coefficient ($\Gamma$ is Euler's gamma function) arises from the unity in the Fourier representation. This paper is devoted to testing this hypothesis, since only experiment (including numerical modeling) can confirm or disprove a theoretical postulate. Let us discuss this feature.

Equation (1), like the classical diffusion equation, is easy to solve because of its locality in Fourier space, and describes a similar *information loss*—the evolution of any initial impurity-distribution profile to a finite-parameter self-similar solution. Its fundamental difference, however, consists in the presence of an exponential tail in this solution as $l\to\infty$, i.e., a divergence of certain exponential moments in the self-similar function $T$, which is a special case of the Levi function.[5,7,8] In practice, this property denotes the exponential smallness of the probability of detecting in a given experiment a cloud with a diameter that substantially exceeds the mean Richardson value, but not at all the exponential character of the falloff of the impurity concentration at the periphery of the given cloud—as already mentioned, Eq. (1) does not strictly describe the process of turbulent mixing, just as one plane projection does not give a complete repre-

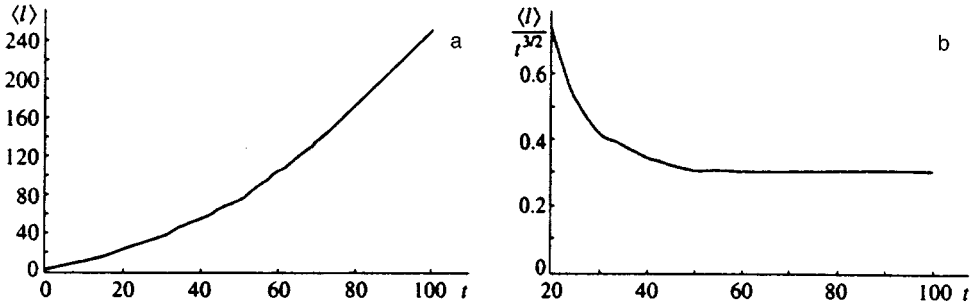



FIG. 1. Realization of Richardson's law: (a) original data, (b) analysis.

 1063-7761/97/070087-02$10.00  

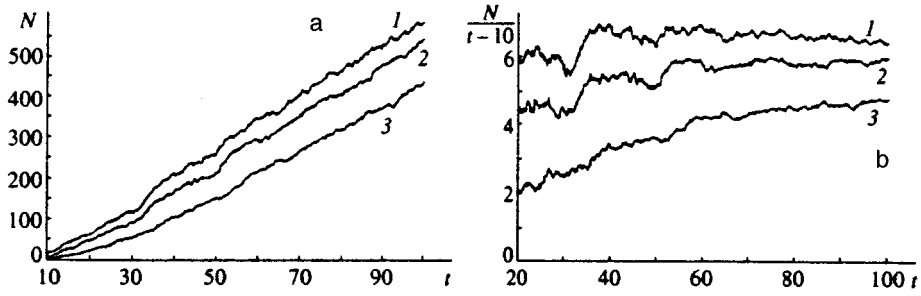



FIG. 2. Number of pairs with an interparticle distance greater than a fixed $l_0$ (*1*—$l_0=80$, *2*—100, *3*—150) vs. time: a) original data, b) analysis.

sentation of a three-dimensional body (and a more appropriate analogy in the given case is one with an infinite-dimensional body); see also Refs. 2–4.

It is especially simple to see the presence of such a tail by rewriting Eq. (1) in the desired limit, taking $|\mathbf{l}-\mathbf{l}'|^{5/3}\simeq l^{5/3}$ out from under the integral sign on the left-hand side and using the normalization condition $\int T d^3\mathbf{l}\equiv 1$:

$$\frac{\partial T}{\partial t}=\frac{\sqrt{3}}{4\pi^2}\,\Gamma\left(\frac{2}{3}\right)\Delta\,\frac{1}{l^{5/3}}\propto\frac{1}{l^{11/3}},\quad l\gg t^{3/2}. \tag{2}$$

One more characteristic property of the relative diffusion process described by Eq. (1) immediately follows from Eq. (2)—the probability of finding a pair of particles separated by a distance significantly greater than the Richardson distance is a linear function of time. It is easy to see that this property, unlike the exponent on $l$, depends neither on the turbulence spectrum nor the dimensionality of the problem, but is associated only with the hypothesis of the nonlocality of the process. It is precisely this, as the most stable property, that is convenient to choose as a verifiable parameter.

Starting from the possibilities of the authors, such a test was made on the basis of numerical modeling. The turbulence was modeled by a given two-dimensional incompressible flow with a quasi-Kolmogorov spectrum. Specifically, the flux function $\Psi(\mathbf{v}=\{v_x,v_y\}=\mathbf{e}_z\times\nabla\Psi)$ was chosen in the form of a sum of 140 harmonics combined into groups of seven each in twenty scale classes; i.e.,

$$\Psi=\sum_{i=1}^{20}\Psi_i,\quad \Psi_i=A_i\sum_{j=1}^{7}\sin(\omega_{ij}t-\mathbf{k}_{ij}\cdot\mathbf{r}+\alpha_{ij}),$$

where the moduli of all the wave vectors in each group are identical, $|\mathbf{k}_{ij}|=k_i$, and their ratio in adjacent classes is $k_i/k_{i+1}=1.4\simeq\sqrt{2}$ $(k_1=1)$. Such a choice made it possible to cover the scale range of the turbulent fluctuations in three orders of magnitude (i.e., $1.4^{20}$). Amplitudes $A_i$, according to the Kolmogorov–Obukhov law, equalled $k_i^{-4/3}$. The angles of rotation of $\mathbf{k}_{ij}$ relative to the $x$ and $y$ axes, as well as phases $\alpha_{ij}$, were chosen randomly in the interval $(0,2\pi)$, and the number of harmonics in each class—seven—was considered sufficient to ensure that the turbulence is isotropic. The frequencies $\omega_{ij}$ were of the order of $k_iv_i$; i.e., they were determined from $\omega_{ij}=\beta_{ij}k_i^{2/3}$, where $\beta_{ij}$ are random quantities in the interval (1/2, 3/2).

The motion of 1004 pairs of points was studied in this given field of velocities, with the initial distance in each pair being equal to 3 (i.e., of the order of a half-wave of the smallest-scale harmonic $\Psi$). To average over the implementations of turbulent flow, the pairs were placed at a distance of 3000 from each other (i.e., of the order of the half-wave of the largest-scale harmonic $\Psi$)—here, by the way, the difference between the process being modeled and the blurring of any fixed cloud is again seen.

The results of numerical modeling are shown in Figs. 1 and 2. The former demonstrates the accuracy of the agreement of Richardson's law for the given model. It can be seen that the self-similar regime is not reached too rapidly—in a time greater than 30, when the particles in pairs diverge on the average by a distance of about 40. The second, key figure reflects the degree of correspondence of the hypothesis concerning the nonlocality of the numerical calculation process (the time shift by 10 in the processing does not contradict the linearity and is evoked by the initial data). The theory of Ref. 5 corresponds to constancy of the functions in Fig. 2b (and their proportionality to $l_0^{-2/3}$) at least until $l_0\gg\langle l\rangle$ (see Fig. 1). The agreement obviously looks rather persuasive. In other words, Eq. (1) gives an extremely good description of the process of relative ''diffusion.''

The authors are grateful to V. V. Yan'kov, discussions with whom stimulated the performance of this work. It was also supported by the Russian Fund for Fundamental Research (Project No. 96-02-17249a) and by the ''Nonlinear Dynamics'' program of the Ministry of Science.

[1] L. F. Richardson, Proc. R. Soc. London, Ser. A **110**, 709 (1926).
[2] G. K. Batchelor, Proc. Cambridge Philos. Soc. **48**, 345 (1952).
[3] A. S. Monin and A. M. Yaglom, *Statistical Hydromechanics*, Nauka, Moscow (1967), part 2, sect. 24.
[4] H. G. E. Hentshel and I. Procaccia, Phys. Rev. A **29**, 1461 (1984).
[5] K. V. Chukbar, JETP Lett. **58**, 90 (1993).
[6] M. F. Shlesinger, B. J. West, and J. Klafter, Phys. Rev. Lett. **58**, 1100 (1987).
[7] J. Klafter, M. F. Schlesinger, and G. Zumofen, Phys. Today **49(2)**, 33 (1996).
[8] E. W. Montroll and M. F. Shlesinger, in *Studies in Statistical Mechanics*, vol. 11, J. Leibowitz and E. W. Montroll (eds.) (North-Holland, Amsterdam 1984), p. 1.

Translated by W. J. Manthey

# Excitonic light-absorption and amplification bands in the presence of laser radiation

S. A. Moskalenko and V. G. Pavlov

*Institute of Applied Physics, Moldavian Academy of Sciences, MD-2028 Kishinev, Republic of Moldova*
(Submitted 17 June 1996)
Zh. Éksp. Teor. Fiz. **112**, 167–179 (July 1997)

We examine the absorption and amplification bands of a weak probe signal in the presence of Bose–Einstein condensation of excitons that emerges in nonequilibrium conditions in the field of coherent laser radiation with a wave vector $\mathbf{k}_0$. We assume that the detuning $\tilde{\Delta}$ from resonance between the energy $\hbar\omega_{ex}(\mathbf{k}_0)+L_0$ of the exciton level, which is shifted because of exciton–exciton interaction, and the laser photon energy $\hbar\omega_L$, is generally nonzero. The elementary excitation spectrum consisting of the quasiexcitonic and quasienergy branches determines the optical properties of the system. When there is real induced Bose–Einstein condensation, at $\tilde{\Delta}=0$ the two branches touch, as they do in spontaneous Bose–Einstein condensation. In virtual induced Bose–Einstein condensation, when $\tilde{\Delta}<0$, instabilities emerge in the spectrum in certain regions of the **k**-space. These instabilities are caused by a real transformation of two laser photons into two extracondensate particles. Nonequilibrium extracondensate excitons strongly affect the absorption and amplification of the probe light signal. We show that light absorption is due to the quantum transition from the ground state of the crystal to the quasiexcitonic branch of the spectrum. On the other hand, amplification of the signal is caused by the transition from the quasienergy branch to the ground state of the crystal. The same transition can be explained by a real transformation of two laser photons into a vacuum photon of frequency $\hbar cq$ and a crystal exciton with a wave vector $2\mathbf{k}_0-\mathbf{q}$. Finally, we show that the excitonic absorption and light-amplification bands are essentially anisotropic at $\tilde{\Delta}\approx 0$ and depend on the orientation of the vectors $\mathbf{q}$ and $\mathbf{k}_0$.  [S1063-7761(97)01407-8]

## 1. INTRODUCTION

The optical Stark effect in the exciton part of the spectrum has been studied both experimentally and theoretically.[1–6] The interpretation of this phenomenon suggested by Schmitt-Rink, Chemla, and Haug[3,4] is based on the idea of an induced Bose–Einstein condensation of excitons being produced by external coherent laser radiation. In contrast to the work of Keldysh and Kozlov,[7] devoted to the spontaneous Bose–Einstein condensation of excitons in the electron–hole setting, here the frequency of the laser radiation acts as the chemical potential. Induced Bose–Einstein condensation can be either real but nonequilibrium,[8] where coherent laser photons excite resonant excitons in the band with the same value as the wave vector, or virtual, where the laser frequency differs considerably from the exciton transition frequency.[3–6] The second variant was realized in the experiments described in Refs. 1 and 2, where the photon energy was much lower than the energy of the lowest exciton level. The experimenters observed a shift in the exciton level after an ultrashort laser pulse was switched on. The level returned to its initial position upon switch-off of the pulse. Theoretical work devoted to this phenomenon[3–6] made it possible to determine a number of features explaining the behavior of excitons in semiconductors in the presence of laser radiation, which induces macroscopic coherent polarization in the crystal. For instance, Schmitt-Rink, Chemla, and Haug[3,4] demonstrated that the filling of the phase space by virtual electrons and holes produced by coherent macroscopic polarization and the exchange electron–hole interaction change the internal electronic structure of the exciton. These phenomena are similar to those occurring in spontaneous Bose–Einstein condensation.[7] As the concentration of electron–hole pairs (or excitons) $n_{ex}$ increases in the low-density limit $n_{ex}a_{ex}^3<1$, where $a_{ex}$ of the exciton Bohr radius, in the Hartree–Fock–Bogolyubov approximation the exciton level shifts into the violet part of the spectrum. However, as shown by Zimmermann,[9] in bulk crystals this shift is balanced by screening and correlation corrections. Hence Bose–Einstein condensation of excitons in bulk crystals has essentially no effect on the position of the exciton level. Nevertheless, the level moves closer to the continuous spectrum of the electron–positron pair and the exciton's binding energy diminishes. The reason is the lowering of the ground-state energy per electron–hole pair in the electron–hole plasma, and an effective decrease in the semiconductor's band gap at concentrations $n_{ex}a_{ex}^3<1$ (see Ref. 9).

Elesin and Kopaev[10] found that the number density of Bose-condensed electron–hole pairs in conditions where $n_{ex}a_{ex}^3<1$ is not a unique function of the laser intensity and the detuning from resonance between the laser frequency and the exciton transition frequency. There are amplitude and frequency hysteresis loops, a situation that indicates the presence of optical bistability.

The effect of an induced Bose condensate of excitons on the energy spectrum of extracondensate quasiparticles was studied in Refs. 5 and 6. In conditions of an optical Stark effect, the elementary excitation spectrum differs substantially from the energy spectrum in the theory of a nonideal

Bose gas, and can be reduced to the latter only in a certain special case.

One important feature of the energy spectrum of extracondensate quasiparticles (excitons, phonons, and vacuum photons) in the presence of laser radiation is its instability. Mathematically, in the system where damping is ignored the energy spectrum becomes complex-valued in certain regions of the **k**-space. One of the two complex conjugate solutions has a positive imaginary part, which means that the amplitude of this solution increases without bound with the passage of time and that the system becomes unstable. Physically the instability is caused by a real transformation of two laser photons into two extracondensate quasiparticles, say, into two excitons or an exciton and a vacuum photon.

These processes are real since they have corresponding conservation laws that the energy and momentum must obey simultaneously, laws that link the initial and final states. The transformations proceed through intermediate virtual states. For instance, two laser photons are transformed into two excitons of the induced Bose condensate, after which the two excitons become virtually transformed into two extracondensate quasiparticles. The real process here is the transformation of two laser photons into two extracondensate quasiparticles.

New waves can build up in the system due to the energy taken from the laser radiation, which for the sake of simplicity is assumed to be given, and inexhaustible. Unlimited buildup of new waves in the system must come as no surprise then. The instabilities can be either convective or absolute. In the first case the build occurs as the wave moves deeper into the medium, with the result that the system operates as an amplifier of waves. In the second the buildup does not leave the region of origin. The wave encompasses an ever growing volume of the crystal and increases with time.

As noted in Ref. 11, the concepts of absolute and convective instability are relative: they depend on the reference frame in which the phenomenon is observed. In the reference frame moving together with the propagating wave the convective instability becomes absolute, and vice versa. What is important is that the occupation numbers of the Bose quasiparticles in the regions of the **k**-space where the instabilities in the elementary excitation spectrum manifest themselves become anomalously large.

In the simplest variant of a system of excitons without damping, the emergence of instabilities is a nonthreshold process. This means that instability appears at arbitrarily low concentration $n_{ex}$ of Bose-condensed excitons. In this simplest variant the occupation numbers of the elementary excitations in the regions of the **k**-space where instabilities appear become infinitely large. Since the process of exciton absorption of light is related to the transition from the ground state of the crystal and creation of a quasiparticle in the final state, its probability is proportional to $1+n_{\mathbf{q}}$, where $n_{\mathbf{q}}$ is the mean occupation number of quasiparticles in the final state. When **q** approaches the instability region in the wave-vector space, this factor tends to infinity, which leads to singularities in the exciton absorption spectrum.

This explains some of the results we arrive at below. The other results are related to the features of the Bose–Einstein distribution function and to the anisotropy resulting from the coherent macroscopic polarization of the medium. In fact, there is damping in the system, so that the onset of instability is a threshold process. Instability can only occur when the number density of coherent excitons, $n_{ex}$, exceeds a threshold value $n_c$ determined by the decay of the exciton levels and the exciton–exciton coupling constant.

The quantum statistical properties of polariton systems near such thresholds of the onset of instabilities was studied by Keldysh and Tikhodeev,[12,13] who used the example of Stokes scattering of coherent polaritons by acoustic phonons, and in Ref. 14 by employing polariton–polariton Raman scattering. Keldysh and Tikhodeev showed that near the threshold of stimulated Brillouin scattering in the wave-vector region where there is Stokes scattering and where instability emerges, the Green's functions describing the scattered polariton and acoustic phonons have singularities of the $1/\lambda$ type, where $\lambda$ is the measure of the deviation of the number density of coherent polaritons from the threshold value:

$$n_{ex}=n_c(1-\lambda), \quad 0<\lambda\ll 1.$$

The occupation numbers of the scattered quasiparticles have the same singularity. This result provides a better understanding of the above conclusion that in the nonthreshold case the mean occupation numbers of elementary excitations with wave vectors lying in the instability region tend to infinity. As shown in Ref. 14, stimulated Raman scattering of coherently excited polaritons has a smeared threshold, which probably smooths out the singularity of the $1/\lambda$ type.

Below we study the probability of absorption and amplification of a weak probe light signal in the transition from the ground state of the crystal to an excitonic state when the crystal is in the field of intense coherent laser radiation, which induces coherent macroscopic polarization in the medium. We are speaking of the probability of a transition in which one more exciton is created in the presence of a large number of real or virtual excitons created in stationary conditions by the coherent laser radiation. The presence of laser radiation and the coherent polarization of the medium are important factors that lead to the possibility of amplifying a weak signal and distinguish the given exciton absorption from the one studied earlier in unexcited crystals.

## 2. HAMILTONIAN AND TRANSITION PROBABILITIES

The Hamiltonian of excitons interacting with each other, with laser radiation, and with vacuum photons can be written in the form:[5,6]

$$H=\sum_{\mathbf{p}} E_{ex}(\mathbf{p})a_{\mathbf{p}}^{+}a_{\mathbf{p}}+\sum_{\mathbf{p}} \hbar cp\mathscr{C}_{\mathbf{p}}^{+}\mathscr{C}_{\mathbf{p}}+\sum_{\mathbf{p}} \lambda_{\mathbf{p}}(a_{\mathbf{p}}^{+}\mathscr{C}_{\mathbf{p}} + a_{\mathbf{p}}\mathscr{C}_{\mathbf{p}}^{+})+\lambda_{\mathbf{k}_0}(a_{\mathbf{k}_0}^{+}\mathscr{C}_{\mathbf{k}_0}+a_{\mathbf{k}_0}\mathscr{C}_{\mathbf{k}_0}^{+}) + \frac{1}{2V}\sum_{\mathbf{p},\mathbf{q},\mathbf{k}} \nu(k)a_{\mathbf{p}}^{+}a_{\mathbf{q}}^{+}a_{\mathbf{q}+\mathbf{k}}a_{\mathbf{p}-\mathbf{k}}, \tag{1}$$

where $E_{ex}$ is the exciton energy; $\hbar cp$ is the photon energy; $a_{\mathbf{p}}^{+}$, $a_{\mathbf{p}}$, $\mathscr{C}_{\mathbf{p}}^{+}$, and $\mathscr{C}_{\mathbf{p}}$ are the creation and annihilation op-

erators for excitons and photons, respectively; $\nu(k)$ and $\lambda_{\mathbf{p}}$ are the exciton–exciton and exciton–photon coupling constants; $\lambda_{\mathbf{k}_0}$ is the exciton–laser coupling constant; and $V$ is the volume occupied by the system.

We assume that the laser radiation is characterized by a wave vector $\mathbf{k}_0$ and photon frequency $\omega_L = ck_0$. Antiresonant interaction terms are ignored. The coherent laser radiation is introduced into (1) by replacing the operators $\mathscr{C}^+_{\mathbf{k}_0}$ and $\mathscr{C}_{\mathbf{k}_0}$ with expressions like

$$C_{\mathbf{k}_0} = \sqrt{F_{\mathbf{k}_0}}\,\exp(-i\omega_L t - i\varphi), \quad F_{\mathbf{k}_0} \sim V. \tag{2}$$

Quantum single-particle states of photons with $\mathbf{p} \neq \mathbf{k}_0$ describe the vacuum electromagnetic field and a weak broadband probe signal. The explicit time dependence in the Hamiltonian (1), which emerges as a result of replacing the operators $\mathscr{C}^+_{\mathbf{k}_0}$ and $\mathscr{C}_{\mathbf{k}_0}$ with (2), can be eliminated by transforming to a system of coordinates rotating with frequency $\omega_L$. This is achieved by introducing the unitary transformation

$$\hat{V} = e^{-i\omega_L t\hat{N}}, \quad \hat{N} = \sum_{\mathbf{p}} (a^+_{\mathbf{p}} a_{\mathbf{p}} + \mathscr{C}^+_{\mathbf{p}} \mathscr{C}_{\mathbf{p}}), \tag{3}$$

and by considering a new Hamiltonian,

$$\mathscr{H} = V^+ H V - \hbar\omega_L \hat{N}, \tag{4}$$

in which free quasiparticles are characterized by an energy spectrum measured from the frequency $\omega_L$:

$$\begin{aligned}\mathscr{H} = &\sum_{\mathbf{p}} \hbar[\omega_{ex}(\mathbf{p}) - \omega_L] a^+_{\mathbf{p}} a_{\mathbf{p}} - \sum_{\mathbf{p}} \hbar(cp - \omega_L)\mathscr{C}^+_{\mathbf{p}} \mathscr{C}_{\mathbf{p}} \\ &+ \sum_{\mathbf{p}} \lambda_{\mathbf{p}} (\mathscr{C}^+_{\mathbf{p}} a_{\mathbf{p}} + a^+_{\mathbf{p}} \mathscr{C}_{\mathbf{p}}) + \sqrt{F_{\mathbf{k}_0}} (a^+_{\mathbf{k}_0} + a_{\mathbf{k}_0}) \\ &+ \frac{1}{2V} \sum_{\mathbf{p},\mathbf{q},\mathbf{k}} \nu(k) a^+_{\mathbf{p}} a^+_{\mathbf{q}} a_{\mathbf{q}+\mathbf{k}} a_{\mathbf{p}-\mathbf{k}}. \end{aligned} \tag{5}$$

The terms that are linear in the operators $a^+_{\mathbf{k}_0}$ and $a_{\mathbf{k}_0}$ can be eliminated by a Bogolyubov shift,[15]

$$a_{\mathbf{p}} = \sqrt{N_{\mathbf{k}_0}} e^{-i\phi} \delta_{\mathbf{p},\mathbf{k}_0} + \alpha_{\mathbf{p}}, \tag{6}$$

where $\alpha_{\mathbf{p}}$ is a small addition. The macroscopic filling factor $N_{\mathbf{k}_0}$ of the exciton mode $\mathbf{k}_0$ is related to $F_{\mathbf{k}_0}$ by[10]

$$\begin{aligned} &n_{\mathbf{k}_0} = \frac{\lambda^2_{\mathbf{k}_0} f_{\mathbf{k}_0}}{\tilde{\Delta}^2 + \gamma^2_{ex}}, \quad n_{\mathbf{k}_0} = \frac{N_{\mathbf{k}_0}}{V}, \quad f_{\mathbf{k}_0} = \frac{F_{\mathbf{k}_0}}{V}, \\ &\tilde{\Delta} = \hbar[\omega_{ex}(\mathbf{k}_0) - \omega_L] + L_0, \quad L_{\mathbf{k}} = \nu(k) n_{\mathbf{k}_0}. \end{aligned} \tag{7}$$

We assume everywhere that repulsion between excitons ($L_{\mathbf{k}} > 0$) is predominant. Here the damping factor $\gamma_{ex}$ has been introduced phenomenologically and enters into the expression that describes optical bistability in the exciton part of the spectrum. After the Hamiltonian (5) has been expanded in the small operators $\alpha^+_{\mathbf{k}_0+\mathbf{k}}$ and $\alpha_{\mathbf{k}_0+\mathbf{k}}$, with $\mathbf{k} \neq 0$, we can separate an additive constant, a quadratic part, and higher-order terms. Here we are interested only in the quadratic part,

$$\begin{aligned}\mathscr{H}^{(2)} = &\sum_{\mathbf{k}} \{\hbar[\omega_{ex}(\mathbf{k}_0 + \mathbf{k}) - \omega_L] + L_0 + L_{\mathbf{k}}\} \alpha^+_{\mathbf{k}_0+\mathbf{k}} \alpha_{\mathbf{k}_0+\mathbf{k}} \\ &+ \frac{1}{2} \sum_{\mathbf{k}} L_{\mathbf{k}} (e^{-2i\phi} \alpha^+_{\mathbf{k}_0+\mathbf{k}} \alpha^+_{\mathbf{k}_0-\mathbf{k}} \\ &+ e^{2i\phi} \alpha_{\mathbf{k}_0+\mathbf{k}} \alpha_{\mathbf{k}_0-\mathbf{k}}) + \sum_{\mathbf{k}} \hbar(c|\mathbf{k}_0 + \mathbf{k}| \\ &- \omega_L) \mathscr{C}^+_{\mathbf{k}_0+\mathbf{k}} \mathscr{C}_{\mathbf{k}_0+\mathbf{k}} + \sum_{\mathbf{k}} \lambda_{\mathbf{k}_0+\mathbf{k}} (\mathscr{C}^+_{\mathbf{k}_0+\mathbf{k}} a_{\mathbf{k}_0+\mathbf{k}} \\ &+ a^+_{\mathbf{k}_0+\mathbf{k}} \mathscr{C}_{\mathbf{k}_0+\mathbf{k}}). \end{aligned} \tag{8}$$

Below we consider vacuum photons to be responsible for quantum transitions and use the procedure adopted in Ref. 6 to diagonalize the exciton part of the Hamiltonian (8). In this case we can assume, without loss of generality, that $\phi = 0$.

In Ref. 6 diagonalization was achieved by introducing operators $\xi^+_{\mathbf{k}}$ and $\xi_{\mathbf{k}}$ via Bogolyubov's unitary transformation:[15]

$$\xi_k = \frac{\alpha_{\mathbf{k}_0+\mathbf{k}} + A_{\mathbf{k}} \alpha^+_{\mathbf{k}_0-\mathbf{k}}}{\sqrt{1 - |A_{\mathbf{k}}|^2}}, \quad \alpha_{\mathbf{k}_0+\mathbf{k}} = \frac{\xi_{\mathbf{k}} - A_{\mathbf{k}} \xi^+_{-\mathbf{k}}}{\sqrt{1 - |A_{\mathbf{k}}|^2}}, \tag{9}$$

where the coefficients $A_{\mathbf{k}}$ depend on the elementary excitation energy:

$$A_{\mathbf{k}} = \frac{\tilde{\Delta} + T_{\mathbf{k}} + L_{\mathbf{k}} - \mathscr{E}(\mathbf{k})}{L_{\mathbf{k}}}. \tag{10}$$

Here

$$\mathscr{E}(\mathbf{k}) = \sqrt{(\tilde{\Delta} + L_{\mathbf{k}} + T_{\mathbf{k}})^2 - L^2_{\mathbf{k}}} \tag{11}$$

is a component of the total elementary-excitation energy $E(\mathbf{k})$,

$$E(\mathbf{k}) = \mathscr{E}(\mathbf{k}) + \hbar \mathbf{V}_s \mathbf{k}, \quad \mathbf{V}_s = \frac{\hbar \mathbf{k}_0}{m_{ex}}, \tag{12}$$

which depends on the velocity $\mathbf{V}_s$ of the induced condensate. The velocity is determined by the momentum $\hbar\mathbf{k}_0$ of a laser photon and the translational exciton mass $m_{ex}$. For the coefficients $A_{\mathbf{k}}$ to meet the condition

$$|A_{\mathbf{k}}| \leqslant 1 \tag{13}$$

in the entire $\mathbf{k}$-space, the square root in (11) must have the same sign as the sum $\tilde{\Delta} + T_{\mathbf{k}} + L_{\mathbf{k}}$. We denote this solution by $\mathscr{E}_1(\mathbf{k})$ and define it according to the following rule:

$$\operatorname{sgn} \mathscr{E}_1(k) = \operatorname{sgn}(\tilde{\Delta} + T_{\mathbf{k}} + L_{\mathbf{k}}). \tag{14}$$

Then the coefficients $A_{\mathbf{k},1}$ meet the condition (13). The energy $E_1(\mathbf{k})$ selected in this manner generally follows the dispersion law of the initial excitonic branch of the spectrum, $\hbar\omega_{ex}(\mathbf{k}_0 + \mathbf{k}) - \hbar\omega_L + L_0 + L_{\mathbf{k}}$, present in the effective Hamiltonian (8). Hence we call the elementary excitations with an energy $E_1(\mathbf{k})$ quasiexcitonic elementary excitations. In addition to $E_1(\mathbf{k})$ there is a second quasienergy branch of the spectrum with an energy $E_2(\mathbf{k})$ determined by the value $\mathscr{E}_2(\mathbf{k}) = -\mathscr{E}_1(\mathbf{k})$ and the property that

$$E_2(\mathbf{k})=\mathscr{E}_2(\mathbf{k})+\hbar\mathbf{V}_s\mathbf{k}=-\mathscr{E}_1(\mathbf{k})+\hbar\mathbf{V}_s\mathbf{k}=-E_1(-\mathbf{k}). \tag{15}$$

This branch generally follows the initial quasienergy branch of the spectrum, $\hbar\omega_L-\hbar\omega_{ex}(\mathbf{k}_0-\mathbf{k})-L_0-L$. The coefficients $A_{\mathbf{k},2}$ can be found by replacing $\mathscr{E}(\mathbf{k})$ with $\mathscr{E}_2(\mathbf{k})$ in (10), and have the properties

$$A_{\mathbf{k},2}A_{\mathbf{k},1}=1, \quad |A_{\mathbf{k},2}|\geqslant 1. \tag{16}$$

Although the two energy levels — the quasiexcitonic level and the quasienergy level — are close, there is only one set of independent operators $\xi_{\mathbf{k}}^+,\xi_{\mathbf{k}}$, specified by (9), with all possible values of $\mathbf{k}$. For such a set we select the operators corresponding to the coefficients $A_{\mathbf{k},1}$ and the elementary-excitation energy $E_1(\mathbf{k})$. To simplify matters we drop the subscript ''1'' on the operators $\xi_{\mathbf{k},1}^+,\xi_{\mathbf{k},1}$ and the coefficients $A_{\mathbf{k},1}$ but retain it on the branch $E_1(\mathbf{k})$.

The complete set of operators $\xi_{\mathbf{k}}^+,\xi_{\mathbf{k}},\xi_{-\mathbf{k}}^+,\xi_{-\mathbf{k}}$ with coefficients $|A_{\mathbf{k}}\leqslant 1|$ and energies $E_1(\mathbf{k})$ proves to be sufficient for describing elementary excitations of both types, and quantum transitions in the system.

This requires, as we will shortly see, allowing for both resonant and antiresonant quantum transitions. After the exciton part of the quadratic Hamiltonian (8) is diagonalized and the new photon operators

$$\eta_{\mathbf{k}}=\mathscr{C}_{\mathbf{k}_0+\mathbf{k}}$$

are introduced, the Hamiltonian $\mathscr{H}^{(2)}$ assumes the form

$$\mathscr{H}^{(2)}=\sum_{\mathbf{k}} E_1(\mathbf{k})\xi_{\mathbf{k}}^+\xi_{\mathbf{k}}+\sum_{\mathbf{k}}\hbar(c|\mathbf{k}_0+\mathbf{k}|-\omega_L)\eta_{\mathbf{k}}^+\eta_{\mathbf{k}}$$
$$+\sum_{\mathbf{k}}\frac{\lambda_{\mathbf{k}_0+\mathbf{k}}}{\sqrt{1-|A_{\mathbf{k}}|^2}}(\xi_{\mathbf{k}}^+\eta_{\mathbf{k}}+\eta_{\mathbf{k}}^+\xi_{\mathbf{k}}$$
$$-A_{\mathbf{k}}^*\xi_{-\mathbf{k}}\eta_{\mathbf{k}}-A_{\mathbf{k}}\xi_{\mathbf{k}}^+\eta_{-\mathbf{k}}^+). \tag{17}$$

The ground state of the system of excitons in a crystal polarized by coherent laser radiation is the vacuum state for the elementary excitation operators $\xi_{\mathbf{k}}$:

$$\xi_{\mathbf{k}}|0\rangle_{ex}=0. \tag{18}$$

Following the results of Ref. 14, we state that the nonequilibrium distribution function of elementary excitations depends on the absolute value of the energy of these excitations, $N(E_1(\mathbf{k}))=N(|E_1(\mathbf{k})|)$. Hence the ground state (18) is stable even if the values of $E_1(\mathbf{k})$ in the rotating system of coordinates are negative. Even in this case, elementary excitations do not appear spontaneously.

However, this state of the crystal is characterized by nonzero occupation numbers of the initial excitonic states:

$$\langle\alpha_{\mathbf{k}_0+\mathbf{k}}^+\alpha_{\mathbf{k}_0+\mathbf{k}}\rangle=n_{\mathbf{k}_0+\mathbf{k}}^{ex}=\frac{|A_{\mathbf{k}}|^2}{|1-|A_{\mathbf{k}}|^2|}, \tag{19}$$

$$\langle\alpha_{\mathbf{k}_0+\mathbf{k}}\alpha_{\mathbf{k}_0+\mathbf{k}}^+\rangle=1+n_{\mathbf{k}_0+\mathbf{k}}^{ex}=\frac{1}{|1-|A_{\mathbf{k}}|^2|}. \tag{20}$$

Thus, in the presence of resonant or nonresonant laser radiation, a high-density exciton gas consisting of real or virtual excitons appears. The difference between virtual excitons and real quasiparticles is that the former exist only in the course of the laser pulse and disappear after switch-off of the pulse. Real excitons exist during their entire lifetime. In the zeroth approximation, the states of the polarized crystal and those of the vacuum photon fields are independent, since the exciton–photon Hamiltonian in this case is only responsible for quantum transitions. The polarization effect in the presence of coherent macroscopic polarization was taken into account in Ref. 5.

Following the results of Ref. 16, we consider a light-induced quantum transition from the ground state $|0\rangle$ of a coherently polarized crystal into the quasiexcitonic state with wave vector $\mathbf{P}$, which we denote by $\xi_{\mathbf{P}}^+|0\rangle$. We write the initial and final states of a two-component system consisting of the polarized crystal and the vacuum field in the rotating coordinate system as

$$|i\rangle=|0\rangle_{ex}\eta_{\mathbf{Q}}^+|0\rangle_{ph}, \quad E_i=\hbar(c|\mathbf{k}_0+\mathbf{Q}|-\omega_L),$$
$$|f\rangle=\xi_{\mathbf{P}}^+|0\rangle_{ex}|0\rangle_{ph}, \quad E_f=E_1(\mathbf{P})=\mathscr{E}_1(\mathbf{P})+\hbar\mathbf{V}_s\mathbf{P}, \tag{21}$$

where $|0\rangle_{ph}$ is the ground state of vacuum photons.

The amplitude of the transition involving the interaction Hamiltonian that enters into (17) is

$$\langle i|H_{\text{int}}|f\rangle=\delta_{\mathbf{Q},\mathbf{P}}\frac{\lambda_{\mathbf{k}_0+\mathbf{P}}}{\sqrt{1-|A_{\mathbf{P}}|^2}}. \tag{22}$$

The transition probability summed over final states $\mathbf{P}$ for a fixed vector $\mathbf{Q}$ of the photon in the initial state is

$$P_{\text{absorb}}(\mathbf{Q})=\frac{2\pi}{\hbar}\frac{|\lambda_{\mathbf{k}_0+\mathbf{Q}}|^2}{|1-|A_{\mathbf{Q}}|^2|}$$
$$\times\delta(\hbar c|\mathbf{k}_0+\mathbf{Q}|-\hbar\omega_L-E_1(\mathbf{Q})). \tag{23}$$

The transition probability is a function of the vacuum-photon wave vector $\mathbf{Q}$, reckoned from the laser-photon wave vector $\mathbf{k}_0$. The total vacuum-photon wave vector is $\mathbf{q}=\mathbf{k}_0+\mathbf{Q}$, and the energy of this photon is $\hbar\omega=\hbar cq=\hbar c|\mathbf{k}_0+\mathbf{Q}|$. The probability of such a transition as a function of the wave vector $\mathbf{q}$ is

$$P_{\text{absorb}}(\mathbf{q})=\frac{2}{\hbar}\frac{|\lambda_{\mathbf{q}}|^2}{|1-|A_{\mathbf{q}-\mathbf{k}_0}|^2|}$$
$$\times\frac{\Gamma(\mathbf{q}-\mathbf{k}_0)}{(\hbar\omega-\hbar\omega_L-E_1(\mathbf{q}-\mathbf{k}_0))^2+\Gamma^2(\mathbf{q}-\mathbf{k}_0)}, \tag{24}$$

with $\hbar cq=\hbar\omega$. Here we have introduced the damping factor $\Gamma(\mathbf{Q})$ for an elementary excitation with energy $E_1(\mathbf{Q})$, and represented the delta function by a Lorentzian:

$$\delta(\chi)=\frac{1}{\pi}\frac{\Gamma}{\chi^2+\Gamma^2}, \quad \Gamma\to 0. \tag{25}$$

The transition probability strongly depends on the orientation of $\mathbf{q}$ relative to $\mathbf{k}_0$. This dependence is especially apparent in the coefficients $|A_{\mathbf{q}-\mathbf{k}_0}|^2$ and is discussed in Sec. 3.

However, the quantum transition with photon absorption and a Stokes mechanism of creation of a quasiexcitonic elementary excitation is not the only possible transition. There

are three more possibilities. One is the anti-Stokes process of photon absorption and quasiexcitonic elementary excitation. This is an antiresonant quantum transition. The probability of such a transition is nil, however, the reason being that at $T=0$ there are no quasiexcitonic elementary excitations in the ground state.

What remains is the Stokes process of light emission with simultaneous creation of a quasiexcitonic elementary excitation. Here the initial state is the vacuum one and the final state is the two-particle state

$$|i\rangle=|0\rangle_{ex}|0\rangle_{ph}, \quad |f\rangle=\xi_{\mathbf{P}}^{+}|0\rangle_{ex}\eta_{\mathbf{Q}}^{+}|0\rangle_{ph},$$

$$E_f=\hbar c|\mathbf{k}_0+\mathbf{Q}|-\hbar\omega_L+E_1(\mathbf{P}), \quad E_i=0. \tag{26}$$

This transition is also quasiresonant. In the final two-particle state there is a photon with a fixed wave vector $\mathbf{Q}$, and a second quasiparticle, a quasiexcitonic elementary excitation with arbitrary wave vector $\mathbf{P}$. The probability of a transition involving the third term in the interaction Hamiltonian (17), summed over final states ($\Sigma_{\mathbf{P}}$) and averaged over initial states, which here amount to only one state, has the form

$$P_{\text{emit}}(\mathbf{Q})=\frac{2\pi|\lambda_{\mathbf{k}_0+\mathbf{Q}}|^2}{\hbar}\frac{|A_{\mathbf{Q}}|^2}{|1-|A_{\mathbf{Q}}|^2|}\,\delta(\hbar c|\mathbf{k}_0+\mathbf{Q}|-\hbar\omega_L+E_1(-\mathbf{Q})). \tag{27}$$

The law of energy conservation corresponding to this two-particle antiresonant process can be written

$$\hbar c|\mathbf{k}_0+\mathbf{Q}|-\hbar\omega_L+E_1(-\mathbf{Q})=0,$$

$$\hbar c|\mathbf{k}_0+\mathbf{Q}|=\hbar\omega_L+E_2(\mathbf{Q}). \tag{28}$$

Bearing in mind that the branch $\hbar\omega_L+E_1(-\mathbf{Q})$ approximately follows the exciton dispersion law $\hbar\omega_{ex}(\mathbf{k}_0-\mathbf{Q})$, we can rewrite Eq. (28) approximately as

$$\hbar c|\mathbf{k}_0+\mathbf{Q}|+\hbar\omega_{ex}(\mathbf{k}_0-\mathbf{Q})\approx 2\hbar\omega_L. \tag{29}$$

This equation indicates that two laser photons in the system are transformed into a weak-source photon and an extracondensate exciton:

$$\text{photon}(\mathbf{k}_0)+\text{photon}(\mathbf{k}_0)$$

$$=\text{photon}(\mathbf{k}_0+\mathbf{Q})+\text{exciton}(\mathbf{k}_0-\mathbf{Q}). \tag{30}$$

This all suggests that the only reason for emission and amplification of a weak signal is the presence of external laser radiation. Here it is advisable to recall the results arrived at in Refs. 17 and 18 in the studies of the shapes of absorption and luminescence bands at absolute zero in conditions of spontaneous quasiequilibrium Bose–Einstein condensation of excitons in semiconductors. The bands consisted of narrow central peaks at frequencies close to the energy of the Bose-condensed excitons, with wider wings. In relation to the central peak, the absorption band wing was on the higher-energy side. The intensity of the wing was determined by the coefficients $U_{\mathbf{q}}^2=1+n_{\mathbf{q}}^{ex}$. An elementary excitation was found to be produced simultaneously with the absorption of light. The luminescence band wing was on the lower-energy side of the central peak, since an elementary excitation was created simultaneously with photon emission. The intensity of the wing was proportional to the coefficients $V_{\mathbf{q}}^2=n_{\mathbf{q}}^{ex}$, which are less than $U_{\mathbf{q}}^2$.

Equation (31) below describes a similar situation. Note that the probability of anti-Stokes emission of a photon accompanied by absorption of a quasiexcitonic elementary excitation is nil, since at $T=0$ there can be no elementary excitations. Thus, we have the probability of photon absorption (Eq. (23)) and the probability of photon emission (Eq. (27)) in the same system. The probability of light absorption minus the probability of light emission yields the probability of true absorption:

$$P_{\text{net absorb}}(\mathbf{Q})=P_{\text{absorb}}(\mathbf{Q})-P_{\text{emit}}(\mathbf{Q})=\frac{2\pi|\lambda_{\mathbf{k}_0+\mathbf{Q}}|^2}{\hbar}\times\left[\frac{1}{|1-|A_Q|^2|}\,\delta(\hbar c|\mathbf{k}_0+\mathbf{Q}|-\hbar\omega_L-E_1(\mathbf{Q}))-\frac{|A_Q|^2}{|1-|A_Q|^2|}\,\delta(\hbar c|\mathbf{k}_0+\mathbf{Q}|-\hbar\omega_L+E_1(-\mathbf{Q}))\right]. \tag{31}$$

In ranges of frequency $\hbar\omega$ where this difference is positive there is true light absorption, and in ranges where it is negative there is true emission, or light amplification. After the delta functions are replaced by Lorentzians, we have the desired probability at $T=0$:

$$P_{\text{net absorb}}(\mathbf{Q})=\frac{2|\lambda_{\mathbf{k}_0+\mathbf{Q}}|^2}{\hbar}\left[\frac{1}{|1-|A_Q|^2|}\times\frac{\Gamma(\mathbf{Q})}{(\hbar c|\mathbf{k}_0+\mathbf{Q}|-\hbar\omega_L-E_1(\mathbf{Q}))^2+\Gamma^2(\mathbf{Q})}-\frac{|A_Q|^2}{|1-|A_Q|^2|}\times\frac{\Gamma(-\mathbf{Q})}{(\hbar c|\mathbf{k}_0+\mathbf{Q}|-\hbar\omega_L+E_1(-\mathbf{Q}))^2+\Gamma^2(-\mathbf{Q})}\right]. \tag{32}$$

### 3. ANISOTROPY OF THE EXCITON BANDS OF ABSORPTION AND EMISSION IN A COHERENTLY POLARIZED CRYSTAL

The energy spectrum $\mathscr{E}_i(\mathbf{k})$ (Eq. (11)), the coefficients $|A_{\mathbf{k}i}|$ (Eq. (10)), and the factors $(1-|A_{\mathbf{k}i}|^2)^{-1}$ are depicted in Fig. 1 as functions of the wave vector $\mathbf{k}$ for several values of the detuning $\tilde{\Delta}$ (Eq. (7)): $\tilde{\Delta}=L_0$, 0, and $-L_0$. The functions $\mathscr{E}_i(\mathbf{k})$ were derived in Ref. 6, but the coefficients $A_{\mathbf{k}i}$ and quantum transition probabilities have not been studied.

The same expressions, viewed as functions of $|\mathbf{q}-\mathbf{k}_0|$, depend on the orientation of the wave vector $\mathbf{q}$ of the probe light relative to the wave vector $\mathbf{k}_0$ of the laser light that produces coherent macroscopic polarization in the medium. Three possible observational geometries for the light absorption coefficient are of interest here: $\mathbf{q}\uparrow\uparrow\mathbf{k}_0$, $\mathbf{q}\perp\mathbf{k}_0$, and $\mathbf{q}\downarrow\uparrow\mathbf{k}_0$, i.e., in which the probe radiation propagates in the

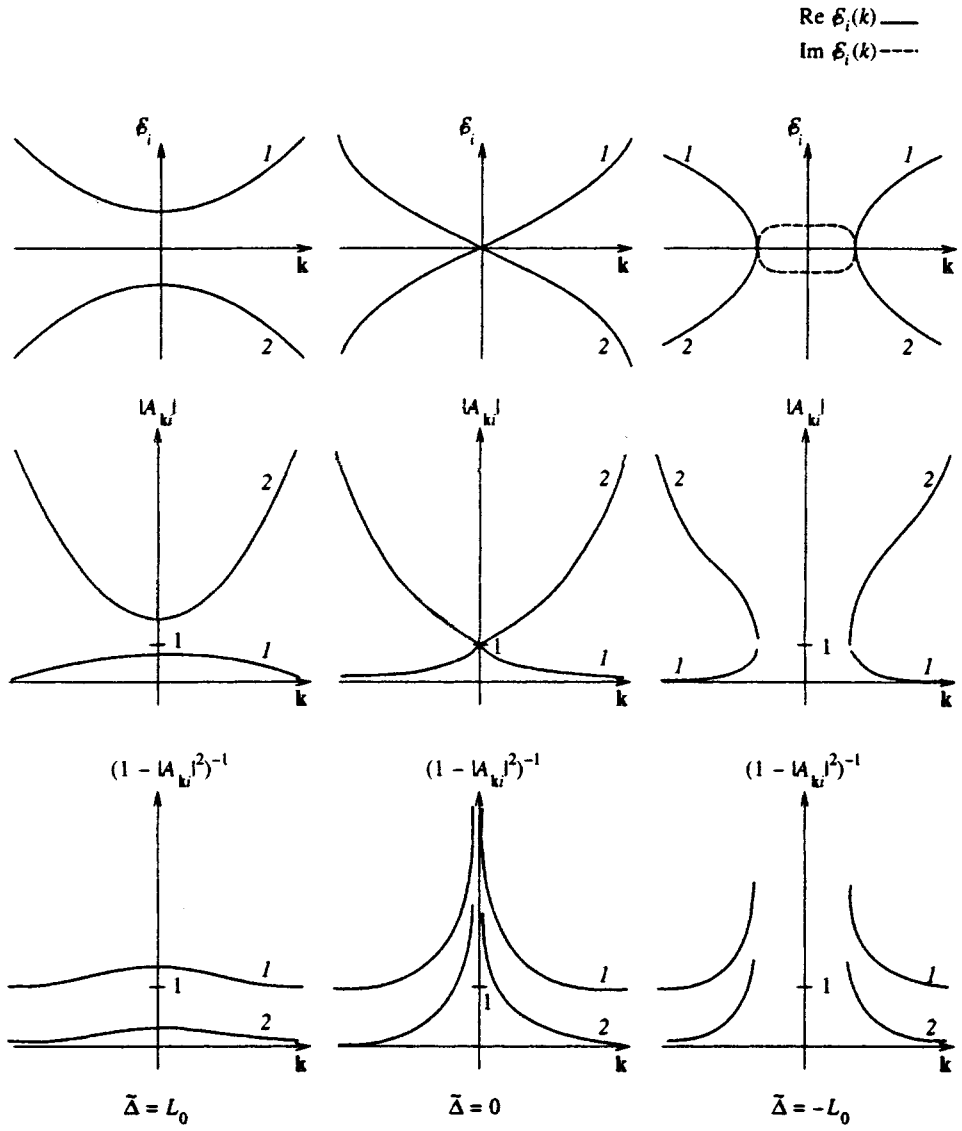



FIG. 1. Energy spectrum $\mathscr{E}_i(\mathbf{k})$, coefficients $|A_{\mathbf{k}i}|$, and factors $(1-|A_{\mathbf{k}i}|^2)^{-1}$ as functions of the wave vector $\mathbf{k}$ for various values of $\tilde{\Delta}$. Curves *1* correspond to solutions of the first type, and curves *2* to solutions of the second type.

same direction as the laser radiation, perpendicular to the laser radiation, and antiparallel to the laser radiation.

The quantity $|(\mathbf{q}-\mathbf{k}_0)|/|\mathbf{k}_0|$ takes on differing values for the same value of $|\mathbf{q}|=x|\mathbf{k}_0|$, with $0<x<\infty$, depending on the orientation of $\mathbf{q}$ relative to $\mathbf{k}_0$. The values are $|x-1|$, $\sqrt{x^2+1}$, and $x+1$ for parallel, perpendicular, and antiparallel orientations, respectively.

For this reason the energy spectrum $\mathscr{E}(\mathbf{q}-\mathbf{k}_0)$, coefficients $|A_{\mathbf{q}-\mathbf{k}_0}|$, and factors $(1-|A_{\mathbf{q}-\mathbf{k}_0}|)^{-1}$ as functions of $x$ are given for three orientations of $\mathbf{q}$ relative to $\mathbf{k}_0$. A dependence on $x$ also means a dependence on the frequency of the absorbed light, since $\hbar\omega=\hbar cq=x\hbar ck_0=x\hbar\omega_L$, where $x>0$. All this is shown in Figs. 2a, 2b, and 2c. Each figure is drawn for a definite detuning from resonance, $\tilde{\Delta}$, and contains the frequency curves $\mathscr{E}_1(x)$, $|A(x)|$, and $(1-|A(x)|^2)^{-1}$ for three observation geometries.

The frequency dependence of $(1-|A(x)|^2)^{-1}$ depicted in Fig. 2 is markedly anisotropic. The anisotropy shows up for positive resonance offset $\tilde{\Delta}=L_0$, and amounts to about 2%. At large positive values of $\tilde{\Delta}$ the anisotropy becomes vanishingly small, since the coefficients $|A_{\mathbf{k}}|$ for the exciton-like branch of the spectrum become much less than unity. At $\tilde{\Delta}=0$ we have a real but nonequilibrium induced Bose condensate of excitons with wave vector $\mathbf{k}_0$. The exciton occupation numbers $n^{ex}_{\mathbf{k}_0+\mathbf{k}}$ (Eq. (19)) tend to infinity as $k\to 0$, as follows from the Bose–Einstein distribution function with a chemical potential equal to zero. This leads to a large discrepancy between the occupation numbers

$$n^{ex}_{\mathbf{k}_0+\mathbf{k}}=n^{ex}_{\mathbf{q}}=\frac{|A_{\mathbf{q}-\mathbf{k}_0}|^2}{1-|A_{\mathbf{q}-\mathbf{k}_0}|^2}, \quad \mathbf{q}=\mathbf{k}_0+\mathbf{k}, \tag{33}$$

when $\mathbf{q}$ tends to $\mathbf{k}_0$ and when $\mathbf{q}$ tends to $-\mathbf{k}_0$. Since the light absorption coefficients are proportional to $1+n^{ex}_{\mathbf{q}}$, the absorption of light becomes highly anisotropic.

At negative $\tilde{\Delta}$ the spectrum becomes unstable, a situation discussed above. In wave vector ranges harboring instability, the coefficients $|A_{\mathbf{k}}|^2=1$, and the corresponding occupation numbers tend to infinity.

In these wave-vector ranges, the canonical transformations (9) become invalid. New waves are generated and stimulated exciton scattering occurs. In our case the scatter-

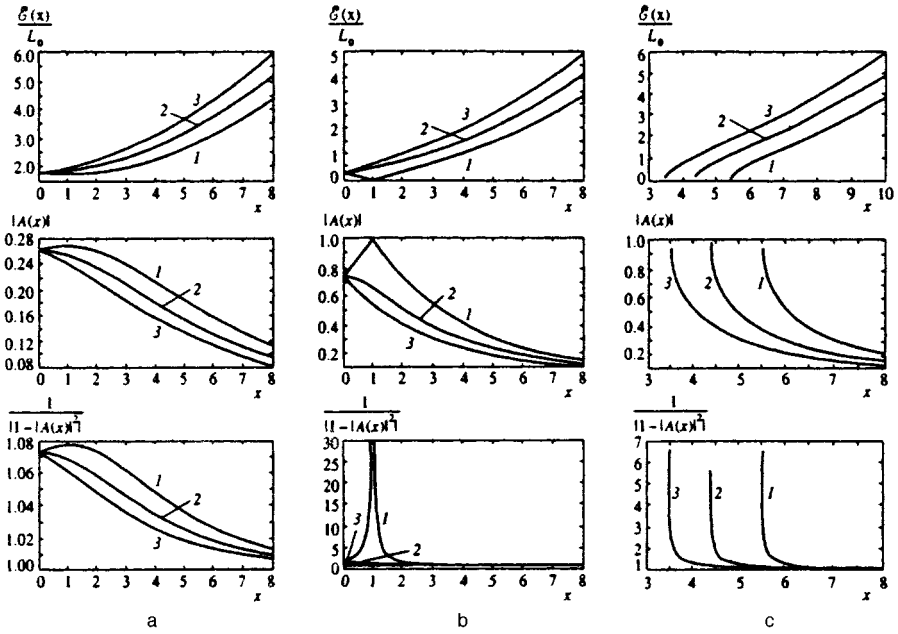



FIG. 2. $\mathscr{E}_1(x)$, $|A(x)|$, and $(1-|A(x)|^2)^{-1}$ as functions of the frequency for various values of $\tilde{\Delta}$ ((a) $\tilde{\Delta}=L_0$, (b) $\tilde{\Delta}=0$, and (c) $\tilde{\Delta}=-L_0$) and for three orientations of the wave vector $\mathbf{q}$ of the probe light relative to the wave vector $\mathbf{k}_0$ of the laser radiation: curves *1* correspond to $\mathbf{q}\uparrow\uparrow\mathbf{k}_0$, curves *2* to $\mathbf{q}\perp\mathbf{k}_0$, and curves *3* to $\mathbf{q}\downarrow\uparrow\mathbf{k}_0$.

ing is of the nonthreshold type, since we have not allowed for the damping of the energy spectrum of the initial excitons. Since the number of nonequilibrium excitons in these regions of **k**-space becomes anomalously high, so does the light absorption coefficient. Figure 2c clearly demonstrates this. Interestingly, the regions where anomalous light absorption is possible also are shifted along the energy scale, depending on the observation geometry. In addition to the factor $(1-|A(x)|^2)^{-1}$, the probabilities of true light absorption, Eqs. (31) and (32), contain a second factor, which in one case is a delta function and in the other a Lorentzian. A Lorentzian can also possess a varying frequency dependence for different orientations of $\mathbf{q}$ and $\mathbf{k}_0$, but it also depends on the small difference of such quantities as $\hbar\omega$ and $\hbar\omega_L$, each of which is large compared to the elementary excitation spectrum $E_1(\mathbf{q}-\mathbf{k}_0)=\mathscr{E}_1(\mathbf{q}-\mathbf{k}_0)+\hbar\mathbf{V}_s(\mathbf{q}-\mathbf{k}_0)$. Anisotropy therefore shows up more clearly in terms of the factor $(1-|A_{\mathbf{q}-\mathbf{k}_0}|^2)^{-1}$. Furthermore, if the arguments of the delta functions in Eq. (31) or of the corresponding Lorentzians are identical, all anisotropy of the sort discussed above disappears. In this case, the regions with pure absorption and pure emission coincide, and the anisotropies cancel perfectly.

Thus, anisotropy becomes imperceptible if the Lorentzians have a halfwidth larger than $2|\mathscr{E}_1(\mathbf{Q})|$. The most favorable conditions for observing the anisotropy in the absorption and luminescence bands correspond to $\Gamma(\mathbf{Q})<|\mathscr{E}_1(\mathbf{Q})|$ and small $\tilde{\Delta}$, whereupon the $|A_{\mathbf{Q}}|^2$ are close to unity.

Note that the anisotropy of two-photon absorption in the transition from the ground state of a crystal to a biexcitonic state in the presence of laser radiation was studied in Ref. 19. When the exciton levels are degenerate, they can split and the corresponding quantum transitions can become polarized. These aspects under the conditions of the optical Stark effect were studied by Combescot.[20] In contrast, in our case the exciton level is nondegenerate and the anisotropy of the quantum transitions depends on the direction of propagation of the probe signal with respect to the wave vector of the laser radiation.

In conclusion we note that the properties of the absorption and light-amplification bands depend strongly on the values of $n_{\mathbf{k}_0}$ and $\tilde{\Delta}$. These quantities are interrelated and can be determined self-consistently by Eqs. (7). Generally they follow a bistable curve characterized by the presence of amplitude and frequency hysteresis loops. The stationary values of $n_{\mathbf{k}_0}$ in the hysteresis loops or near such loops are usually not Lyapunov stable. In such cases, the amplitude of the coherent wave experiences self-pulsations, and the steady state becomes self-oscillatory. The self-pulsations may be cyclic if the phase trajectory describes a limit cycle. In more complicated cases the phase trajectory can move along the surface of a torus or even be a strange attractor.[21]

Such self-pulsations have a higher probability of forming when the coherent pump radiation is turned on or off. In the case of stimulated Brillouin scattering this phenomenon was studied by Keldysh and Tikhodeev,[13] who found that self-pulsations of the amplitude of the anti-Stokes component have a frequency determined by the splitting of phonoriton curves. The amplitude of the Stokes component increases without limit as a function of time in the number density of coherent polaritons exceeds a certain threshold. This is due to the emergence of absolute or convective instability in the polariton–phonon system. Similar phenomena can occur in the case studied here.

The authors are grateful to S. G. Tikhodeev form drawing their attention to these effects.

The present work was made possible by a grant from the INTAS international project (Grant No. 94-324). The authors are grateful to the coordinator of the project, K. Klingshearn, and to M. S. Brodin for collaboration.

[1] A. Mysyrowicz, D. Hulin, A. Antonetti, A. Migus, W. T. Masselink and H. Morkoc, Phys. Rev. Lett. **56**, 2748 (1986).
[2] A. Von Lehmen, D. S. Chemla, J. E. Zucher, and J. P. Heritage, Opt. Lett. **11**, 609 (1986).
[3] S. Schmitt-Rink and D. S. Chemla, Phys. Rev. Lett. **57**, 2752 (1986).
[4] S. Schmitt-Rink, D. S. Chemla, and H. Haug, Phys. Rev. B **37**, 941 (1988).
[5] V. R. Mis'ko, S. A. Moskalenko, and M. I. Shmiglyuk, Fiz. Tverd. Tela (St. Petersburg) **35**, 3213 (1993) [Phys. Solid State **36**, 1580 (1993)].
[6] S. A. Moskalenko and V. R. Mis'ko, Ukr. Fiz. Zh. **37**, 1812 (1992).
[7] L. V. Keldysh and A. N. Kozlov, Zh. Éksp. Teor. Fiz. **54**, 978 (1968) [Sov. Phys. JETP **27**, 521 (1968)].
[8] S. A. Moskalenko, Fiz. Tverd. Tela (Leningrad) **4**, 276 (1962) [Sov. Phys. Solid State **4**, 199 (1962)].
[9] R. Zimmermann, Phys. Status Solidi B **76**, 191 (1976).
[10] V. F. Elesin and Yu. V. Kopaev, Zh. Éksp. Teor. Fiz. **63**, 1447 (1972) [Sov. Phys. JETP **35**, 766 (1973)].
[11] E. M. Lifshitz and Pitaevskiĭ, *Physical Kinetics*, Pergamon Press, Oxford (1981), §62–64.
[12] L. V. Keldysh and S. G. Tikhodeev, Zh. Éksp. Teor. Fiz. **90**, 1852 (1986) [Sov. Phys. JETP **63**, 1086 (1986)].
[13] L. V. Keldysh and S. G. Tikhodeev, Zh. Éksp. Teor. Fiz. **91**, 78 (1986) [Sov. Phys. JETP **64**, 45 (1986)].
[14] V. P. Mis'ko, S. A. Moskalenko, A. Kh. Rotaru, and Yu. M. Shvera, Zh. Éksp. Teor. Fiz. **99**, 1215 (1991) [Sov. Phys. JETP **72**, 676 (1991)].
[15] N. N. Bogolyubov, *Collected Scientific Works* [in Russian], Naukova Dumka, Kiev (1971).
[16] A. I. Bobrysheva, M. I. Shmiglyuk, and S. S. Russu, Proc. SPIE **1807**, 79 (1993).
[17] A. V. Lelyakov and S. A. Moskalenko, Fiz. Tverd. Tela (Leningrad) **11**, 3260 (1969) [Sov. Phys. Solid State **11**, 2642 (1969)].
[18] S. A. Moskalenko, *Bose–Einstein Condensation of Excitons and Biexcitons* [in Russian], RIO, Kishinev (1970).
[19] A. I. Bobrysheva, S. A. Moskalenko, and Hoang Ngok Kam, Zh. Éksp. Teor. Fiz. **103**, 301 (1993) [JETP **76**, 163 (1993)].
[20] M. Combescot, Phys. Rev. B **41**, 3517 (1990).
[21] V. A. Zalozh, S. A. Moskalenko, and A. Kh. Rotaru, Zh. Éksp. Teor. Fiz. **95**, 601 (1989) [Sov. Phys. JETP **68**, 338 (1989)].

Translated by Eugene Yankovsky

# The effect of a local field on Raman scattering in a uniaxial crystal

M. V. Gorkunov and M. I. Ryazanov

*Moscow Engineering-Physical Institute, 115409 Moscow, Russia*
(Submitted 19 September 1996)
Zh. Éksp. Teor. Fiz. **112**, 180–191 (July 1997)

This paper discusses the Raman scattering of light in an anisotropic crystal in the crystal optics approximation, taking into account local fields acting on the molecules. It shows that the effect of the local field reduces to the introduction of the effective Raman polarizability tensor of the molecules, which depends both on the properties of the molecules themselves and on the characteristics of the crystal at the frequencies of the incident and scattered waves. Raman scattering cross sections are obtained in a uniaxial crystal for various types of incident waves. It is shown that, in the case of an extraordinary incident wave, the local field substantially affects how the cross section depends on the direction of incidence. © *1997 American Institute of Physics.* [S1063-7761(97)01507-2]

## 1. INTRODUCTION

Ordered placement of the atoms of a crystal is essential for those processes in which the atoms participate coherently. Raman scattering in which the state of the atomic electron varies is not related to such processes. Therefore, the existence of the crystal lattice affects Raman scattering only because the properties of electromagnetic waves in the crystal depend on it.[1] For waves of optical frequencies, it is possible to use the crystal optics approximation, which treats a crystal as a homogeneous anisotropic substance. This means that only the ordering of the orientations of anisotropic molecules of the crystal is taken into account, but the ordering of the spatial placement of the centers of inertia of the molecules is neglected in this case. The properties of the crystal in this case do not differ from the properties of a homogeneous amorphous substance with identically oriented anisotropic molecules. Taking into account the ordering of the spatial placement of the molecules results in small corrections, of order the ratio of the lattice constant to the wavelength of the field, which is three orders of magnitude smaller than the quantities obtained in the crystal optics approximation. Therefore, Raman scattering in a crystal can be treated with good accuracy in the crystal optics approximation. A rigorous treatment of the crystal in the case of Raman scattering is much more involved but results in only insubstantial additional corrections by comparison with crystal-optics corrections.

Raman scattering is considered below in the crystal optics approximation as in a homogeneous amorphous substance with ordered orientation of the anisotropic molecules. With such a treatment, one has to appeal to an amorphous medium at intermediate stages, so that it is natural to use terms that are used in considering an amorphous substance and sometimes do not coincide with the terms used in a rigorous treatment of a crystal.[2,3]

Spontaneous Raman scattering in a gas differs from that in a condensed state of the same substance for two reasons. First, the electrons of the valence bonds are in different states in the gas and in the condensed state, which causes the Raman scattering cross sections to be different at an atom of the gas and an atom of the condensed substance. Second, the action of adjacent molecules is substantial in a dense medium, so that the mean field acting on a molecule (the local field) differs from the mean macroscopic field. In a gas, the local field coincides with the mean field.

In considering Raman scattering in a crystal, one is usually given the macroscopic field of the incident wave and the scattering amplitude at an individual molecule of the crystal, and one is required to find the macroscopic field of the scattered wave. The first stage of the solution of this problem reduces to determining the local field acting on the given molecule. Knowing the scattering amplitude, it is possible to find the transition current that appears in a molecule under the action of the local field. The fact that the scattered wave is emitted from where the atom is situated makes it necessary at the second stage to solve the inverse problem—finding the macroscopic field of the scattered wave from a given dipole moment.

In the final analysis, scattering at a molecule that forms part of a crystal occurs as if this molecule possessed some effective Raman polarizability that depends on both the Raman polarizability of the molecule and the macroscopic characteristics of the crystal.

Earlier, the local field was taken into account in Refs. 4–7 for an isotropic substance when there is no dependence on the orientation of the fields and everything reduces to correcting factors in the overall characteristics.

It is interesting to estimate the local field effects in a uniaxial crystal when they affect the angular distribution and depend on the orientation of the field of the incident wave.

## 2. THE LOCAL FIELD IN A UNIAXIAL CRYSTAL

We shall regard a uniaxial crystal as a system of homogeneously distributed molecules that are assumed to be axially symmetric and oriented parallel to the optical axis of the crystal.

A molecule located at point $\mathbf{r}$ in a crystal lying in an electromagnetic field acquires a dipole moment whose Fourier transform $d_i(\mathbf{r},\omega)$ is associated with the polarizability $\alpha_{ij}(\omega)$ of the molecule and the Fourier transform of the mi-

 1063-7761/97/070097-07$10.00  

croscopic field $E_j^{\rm mic}(\mathbf{r},\omega)$ acting on this molecule:

$$d_i(\mathbf{r},\omega)=\alpha_{ij}(\omega)E_j^{\rm mic}(\mathbf{r},\omega).$$

The dipole moments induced in each molecule are sources of secondary fields. The microscopic field acting on the molecule at point $\mathbf{R_a}$ consists of the primary field $\mathbf{E}^0$, no longer interacting with a single molecule, and the secondary fields of all the other molecules. Therefore, for the Fourier transforms of the fields, one can write

$$E_i^{\rm mic}(\mathbf{R_a},\omega)=E_i^0(\mathbf{R_a},\omega)+[2\pi^2]^{-1}\sum_b \int d^3qQ_{is}(\mathbf{q},\omega) \times\alpha_{sj}(\omega)E_j^{\rm mic}(\mathbf{R_b},\omega)\exp\{i\mathbf{q}\cdot(\mathbf{R_a}-\mathbf{R_b})\}, \quad (1)$$

where

$$Q_{is}(\mathbf{q},\omega)=\frac{q^2\delta_{is}-q_iq_s}{q^2-(\omega/c)^2}.$$

If the wavelength of the field is large by comparison with the lattice constant $d$, the effective field acting on a molecule is formed by adding the fields of many molecules lying in some volume of the crystal with linear dimensions of $L$. When the inequality

$$c/\omega\gg L\gg d \quad (2)$$

is valid, the value of the field acting on a molecule is close to the field of the other molecules averaged over position, usually called the local field. In the crystal optics approximation, Eq. (1) can be averaged over the coordinates of all the other molecules as for an amorphous substance. In this case, the microscopic field is replaced by the local field, and, to average the sum, it is sufficient to multiply each term by the probability of finding the molecule in volume $dV$ at distance $\mathbf{R}'=\mathbf{R_a}-\mathbf{R_b}$ from the given molecule:

$$dW(\mathbf{R}')=g(\mathbf{R}')dV/V=[1-f(\mathbf{R}')]dV/V \quad (3)$$

and to integrate over all $\mathbf{R}'$. After this, the summation reduces to multiplying by the number of identical terms, $N-1\approx N$. The distribution function $g(\mathbf{R}')$ vanishes for $R'\ll d$ because the probability of arbitrarily close approach of two molecules vanishes. On the other hand, there is no correlation of the position of two molecules at large distances, so that $f(\mathbf{R}')=0$ for $R'\gg d$ and $f(\mathbf{R}')=1$ for $R'\ll d$.

Introducing the polarization $\mathbf{P}$ of the substance, associated with the local field by

$$P_i(\mathbf{q},\omega)=n_0\alpha_{ij}(\omega)E_j^{\rm loc}(\mathbf{q},\omega), \quad (4)$$

where $n_0=N/V$ is the number of molecules in unit volume, we can write the result of averaging Eq. (1) in the form of an equation for the Fourier transforms of the fields over the spatial coordinates:

$$P_i(\mathbf{q},\omega)=n_0\alpha_{ij}(\omega)E_j^0(\mathbf{q},\omega)+4\pi n_0\alpha_{ij}Q_{js}(\mathbf{q},\omega) \times P_s(\mathbf{q},\omega)-4\pi n_0\alpha_{ij}(\omega)\int d^3lQ_{js}(\mathbf{l},\omega) \times f(\mathbf{q}-\mathbf{l})P_s(\mathbf{q},\omega), \quad (5)$$

where

$$f(\mathbf{q}-\mathbf{l})=(2\pi)^{-3}\int d^3Rf(\mathbf{R})\exp\{-i(\mathbf{q}-\mathbf{l})\cdot\mathbf{R}\}.$$

Equation (5) connects the polarization of a substance with the primary field, i.e., with the field that satisfies Maxwell's equations in a vacuum with the same current density as in the equations for the mean macroscopic field $\mathbf{E}$. This circumstance makes it possible to express the primary field $\mathbf{E}^0$ in terms of $\mathbf{E}$ and $\mathbf{P}$:

$$E_i^0(\mathbf{q},\omega)=E_i(\mathbf{q},\omega)+4\pi P_i(\mathbf{q},\omega) -4\pi Q_{is}(\mathbf{q},\omega)P_s(\mathbf{q},\omega). \quad (6)$$

Eliminating $\mathbf{E}^0$ from Eqs. (5) and (6), it is possible to obtain an equation that connects the mean macroscopic field with the polarization:

$$n_0\alpha_{ij}E_j(\mathbf{q},\omega)=P_i(\mathbf{q},\omega)-4\pi n_0\alpha_{ij}P_j(\mathbf{q},\omega)+4\pi n_0\alpha_{ij} \times\int d^3lQ_{js}(\mathbf{l},\omega)f(\mathbf{q}-\mathbf{l})P_s(\mathbf{l},\omega). \quad (7)$$

Introducing, as is usually done, the electric susceptibility tensor $\chi_{ij}$ connecting the polarization with the macroscopic field:

$$P_i(\mathbf{q},\omega)=\chi_{ij}(\mathbf{q},\omega)E_j(\mathbf{q},\omega),$$

we get for it

$$\chi_{ij}(\mathbf{q},\omega)=\{\alpha_{ij}^{-1}(\omega)/n_0+4\pi[T_{ij}(\mathbf{q},\omega)-\delta_{ij}]\}^{-1}, \quad (8)$$

where

$$T_{ij}(\mathbf{q},\omega)=\int d^3lf(\mathbf{q}-\mathbf{l})\,\frac{l^2\delta_{ij}-l_il_j}{l^2-(\omega/c)^2}.$$

Taking into account the inequalities (2), the integral in $T_{ij}$ is built up when $l\gg q,\ \omega/c$. It is therefore possible to assume the approximation $f(\mathbf{q}-\mathbf{l})\approx f(\mathbf{l})$ and neglect $(\omega/c)^2$ in the denominator. Then the dependence on $\mathbf{q}$ disappears, and Eq. (8) gives an electric susceptibility without spatial dispersion. In this case, it is possible to transform from integration over wave vectors to integration over coordinates, and then

$$T_{ij}=\frac{2}{3}\,\delta_{ij}-\int d^3Rf(\mathbf{R})\,\frac{R^2\delta_{ij}-3R_iR_j}{R^5}.$$

In a uniaxial crystal, $f(\mathbf{R})$ is axially symmetric, while the molecules are strictly oriented along the unit vector $\mathbf{e}$ of the optical axis. Consequently, the optical axis of the crystal is the principal axis of the axially symmetric tensors $\alpha_{ij}$ and $T_{ij}$; i.e., they can be represented in the form

$$\alpha_{ij}=\alpha_\perp(\delta_{ij}-e_ie_j)+\alpha_\parallel e_ie_j,$$
$$T_{ij}=T_\perp(\delta_{ij}-e_ie_j)+T_\parallel e_ie_j, \quad (9)$$

with

$$T_\parallel=\frac{2}{3}-\frac{1}{2}\int d^3Rf(\mathbf{R})R^{-5}[R^2-3(\mathbf{R}\cdot\mathbf{e})^2],$$

$$T_\perp=\frac{2}{3}+\frac{1}{4}\int d^3Rf(\mathbf{R})R^{-5}[R^2-3(\mathbf{R}\cdot\mathbf{e})^2].$$

The electric susceptibility accordingly takes the form

$$\chi_{ij}(\omega)=\chi_{\perp}(\omega)(\delta_{ij}-e_ie_j)+\chi_{\parallel}(\omega)e_ie_j,$$

where

$$\chi_{\perp}(\omega)=\frac{n_0\alpha_{\perp}(\omega)}{1-4\pi n_0\alpha_{\perp}(\omega)(1-T_{\perp})},$$
$$\chi_{\parallel}(\omega)=\frac{n_0\alpha_{\parallel}(\omega)}{1-4\pi n_0\alpha_{\parallel}(\omega)(1-T_{\parallel})}. \tag{10}$$

Finally, for the permittivity of a uniaxial crystal, we have

$$\varepsilon_{ij}=\delta_{ij}+4\pi\chi_{ij}(\omega)=\varepsilon_{\perp}(\delta_{ij}-e_ie_j)+\varepsilon_{\parallel}e_ie_j,$$
$$\varepsilon_{\perp}=\frac{1+4\pi n_0T_{\perp}\alpha_{\perp}}{1-4\pi n_0\alpha_{\perp}(1-T_{\perp})},$$
$$\varepsilon_{\parallel}=\frac{1+4\pi n_0T_{\parallel}\alpha_{\parallel}}{1-4\pi n_0\alpha_{\parallel}(1-T_{\parallel})}. \tag{11}$$

In the particular case of an isotropic medium, when $\alpha_{\perp}=\alpha_{\parallel}$ and $f(\mathbf{R})=f(R)$, Eq. (9) gives simply $T_{\perp}=T_{\parallel}=2/3$ and Eq. (11) goes over to the well-known formula for the permittivity in the Lorentz–Lorenz model.

### 3. THE RAMAN SCATTERING AMPLITUDE IN A CRYSTAL

Since Raman scattering is an incoherent process, for which the intensities of the scattered waves from each molecule add up, we calculate the scattering at one molecule of the crystal, located at the origin of coordinates.

We shall consider the incident wave with frequency $\omega$ and amplitude $\mathbf{F}$ strong by comparison with the scattered waves having frequency $\omega'\neq\omega$. Neglecting backscattering from $\omega'$ into $\omega$, we can write the dipole moment of the molecule at the required frequencies in the form

$$d_i(\mathbf{R}=0,\omega)=\alpha_{ij}(\omega)E_j^{\mathrm{mic}}(\mathbf{R}=0,\omega),$$
$$d_i(\mathbf{R}=0,\omega')=\alpha_{ij}(\omega')E_j^{\mathrm{mic}}(\mathbf{R}=0,\omega')+\beta_{ij}(\omega,\omega')E_j^{\mathrm{mic}}(\mathbf{R}=0,\omega). \tag{12}$$

The following formulas of perturbation theory are valid for the polarizability tensors:

$$\alpha_{ij}(\omega)=\sum_l\left\{\frac{\langle n|d_i|l\rangle\langle l|d_j|n\rangle}{\omega_{ln}-\omega-i0}+\frac{\langle n|d_j|l\rangle\langle l|d_i|n\rangle}{\omega_{ln}+\omega+i0}\right\},$$
$$\beta_{ij}(\omega,\omega')=\delta(\omega'-\omega+\omega_{mn})\sum_l\left\{\frac{\langle m|d_i|l\rangle\langle l|d_j|n\rangle}{\omega_{ln}-\omega-i0}+\frac{\langle m|d_j|l\rangle\langle l|d_i|n\rangle}{\omega_{lm}+\omega+i0}\right\}. \tag{13}$$

If states $n$ and $m$ are axially symmetric, the tensors can be represented in the form of Eq. (9) and, consequently,

$$\alpha_{\perp}=e^2\sum_l|\langle l|x|n\rangle|^2 2\omega_{ln}(\omega_{ln}^2-\omega^2)^{-1},$$
$$\alpha_{\parallel}=e^2\sum_l|\langle l|z|n\rangle|^2 2\omega_{ln}(\omega_{ln}^2-\omega^2)^{-1},$$
$$\beta_{\perp}=e^2\sum_l\frac{\langle m|x|l\rangle\langle l|x|n\rangle(\omega_{ln}+\omega_{lm})}{(\omega_{ln}-\omega)(\omega_{lm}+\omega)},$$
$$\beta_{\parallel}=e^2\sum_l\frac{\langle m|z|l\rangle\langle l|z|n\rangle(\omega_{ln}+\omega_{lm})}{(\omega_{ln}-\omega)(\omega_{lm}+\omega)}. \tag{14}$$

In order to find the macroscopic field of the scattered wave, we write the analog of Eq. (1) at frequency $\omega'$, taking into account that Eq. (12) is now correct and that there is no primary field at this frequency:

$$E_i^{\mathrm{mic}}(\mathbf{R_a},\omega')=(2\pi^2)^{-1}\sum_b\int d^3qQ_{ij}(\mathbf{q},\omega')\alpha_{js}(\omega')E_s^{\mathrm{mic}}\times(\mathbf{R_b},\omega')\exp[i\mathbf{q}\cdot(\mathbf{R_a}-\mathbf{R_b})]+(2\pi^2)^{-1}\int d^3qQ_{ij}(\mathbf{q},\omega')\beta_{js}(\omega,\omega')\times E_s^{\mathrm{mic}}(0,\omega)\exp[i\mathbf{q}\cdot\mathbf{R_a}]. \tag{15}$$

The second term reflects the presence of the dipole moment induced at the origin of coordinates during the Raman scattering. Averaging over the position of the $N-1\approx N$ molecules, we get for the Fourier transform of the local field

$$E_i^{\mathrm{loc}}(\mathbf{q},\omega')=4\pi n_0[Q_{ij}(\mathbf{q},\omega')-T_{ij}(\mathbf{q},\omega')]\{\alpha_{js}(\omega')\times E_s^{\mathrm{loc}}(\mathbf{q},\omega')+\beta_{js}(\omega,\omega')E_s^{\mathrm{loc}}(\mathbf{R}=0,\omega)\}. \tag{16}$$

Taking Eq. (4) into account, we can write

$$E_j^{\mathrm{loc}}(\mathbf{R}=0,\omega)=\frac{1}{n_0}\alpha_{jk}^{-1}(\omega)P_k(\mathbf{R}=0,\omega)=\frac{1}{n_0}\alpha_{jk}^{-1}(\omega)\chi_{kl}(\omega)F_l. \tag{17}$$

The polarization of the substance at frequency $\omega'$ is made up of two components:

$$P_i(\mathbf{q},\omega)=n_0\alpha_{ij}(\omega')E_j^{\mathrm{loc}}(\mathbf{q},\omega')+P_i^0(\mathbf{q},\omega'), \tag{18}$$

where $\mathbf{P}^0$ is the Raman addition to the polarization, which according to Eq. (17) equals

$$P_i^0(\mathbf{q},\omega')=\beta_{ij}(\omega,\omega')\frac{1}{n_0}\alpha_{jk}^{-1}(\omega)\chi_{kl}(\omega)F_l.$$

Equation (6), which is a consequence of Maxwell's equations for a macroscopic field, because there is no primary field at frequency $\omega'$, now takes the form

$$E_i(\mathbf{q},\omega')=4\pi[Q_{ij}(\mathbf{q},\omega')-\delta_{ij}]P_j(\mathbf{q},\omega'). \tag{19}$$

Equations (16), (18), and (19) relate the macroscopic field, the local field, and the polarization to each other. In order to find the macroscopic field of the scattered wave, we eliminate the last two unknowns. Combining Eqs. (16) and (18), we get

$$\{\delta_{ij}-4\pi n_0\alpha_{is}(\omega')[Q_{sj}(\mathbf{q},\omega')-T_{sj}(\mathbf{q},\omega')]\}P_j=P_i^0. \tag{20}$$

Using Eq. (8) to introduce the inverse electric susceptibility tensor, we can rewrite Eq. (20) as

$$n_0\alpha_{ij}(\omega')\{\chi_{js}^{-1}(\omega')+4\pi[\delta_{js}-Q_{js}(\mathbf{q},\omega')]\}P_s=P_i^0. \tag{21}$$

Substitution of Eq. (19) gives the desired expression for the macroscopic field:

$$\{[Q_{ij}(\mathbf{q},\omega')-\delta_{ij}]^{-1}-4\pi\chi_{ij}(\omega')\}E_j = 4\pi\frac{1}{n_0}\alpha_{js}^{-1}(\omega')\chi_{ij}(\omega')P_s^0.$$

Inverting the tensor,

$$[Q_{ij}-\delta_{ij}]^{-1}=\left[\frac{(\omega'/c)^2\delta_{ij}-q_iq_j}{q^2-(\omega'/c)^2}\right]^{-1} =\left(\frac{c}{\omega'}\right)^2\left[\left(q^2-\left(\frac{\omega'}{c}\right)^2\right)\delta_{ij}-q_iq_j\right],$$

and introducing the permittivity of the crystal, we finally have

$$\left[q^2\delta_{ij}\varepsilon_{ij}(\omega')\left(\frac{\omega'}{c}\right)^2-q_iq_j\right]E_j(\mathbf{q},\omega') =4\pi\left(\frac{\omega'}{c}\right)^2\chi_{ij}(\omega')\frac{1}{n_0}\alpha_{jk}^{-1}(\omega')\beta_{ks}(\omega,\omega') \times\frac{1}{n_0}\alpha_{sl}^{-1}(\omega)\chi_{ln}(\omega)F_n. \tag{22}$$

It is easy to see by analyzing this result that it is analogous to the result of calculating the radiation of a point dipole **d** in a crystal, using the usual Maxwell's equations, when the right-hand side of the analogous expression has the form

$$4\pi(\omega'/c)^2d_i.$$

Comparing this with Eq. (22), it is easy to see that a molecule of the crystal during the Raman scattering actually behaves as if it had an effective Raman polarizability of

$$\beta_{ij}^{\text{eff}}(\omega,\omega')=\frac{1}{n_0}\alpha_{sk}^{-1}(\omega')\chi_{is}(\omega')\beta_{kl}(\omega,\omega') \times\frac{1}{n_0}\alpha_{ln}^{-1}(\omega)\chi_{nj}(\omega). \tag{23}$$

We should point out that this is valid in a crystal with arbitrary symmetry. For a uniaxial crystal, Eq. (23) has a simpler form:

$$\beta_{ij}^{\text{eff}}(\omega,\omega')=\beta_{\perp}^{\text{eff}}(\delta_{ij}-e_ie_j)+\beta_{\parallel}^{\text{eff}}e_ie_j,$$

where

$$\beta_{\perp}^{\text{eff}}=\frac{\chi_{\perp}(\omega')\beta_{\perp}(\omega,\omega')\chi_{\perp}(\omega)}{n_0\alpha_{\perp}(\omega')n_0\alpha_{\perp}(\omega)}, \qquad \beta_{\parallel}^{\text{eff}}=\frac{\chi_{\parallel}(\omega')\beta_{\parallel}(\omega,\omega')\chi_{\parallel}(\omega)}{n_0\alpha_{\parallel}(\omega')n_0\alpha_{\parallel}(\omega)}. \tag{24}$$

In an isotropic crystal, in the Lorentz–Lorenz model, Eq. (23) gives

$$\beta_{ij}^{\text{eff}}(\omega,\omega')=\delta_{ij}\frac{\varepsilon(\omega)+2}{3}\frac{\varepsilon(\omega')+2}{3}\beta(\omega,\omega'),$$

which coincides with the results of Refs. 4–7.

## 4. RAMAN SCATTERING IN A UNIAXIAL CRYSTAL

In a uniaxial crystal, when the permittivity has the form of Eq. (11), inverting the tensor on the left-hand side of Eq. (22) gives

$$E_i(\mathbf{q},\omega')=\frac{4\pi(\omega'/c)^2}{q^2-\varepsilon_{\perp}(\omega'/c)^2} \times\left[\delta_{ij}-\frac{A_{ij}(\mathbf{q},\omega')}{\varepsilon_{\perp}q_{\perp}^2+\varepsilon_{\parallel}q_z^2-\varepsilon_{\parallel}\varepsilon_{\perp}(\omega'/c)^2}\right] \times\beta_{js}^{\text{eff}}F_s. \tag{25}$$

Tensor $A_{ij}$ is defined as

$$A_{ij}(\mathbf{q},\omega')=\varepsilon_{\perp}(\varepsilon_{\perp}-\varepsilon_{\parallel})\left(\frac{\omega'}{c}\right)^2e_ie_j+\left(\frac{c}{\omega'}\right)^2 \times\left[q^2-\varepsilon_{\parallel}\left(\frac{\omega'}{c}\right)^2\right]q_iq_j+(\varepsilon_{\parallel}-\varepsilon_{\perp}) \times(\mathbf{q}\cdot\mathbf{e})(e_iq_j+e_jq_i),$$

while $\mathbf{q}_{\perp}$ and $q_z$ are, respectively, the transverse and longitudinal components of the wave vector relative to the optic axis $z$ of the crystal.

The inverse Fourier transformation can be broken up into two stages—integration over the longitudinal and the transverse wave vectors:

$$E_i(\mathbf{r},\omega')=(2\pi)^{-3}\int d^2q_{\perp}\,\exp[i\mathbf{q}_{\perp}\cdot\mathbf{r}]\int dq_zE_i(\mathbf{q}_{\perp},q_z,\omega') \times\exp[iq_zz].$$

The integral over $q_z$ is determined by the poles of $E(\mathbf{q},\omega')$. The two types of poles in Eq. (25) correspond to the two types of waves in a uniaxial crystal—ordinary and extraordinary. The scattered wave consequently is the sum

$$\mathbf{E}(\mathbf{r},\omega')=\mathbf{E}^{\text{or}}(\mathbf{r},\omega')+\mathbf{E}^{\text{ex}}(\mathbf{r},\omega').$$

The ordinary wave is determined by the poles

$$q_z=\pm\sqrt{\varepsilon_{\perp}(\omega'/c)^2-q_{\perp}^2}\pm i0.$$

For positive $z$, the contour of integration can be closed via a half-circle of infinite radius in the upper complex half-plane, and likewise for negative $z$ in the lower half-plane. Taking into account the corresponding contributions gives

$$E_i^{\text{or}}(\mathbf{r},\omega')=(2\pi)^{-1}i\left(\frac{\omega'}{c}\right)^2\int d^2q_{\perp} \times\exp[i\mathbf{q}_{\perp}\cdot\mathbf{r}+z\sqrt{\varepsilon_{\perp}(\omega'/c)^2-q_{\perp}^2}] \times\left[\delta_{ij}-\frac{A_{ij}(z/|z|\sqrt{\varepsilon_{\perp}(\omega'/c)^2-q_{\perp}^2},\mathbf{q}_{\perp},\omega')}{(\varepsilon_{\perp}-\varepsilon_{\parallel})q_{\perp}^2}\right] \times\frac{\beta_{js}^{\text{eff}}F_s}{\sqrt{\varepsilon_{\perp}(\omega'/c)^2-q_{\perp}^2}}. \tag{26}$$

To determine the waves at large distances in the wave zone, we use the stationary-phase approximation. The extremum of the expression in the exponential is found at the point

$$\mathbf{q}_\perp^0=\frac{\mathbf{r}_\perp}{r}\,(\omega'/c)\sqrt{\varepsilon_\perp},$$

which corresponds to $q_z^0=zr^{-1}(\omega'/c)\sqrt{\varepsilon_\perp}$ or $\mathbf{q}^0=\mathbf{r}r^{-1}(\omega'/c)\sqrt{\varepsilon_\perp}$. Use of the well-known formulas of the stationary-phase method in the two-dimensional case gives

$$E_i^{\rm or}(\mathbf{r},\omega')=\frac{1}{r}\exp\left[i\sqrt{\varepsilon_\perp}\,\frac{\omega'}{c}\,r\right]\left(\frac{\omega'}{c}\right)^2\times\left[\delta_{ij}-\frac{A_{ij}(\mathbf{r}r^{-1}\sqrt{\varepsilon_\perp}(\omega'/c),\omega')}{(\varepsilon_\perp-\varepsilon_\parallel)\varepsilon_\perp(\omega'/c)^2(r_\perp/r)^2}\right]\beta_{js}^{\rm eff}F_s. \tag{27}$$

Let the incident wave be polarized in the $xz$ plane. Using $\mathbf{n}$ to denote the unit vector of the $x$ axis, we write

$$\mathbf{F}=F_\parallel\mathbf{e}+F_\perp\mathbf{n},$$

and, consequently,

$$\beta_{js}^{\rm eff}F_s=\beta_\parallel^{\rm eff}F_\parallel e_j+\beta_\perp^{\rm eff}F_\perp n_j.$$

Equation (27) in this case can be simplified:

$$E_i^{\rm or}(\mathbf{r},\omega')=\frac{1}{r}\exp\left[ir\sqrt{\varepsilon_\perp}\,\frac{\omega'}{c}\right]\left(\frac{\omega'}{c}\right)^2\beta_\perp^{\rm eff}F_\perp\times\left(n_i-\frac{xr_i}{r^2-z^2}+e_i\,\frac{xz}{r^2-z^2}\right). \tag{28}$$

The energy flux into an element of solid angle is

$$\frac{dI}{d\Omega}=r^2S=r^2\,\frac{c}{8\pi}\,|\mathbf{E}\cdot\mathbf{H}^*|=\frac{c^2}{8\pi\omega'}\,|\mathbf{E}\times(\mathbf{q}^0\times\mathbf{E}^*)|,$$

where $\mathbf{q}^0$ is the stationary-phase point. Substituting Eq. (28), we get

$$\frac{dI^{\rm or}}{d\Omega}=\frac{c}{8\pi}\left(\frac{\omega'}{c}\right)^4\sqrt{\varepsilon_\perp(\omega')}|\beta_\perp^{\rm eff}(\omega,\omega')F_\perp|^2\,\sin^2\,\varphi, \tag{29}$$

$\varphi$ is the azimuthal angle measured from the $x$ axis.

Similar calculations can be made for the extraordinary waves. In this case, the integral over $q_z$ includes the contribution of the pole

$$q_z=\pm\sqrt{\varepsilon_\perp(\omega'/c)^2-q_\perp^2\varepsilon_\perp/\varepsilon_\parallel}\pm i0.$$

The stationary-phase method makes it possible to integrate over the transverse wave vectors, and we finally have

$$E_i^{\rm ex}(\mathbf{r},\omega')=\exp\left[i\sqrt{\varepsilon_\perp}\,\frac{\omega'}{c}\,\sqrt{z^2+r^2\,\frac{\varepsilon_\parallel}{\varepsilon_\perp}}\right]\times\left[z^2+r^2\,\frac{\varepsilon_\parallel}{\varepsilon_\perp}\right]^{-3/2}\left(\frac{\omega'}{c}\right)^2\left[\beta_\parallel^{\rm eff}F_\parallel-\frac{zx}{r}\,\beta_\perp^{\rm eff}F_\perp\right][e_ir_\perp^2-r_{\perp i}z]. \tag{30}$$

The radiation intensity into an element of solid angle in this case is

$$\frac{dI^{\rm ex}}{d\Omega}=\frac{c}{8\pi}\,\sqrt{\varepsilon_\perp}\left(\frac{\omega'}{c}\right)^4\sin^2\,\theta\times\frac{|\beta_\parallel^{\rm eff}F_\parallel-\beta_\perp^{\rm eff}F_\perp\,\cos\,\varphi\,\cot\,\theta|^2}{[\cos^2\,\theta+(\varepsilon_\parallel/\varepsilon_\perp)\sin^2\,\theta]^{5/2}}. \tag{31}$$

For the values $\theta=\pi$ and $\theta=0$, it should be taken into account that both types of waves in these directions have identical wave vectors and orthogonal polarization. It therefore makes sense to speak only of their total intensity, which is determined by

$$\frac{dI^{\rm tot}}{d\Omega}\,(\theta=0,\pi)=\frac{c}{8\pi}\,\sqrt{\varepsilon_\perp}\left(\frac{\omega'}{c}\right)^4|\beta_\perp^{\rm eff}F_\perp|^2.$$

It can be seen from the resulting formulas that the angular radiation distribution is determined by the components of the effective Raman polarizability tensor and consequently strongly depends on the local-field effects.

## 5. RAMAN SCATTERING CROSS SECTION IN A UNIAXIAL CRYSTAL

The radiation intensity of the two types of waves can be obtained by integrating the fluxes given by Eqs. (29) and (31) over angles:

$$I^{\rm or}=\frac{c}{4}\,\sqrt{\varepsilon_\perp}\left(\frac{\omega'}{c}\right)^4|\beta_\perp^{\rm eff}F_\perp|^2,$$

$$I^{\rm ex}=\frac{c}{12}\,\sqrt{\varepsilon_\perp}\left(\frac{\omega'}{c}\right)^4\left(\frac{\varepsilon_\parallel}{\varepsilon_\perp}\right)^2\left[\left|\beta_\perp^{\rm eff}F_\perp\right|^2\frac{\varepsilon_\perp}{\varepsilon_\parallel}+4\left|\beta_\parallel^{\rm eff}F_\parallel\right|^2\right]. \tag{32}$$

The total radiation intensity of all the waves is given by

$$I^{\rm tot}=\frac{c}{12}\sqrt{\varepsilon_\perp}\left(\frac{\omega'}{c}\right)^4\left\{\left|\beta_\perp^{\rm eff}F_\perp\right|^2\left(3+\frac{\varepsilon_\parallel}{\varepsilon_\perp}\right)+4\left|\beta_\parallel^{\rm eff}F_\parallel\right|^2\left(\frac{\varepsilon_\parallel}{\varepsilon_\perp}\right)^2\right\}. \tag{33}$$

In writing the scattering cross sections, it is necessary to distinguish two cases, corresponding to different types of incident wave. If the incident wave is ordinary, it is always polarized perpendicular to the optical axis; i.e., then $F_\parallel=0$. The energy flux in the incident wave in this case,

$$\mathbf{S}_{\rm or}=\frac{c^2}{8\pi}\,\frac{1}{\omega'}\,\mathbf{q}|\mathbf{F}|^2=\frac{c}{8\pi}\,\sqrt{\varepsilon_\perp}|\mathbf{F}|^2\,\frac{\mathbf{q}}{q},$$

is independent of the propagation direction of the wave. The Raman scattering cross section of the ordinary wave is accordingly

$$\sigma_{\rm or}=\frac{I^{\rm tot}}{S_{\rm or}}=\frac{2\pi}{3}\left(\frac{\omega'}{c}\right)^4|\beta_\perp^{\rm eff}|^2\left(3+\frac{\varepsilon_\parallel(\omega')}{\varepsilon_\perp(\omega')}\right)\sqrt{\frac{\varepsilon_\perp(\omega')}{\varepsilon_\perp(\omega)}}. \tag{34}$$

We should point out that the label showing the type of wave is written below in order to emphasize that one is dealing with a type of incident wave and not a scattered wave, as it was above.

In describing the incident extraordinary wave, it is convenient to introduce the ray vector **s**, defined by

$$\mathbf{F}=\mathbf{H}\times\mathbf{s}, \quad \mathbf{H}=\mathbf{s}\times\mathbf{D}.$$

The energy flux

$$\mathbf{S}=(c/8\pi)|\mathbf{H}|^2\mathbf{s}$$

is directed along the ray vector. Introducing the angle $\gamma$ between the ray vector and the optical axis and also recalling that the vectors **s**, **q**, **F**, and **D** lie in the same plane, we can write

$$\mathbf{s}=s(\mathbf{e}\cos\gamma+\mathbf{n}\sin\gamma),$$

$$\mathbf{F}=F(-\mathbf{e}\sin\gamma+\mathbf{n}\cos\gamma).$$

Then

$$|\mathbf{H}|^2=|\mathbf{s}\times\mathbf{D}|^2=|s_\parallel D_\perp-s_\perp D_\parallel|^2$$
$$=s^2(\varepsilon_\parallel\sin^2\gamma+\varepsilon_\perp\cos^2\gamma)^2|\mathbf{F}|^2.$$

The magnitude of the ray vector depends on its direction in a known way:

$$s^2=(\varepsilon_\perp\cos^2\gamma+\varepsilon_\parallel\sin^2\gamma)^{-1};$$

therefore, finally, the energy flux in the extraordinary wave equals

$$S_{ex}=\frac{c}{8\pi}|\mathbf{F}|^2(\varepsilon_\perp(\omega)\cos^2\gamma+\varepsilon_\parallel(\omega)\sin^2\gamma)^{1/2}.$$

The Raman scattering cross section of the extraordinary wave is accordingly

$$\sigma_{ex}=\frac{2\pi}{3}\left(\frac{\omega'}{c}\right)^4\sqrt{\frac{\varepsilon_\perp(\omega')}{\varepsilon_\perp(\omega)}}\frac{|\beta_\perp^{eff}|^2[3+\varepsilon_\parallel(\omega')/\varepsilon_\perp(\omega')]\cos^2\gamma+4[\varepsilon_\parallel(\omega')/\varepsilon_\perp(\omega')]^2|\beta_\parallel^{eff}|^2\sin^2\gamma}{[\cos^2\gamma+(\varepsilon_\parallel(\omega)/\varepsilon_\perp(\omega))\sin^2\gamma]^{1/2}}. \tag{35}$$

The frequently used extinction coefficient $h$, which characterizes the energy loss of the incident wave per unit length of crystal during scattering, is obtained by simply multiplying the cross sections by the number of molecules in unit volume of the crystal:

$$h=\sigma n_0.$$

As can be seen from Eqs. (34) and (35), the scattered energy depends on the propagation direction only for the extraordinary waves.

### 6. DISCUSSION OF THE RESULTS

To estimate the effect of the local field, let us consider the case in which the frequencies of the incident and scattered waves are close, so that it is possible to assume $\varepsilon_{ij}(\omega')=\varepsilon_{ij}(\omega)$. We assume that the anisotropy of the crystal caused by anisotropy of the molecules in an isotropic lattice is weak. For simplicity, we also assume the intrinsic Raman polarizability of the molecules to be isotropic. Then, in accordance with Eq. (11), we can write

$$\varepsilon_\parallel-\varepsilon_\perp=4\pi n_0\delta\left[\frac{\varepsilon+2}{3}\right]^2\ll\varepsilon,$$

where

$$\delta=\alpha_\parallel-\alpha_\perp\ll\alpha.$$

The effective Raman polarizability tensor takes the form

$$\beta_\perp^{eff}=\beta^{eff}=\beta\left[\frac{\varepsilon+2}{3}\right]^2,$$
$$\beta_\parallel^{eff}=\beta^{eff}+2\beta^{eff}\frac{\varepsilon_\parallel-\varepsilon_\perp}{\varepsilon+2}. \tag{36}$$

Expanding Eq. (35) to terms of first order in the anisotropy, we can obtain in this case the scattering cross section of the extraordinary wave:

$$\sigma_{ex}=\frac{8\pi}{3}\left(\frac{\omega'}{c}\right)^4|\beta^{eff}|^2+\frac{2\pi}{3}\left(\frac{\omega'}{c}\right)^4|\beta^{eff}|^2\frac{\varepsilon_\parallel-\varepsilon_\perp}{\varepsilon}$$
$$\times\left[1+\frac{21\varepsilon+10}{\varepsilon+2}\sin^2\gamma\right]. \tag{37}$$

For comparison, we show the analogous formula calculated without including the local field:

$$\sigma_{ex}=\frac{8\pi}{3}\left(\frac{\omega'}{c}\right)^4|\beta|^2+\frac{2\pi}{3}\left(\frac{\omega'}{c}\right)^4|\beta|^2\frac{\varepsilon_\parallel-\varepsilon_\perp}{\varepsilon}$$
$$\times[1+5\sin^2\gamma].$$

Comparing it with Eq. (37), we can see that, besides the common additional factor $[(\varepsilon+2)/3]^4$, taking into account the local field in a uniaxial crystal results in a relative increase by a factor of 3 or 4 of that part of the cross section that depends on the direction of the incident wave.

Thus, taking into account the effect of the local field on the Raman scattering in a uniaxial crystal reduces to introducing the effective Raman polarizability of the molecules. In this case, the latter is determined both by the intrinsic properties of the molecules and by the macroscopic properties of the crystal at the frequencies of the incident and scattered waves.

This is reflected both in the scattering intensity and in the angular distribution of the scattered waves. In the case of an extraordinary incident wave, when the cross section depends on its propagation direction, it is the effective polarizability that determines this dependence; i.e., it is essential to allow for the local field in this case.

This work was carried out with the partial support of the Russian Fund for Fundamental Research (Project 95-02-06059) and the Ministry of Science (Project 96-7-3).

[1] A. N. Botvich, V. G. Podoprigora, and V. F. Shabanov, *Raman Scattering in Molecular Crystals* (Nauka, Novosibirsk, 1968).
[2] M. I. Ryazanov, Zh. Éksp. Teor. Fiz. **103**, 1840 (1993) [JETP **76**, 910 (1993)].
[3] M. I. Ryazanov, Zh. Éksp. Teor. Fiz. **108**, 1778 (1995) [JETP **81**, 974 (1995)].
[4] V. S. Libov and N. G. Bakshiev, Opt. Spektrosk. **31**, 48 (1971) [Opt. Spectrosc. (USSR) **31**, 24 (1971)].
[5] V. M. Sidorenko, V. S. Libov and N. G. Bakshiev, Opt. Spektrosk. **35**, 270 (1973) [Opt. Spectrosc. (USSR) **35**, 158 (1973)].
[6] V. M. Sidorenko, V. S. Libov and N. G. Bakshiev, Opt. Spektrosk. **37**, 680 (1974) [Opt. Spectrosc. (USSR) **37**, 385 (1974)].
[7] V. M. Sidorenko, V. S. Libov and N. G. Bakshiev, Opt. Spektrosk. **41**, 699 (1976) [Opt. Spectrosc. **41**, 417 (1976)].

Translated by W. J. Manthey